

Análise dos dados faltantes do Sistema de Informações sobre Nascidos Vivos (2022)

Mariana Peres Carneiro

Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR), Universidade Federal de Minas Gerais.

Raquel Zanatta Coutinho

Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR), Universidade Federal de Minas Gerais.

RESUMO

Os dados provenientes do Sistema de Informações sobre Nascidos Vivos (SINASC) permitem o cálculo de indicadores de saúde estratificados por região geográfica, características dos atendimentos e diversas características sociodemográficas da mãe, sendo uma ferramenta de relevância para a saúde materno-infantil. Assim, é importante que a qualidade desses dados seja avaliada constantemente. O objetivo deste estudo foi avaliar, para cada extração temporal de 2022, a proporção de dados faltantes por Unidade Federativa de acordo com a idade, raça/cor e escolaridade materna. Logo, empregou-se a metodologia descritiva, transversal, analisando os microdados preliminares e final do SINASC. Com este estudo, foram identificadas as Unidades Federativas e o perfil demográfico materno que apresentaram as maiores proporções de dados faltantes, proporcionando a oportunidade de focar investimentos em melhorias de preenchimento nas regiões e nos públicos que demandam mais atenção.

Palavras-chave: Sistema de Informações sobre Nascidos Vivos; Saúde Materno-Infantil; Dados faltantes; Qualidade dos dados.

Área temática: Demografia (População e saúde)

Financiamento: Não houve financiamento

INTRODUÇÃO

A disponibilidade de informações completas, oportunas e de qualidade a respeito da saúde materno-infantil é uma necessidade em todos os países e uma premissa para a criação, execução e avaliação de políticas públicas para o desenvolvimento humano. No Brasil, antes da década de 1990, os dados sobre nascidos vivos, como o número de nascimentos ocorridos no país, eram provenientes do Sistema de Registro Civil, no entanto, estes apenas representavam os nascimentos registrados em cartório, evidenciando assim os primeiros problemas no que tange a cobertura e qualidade dos dados. Dessa forma, a partir da década de 1990, o Ministério da Saúde (MS) criou o Sistema de Informações sobre Nascidos Vivos (SINASC) com o intuito de coletar informações abrangentes sobre os nascimentos ocorridos no Brasil (PEDRAZA, 2012; AGRANONIK; JUNG, 2019). O SINASC é um dos subsistemas que compõem os Sistemas de Informações em Saúde (SIS), que são fundamentais para subsidiar a tomada de decisão nas três esferas de governo em relação às condições de saúde da população.

Além disso, ainda em 1990, foi implementada a Declaração de Nascido Vivo (DNV), tornando assim o documento base do SINASC ao padronizar as informações sobre a saúde materno-infantil incluindo dados da mãe, do pré-natal, do recém-nascido e do parto (PEDRAZA, 2012; BRASIL, 2022). A DNV é emitida em três vias, ficando a Via Branca sob responsabilidade do estabelecimento de saúde encaminhá-la para a Secretaria Municipal de Saúde para registro do SINASC, a Via Amarela, que é entregue ao responsável do recém-nascido para que ocorra o registro do nascido vivo no cartório e obtenha a certidão de nascimento, e, por fim, a Via Rosa, que é arquivada no prontuário do paciente no estabelecimento de saúde. Dessa forma, é de responsabilidade das esferas municipais e estaduais o provimento dos dados provenientes da via branca da DNV para o SINASC (PEDRAZA, 2012). Informações estas que dizem respeito, por exemplo, ao peso ao nascer, às características da gestação, como os grupos de Robson, ao índice de Apgar no 1º e no 5º minuto, à presença de anomalias, à duração da gestação, ao tipo de parto, ao número de consultas de pré-natal, à adequação do pré-natal, além de informações demográficas importantes da mãe, como a sua idade, sua raça/cor, seu local de residência, seu grau de instrução, seu estado civil, entre outras informações que possibilitam a análise e acompanhamento em estudos epidemiológicos, demográficos, médicos e sociais (MELLO JORGE; LAURENTI; GOTLIEB, 2007).

Nesse sentido, o SINASC é uma ferramenta de relevância no monitoramento da saúde materno-infantil, pois além da contagem dos nascimentos, permite traçar o perfil epidemiológico da população, planejar e avaliar as ações e políticas públicas, executar o diagnóstico situacional do parto, do nascimento, da mãe e dos recém-nascidos, possibilitando assim conhecer os grupos que necessitam de maior atenção por parte das três esferas de governo (PEDRAZA, 2012; GIRODO et al., 2015). Todavia, em território brasileiro, embora existam diretrizes para o tempo até o lançamento do registro no sistema (BRASIL, 2009) não há um plano pré-estabelecido pelo Ministério da Saúde sobre as formas de monitoramento da qualidade das informações, logo, a qualidade dos registros das informações do SINASC é um tópico que vem sendo estudado por diversos autores e que ainda é considerada uma das limitações mais relevantes do sistema (ROMERO; CUNHA, 2006; GIRODO et al., 2015; STEVANATO; GAÍVA; MATHIAS, 2017; AGRANONIK; JUNG, 2019; MARQUES; SIMÃO; SOUZA, 2019).

No que tange à sua qualidade, pesquisadores têm buscado avaliá-la a partir de dimensões como grau de acessibilidade, clareza metodológica, cobertura, completude, confiabilidade,

consistência e oportunidade dos dados (LIMA *et al.*, 2009; SILVA *et al.*, 2014; AGRANONIK; JUNG, 2019). O objetivo deste artigo foi avaliar, para cada extração do SINASC de 2022 a proporção de dados faltantes (*missing*) por UF, no que tange às características demográficas da mãe como idade, raça-cor e escolaridade a fim de verificar os fatores associados ao mau preenchimento.

METODOLOGIA

Foi empregada a metodologia descritiva, transversal, no qual foram avaliados os registros de nascidos vivos no Brasil por Unidade Federativa provenientes do banco de dados do SINASC no ano de 2022. Como se trata de uma pesquisa que utilizou dados secundários, não foi necessária a submissão à Comissão Nacional de Ética em Pesquisa, conforme a Resolução nº 510, de 7 de abril de 2016 (BRASIL, 2016).

Dados

Foram analisados os microdados do banco de dados preliminar do SINASC de 2022, extraídos do site do Departamento de Análise Epidemiológica e Vigilância de Doenças Não-Transmissíveis da Secretaria de Vigilância em Saúde do MS (BRASIL, 2023) e os microdados finais de 2022 extraídos do Departamento de Informática do Sistema Único de Saúde. A ordem e datas das extrações utilizadas neste estudo estão dispostas no Quadro 1. Parte-se do pressuposto que haveria melhoria no preenchimento na extração final. As variáveis utilizadas no presente estudo foram: UF de nascimento (26 estados e Distrito Federal), faixa etária da mãe por grupos etários quinquenais (10 a 14 anos, 15 a 19 anos, 20 a 24 anos, 25 a 29 anos, 30 a 34 anos, 35 a 39 anos, 40 a 44 anos, 45 a 49 anos, 50 anos ou mais, ignorado), grau de instrução da mãe (0 a 7 anos de estudo, 8 a 11 anos de estudo, 12 anos ou mais de estudo e ignorado), raça/cor (branca, preta, amarela, parda, indígena, ignorado).

Quadro 1 - Ordem e datas das extrações. SINASC, 2022.

Data da extração/disponibilização	Ordem e data da extração		
	1ª (preliminar)	2ª (preliminar)	3ª (consolidada)
Data da extração	março/2023	setembro/2023	janeiro/2024
Data da disponibilização	fevereiro/2023	agosto/2023	janeiro/2024

Nota: A data da extração se refere à data do *download* das bases de dados, enquanto a data da disponibilização se refere a data na qual o Ministério da Saúde disponibilizou as mesmas para *download*.

Métodos

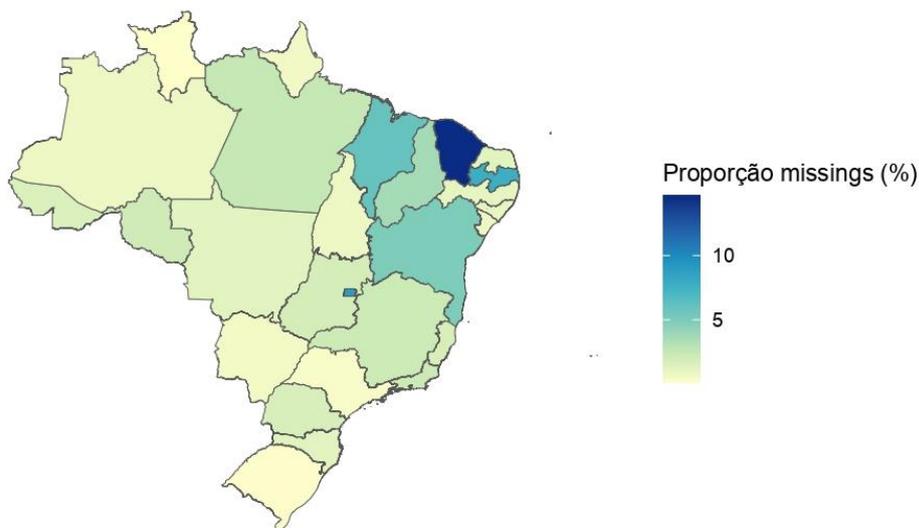
Para alcançar o objetivo proposto, foi calculada a proporção de dados faltantes (*missing*¹) por UF e variável demográfica materna, separadamente para cada extração de 2022. No numerador foi usado o número total de *missing* e no denominador utilizou-se o número total de registros.

¹ Missing é a tradução literal de “dados faltantes” e será utilizada de forma preferencial por uma questão de espaço.

RESULTADOS

A partir da análise dos *missing*, é possível inferir que a qualidade dos dados melhora de uma extração para outra, visto que a quantidade de dados faltantes diminui, no entanto, a presença desses é inaceitável do ponto de vista epidemiológico. Além disso, foi observado que o Distrito Federal e algumas UFs da macrorregião Nordeste (Ceará, Pernambuco, Maranhão e Bahia) tendem a apresentar as maiores proporções de *missing* sob a óptica de raça/cor (Figura 1 e Tabela 1), enquanto as maiores proporções de *missing* de escolaridade foram observadas em Roraima, Pará, Ceará e Bahia (Figuras 2 e Tabela 2), o que evidencia um grave problema de preenchimento dos dados nas macrorregiões Norte e Nordeste. As proporções de *missing* de raça/cor são maiores do que de escolaridade, ou seja, raça/cor é uma variável com pior preenchimento. Os dados faltantes de faixa etária da mãe foram analisados, no entanto, não foram apresentados por representarem apenas 0,001% da amostra.

Figura 1 - Proporção de *missing* da variável raça/cor na terceira extração de 2022



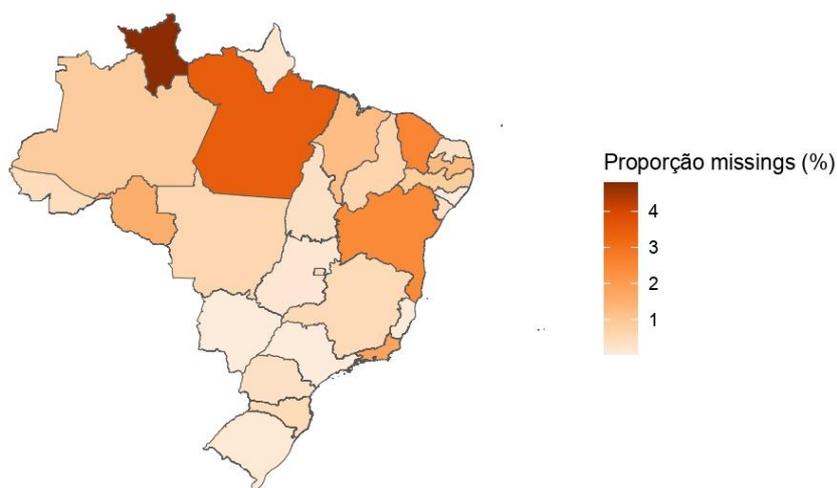
Fonte: SINASC

Tabela 1 - Proporção de *missing* da variável raça/cor materna nas extrações de 2022

UF	Proporção de <i>missing</i> de cada extração (%)		
	1ª extração	2ª extração	3ª extração
Acre	1,76	1,62	1,66
Alagoas	1,64	1,28	1,14
Amapá	0,52	0,58	0,62
Amazonas	0,89	0,88	0,82
Bahia	5,84	5,79	4,97
Ceará	14,76	14,73	14,66
Distrito Federal	9,12	8,83	8,98
Espírito Santo	1,81	1,77	1,78
Goiás	0,93	0,88	1,94
Maranhão	6,01	6,16	6,05
Mato Grosso	1,53	1,52	1,31
Mato Grosso do Sul	0,65	0,60	0,60
Minas Gerais	2,44	2,34	2,32
Pará	2,65	2,63	2,58
Paraíba	7,93	7,74	7,77
Paraná	1,77	1,78	1,76
Pernambuco	1,21	1,18	1,18
Piauí	3,71	3,67	3,70
Rio de Janeiro	2,40	2,29	2,29
Rio Grande do Norte	1,27	1,34	1,31
Rio Grande do Sul	0,54	0,25	0,21
Rondônia	2,08	2,15	2,16
Roraima	0,13	0,11	0,11
Santa Catarina	2,65	1,37	1,34
São Paulo	0,42	0,38	0,35
Sergipe	0,75	0,75	0,75
Tocantins	0,55	0,54	0,85

Fonte: SINASC

Figura 2 - Proporção de *missing* da variável escolaridade na segunda extração de 2022



Fonte: SINASC

Tabela 2 - Proporção de *missing* da variável escolaridade materna nas três extrações de 2022

UF	Proporção de <i>missing</i> de cada extração (%)		
	1ª extração	2ª extração	3ª extração
Acre	0,71	0,61	0,47
Alagoas	0,32	0,04	0,03
Amapá	0,38	0,37	0,21
Amazonas	1,02	0,85	0,91
Bahia	2,84	2,80	2,45
Ceará	2,59	2,54	2,56
Distrito Federal	0,53	0,52	0,51
Espírito Santo	0,23	0,09	0,08
Goiás	0,22	0,13	0,18
Maranhão	1,26	1,26	1,24
Mato Grosso	0,82	0,79	0,61
Mato Grosso do Sul	0,61	0,11	0,07
Minas Gerais	0,65	0,56	0,51
Pará	3,57	3,54	3,48
Paraíba	1,42	1,33	1,33
Paraná	0,52	0,51	0,33
Pernambuco	0,85	0,90	0,88
Piauí	0,76	0,67	0,62
Rio de Janeiro	2,03	1,84	1,83
Rio Grande do Norte	0,60	0,44	0,39
Rio Grande do Sul	0,42	0,18	0,15
Rondônia	1,99	1,95	1,57
Roraima	5,40	4,84	4,79
Santa Catarina	0,61	0,53	0,49
São Paulo	0,18	0,10	0,06
Sergipe	0,36	0,32	0,32
Tocantins	0,35	0,33	0,33

Fonte: SINASC

Por outro lado, foi possível observar algumas reduções na diferença percentual de *missing* nas variáveis de raça/cor e escolaridade (Tabelas 3 e 4) quando se comparam as três extrações. As maiores reduções de *missing* de raça/cor foram observados no Rio Grande do Sul (50,66%; 16,05% e 58,58%), Santa Catarina (47,53% entre a primeira e a segunda extração e 48,35% entre a primeira e a terceira extração), Distrito Federal (25,09% entre a segunda e a terceira extração e 22,98% entre a primeira e a terceira extração) e Alagoas (19,84% entre a primeira e a segunda extração e 28,45% entre a primeira e terceira extração).

Já os *missing* de escolaridade apresentaram maior redução da diferença percentual no estado do Alagoas (87,94% entre a primeira e a segunda extração e 89,36% na primeira e terceira extração), Acre (26,09% entre a segunda e a terceira extração e 32,00% entre a primeira e a terceira extração), Amapá (49,09% entre a segunda e a terceira extração e 47,17% entre a primeira e a terceira extração), Distrito Federal (36,88% entre a segunda e a terceira extração e 23,87% entre a primeira e a terceira extração), Espírito Santo (60,36% entre a primeira e a segunda extração e 61,26% entre a primeira e a terceira extração), Goiás (38,69% entre a primeira e a segunda extração), Mato Grosso (22,49% entre a segunda e a terceira extração e 21,63% entre a primeira e a terceira extração), Mato Grosso do Sul (81,28%; 31,82% e 87,23%), Paraná (35,33% entre a

segunda e a terceira extração e 35,06% entre a primeira e a terceira extração), Rio Grande do Norte (20,98% entre a primeira e a segunda extração e 29,46% entre a primeira e a terceira extração), Rio Grande do Sul (55,56% entre a primeira e a segunda extração; 15,09% entre a segunda e a terceira extração e 62,26% entre a primeira e a terceira extração) e São Paulo (40,76% entre a primeira e a segunda extração; 38,16% entre a segunda e a terceira extração e 63,36% entre a primeira e a terceira extração). Logo, é possível perceber que as maiores reduções ocorreram entre a primeira e a terceira extração. Além disso, é possível observar na Tabela 4 que em algumas UFs houve aumento de dados faltantes, principalmente em Goiás entre a segunda e a terceira extração (53,40%).

Não é possível afirmar, apenas com essa análise, se a redução percentual de dados faltantes entre as extrações se deu pela correção dos dados já existentes e digitados no sistema ou pelo incremento de registros de nascimento que tenham tido melhor preenchimento na fonte.

Tabela 3 - Diferença percentual de dados faltantes de raça/cor entre as extrações de 2022

UF	Diferença percentual entre as extrações (%)		
	1ª e 2ª extrações	2ª e 3ª extrações	1ª e 3ª extrações
Acre	-1,61	-2,04	-3,61
Alagoas	-19,84	-10,75	-28,45
Amapá	20,55	-3,41	16,44
Amazonas	7,29	-5,69	1,19
Bahia	1,82	-12,78	-11,19
Ceará	1,73	-0,65	1,07
Distrito Federal	2,82	-25,09	-22,98
Espírito Santo	4,83	0,99	5,86
Goiás	-4,90	156,26	143,70
Maranhão	5,96	0,39	6,37
Mato Grosso	4,75	-13,72	-9,62
Mato Grosso do Sul	-4,35	0,41	-3,95
Minas Gerais	4,81	-0,26	4,54
Pará	0,98	0,12	1,10
Paraíba	3,41	0,30	3,72
Paraná	2,42	-0,76	1,64
Pernambuco	0,51	0,14	0,65
Piauí	1,62	-4,17	-2,62
Rio de Janeiro	-0,55	-0,15	-0,70
Rio Grande do Norte	14,11	-3,32	10,32
Rio Grande do Sul	-50,66	-16,05	-58,58
Rondônia	5,53	0,75	6,32
Roraima	-12,50	0,00	-12,50
Santa Catarina	-47,53	-1,57	-48,35
São Paulo	-7,11	-8,27	-14,79
Sergipe	0,00	-2,27	-2,27
Tocantins	-2,44	59,17	55,28

Fonte: SINASC

Nota: As proporções negativas representam redução de registros após possível correção de inconsistência feita pelo Ministério da Saúde.

Tabela 4 - Diferença percentual de dados faltantes de escolaridade da mãe das extrações de 2022

UF	Diferença percentual entre as extrações (%)		
	1ª e 2ª extrações	2ª e 3ª extrações	1ª e 3ª extrações
Acre	-8,00	-26,09	-32,00
Alagoas	-87,94	-11,76	-89,36
Amapá	3,77	-49,09	-47,17
Amazonas	-9,04	7,65	-2,07
Bahia	1,25	-10,93	-9,82
Ceará	0,07	0,24	0,31
Distrito Federal	4,12	-26,88	-23,87
Espírito Santo	-60,36	-2,27	-61,26
Goiás	-38,69	53,40	-5,95
Maranhão	3,52	0,83	4,37
Mato Grosso	1,10	-22,49	-21,63
Mato Grosso do Sul	-81,28	-31,82	-87,23
Minas Gerais	-5,76	-7,87	-13,17
Pará	1,04	0,11	1,16
Paraíba	-0,73	0,15	-0,59
Paraná	0,42	-35,33	-35,06
Pernambuco	9,05	-2,26	6,58
Piauí	-8,81	-12,00	-19,76
Rio de Janeiro	-6,00	-0,45	-6,42
Rio Grande do Norte	-20,98	-10,73	-29,46
Rio Grande do Sul	-55,56	-15,09	-62,26
Rondônia	0,00	-19,21	-19,21
Roraima	-3,96	-0,48	-4,42
Santa Catarina	-12,59	-6,03	-17,86
São Paulo	-40,76	-38,16	-63,36
Sergipe	-8,65	-5,26	-13,46
Tocantins	-5,19	2,74	-2,60

Fonte: SINASC

Nota: As proporções negativas representam redução de registros após possível correção de inconsistência feita pelo Ministério da Saúde.

DISCUSSÃO

Este estudo objetivou avaliar, para cada extração do SINASC em 2022, a proporção de dados faltantes (*missing*) por UF, no que tange às características demográficas da mãe, a fim de verificar os fatores associados ao mau preenchimento dos dados. Para isto, utilizou-se os microdados das duas extrações preliminares e da extração final do SINASC do ano de 2022.

A existência de dados faltantes evidencia problemas graves de coleta das informações durante o processo de preenchimento da DNV. Logo, torna-se necessário investir em melhorias no

processo de coleta e registro das informações, bem como fortalecer a capacitação dos profissionais envolvidos nesse processo para que seja possível que a existência dos *missing* se torne insignificante. Para Costa e Frias (2009) e Bonilha *et al.* (2018), a existência de *missing* pode estar relacionado à negligência da importância da DNV pelo profissional responsável pelo preenchimento. Em relação a baixa clareza metodológica do manual de preenchimento do SINASC, citada por Costa e Frias (2009), ao longo do tempo foram publicadas atualizações desse manual, que atualmente está em sua 4ª edição. Apesar das melhorias no esclarecimento acerca dos campos de preenchimento, esse trabalho revela que importantes lacunas ainda existem.

Nesse sentido, descobriu-se que a proporção de *missing* da variável raça/cor da mãe tende a apresentar problemas regionais visto que as maiores proporções foram observadas nas UFs do Nordeste e no Distrito Federal. Por outro lado, no que diz respeito a proporção de *missing* da variável escolaridade materna, descobriu-se que as maiores proporções foram observadas em Santa Catarina, no Espírito Santo, em UFs da macrorregião Nordeste e em algumas UFs da macrorregião Norte.

CONCLUSÃO

Neste estudo foi observado que, apesar da existência do SINASC ser considerada um avanço para a saúde pública do Brasil uma vez que é um instrumento fundamental para a tomada de decisão de gestores de serviços de saúde no que diz respeito à saúde materno-infantil, os dados desse sistema necessitam estar sob constante avaliação no que tange a sua qualidade. A presença de altas proporções de *missing* revela as fragilidades que estão inseridas em todo o processo de coleta, preenchimento, registro, processamento de dados e correção, especialmente na gestão municipal do SINASC, representada pelos estabelecimentos de saúde e Secretarias Municipais de Saúde.

REFERÊNCIAS

AGRANONIK, M.; JUNG, R. O. 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, 2019.

BONILHA, E. A. *et al.* Cobertura, completude e confiabilidade das informações do Sistema de Informações sobre Nascidos Vivos de maternidades da rede pública no município de São Paulo, 2011. **Epidemiologia e Serviços de Saúde**, v. 27, n. 1, p. e201712811, 2018.

BRASIL. Ministério da Saúde. Conselho Nacional de Saúde. **Resolução n° 510, de 7 de abril de 2016**. Brasília, 2016.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise Epidemiológica e Vigilância de Doenças Não Transmissíveis. Coordenação-Geral de Informações e Análises Epidemiológicas. **Sistema de Informações sobre Nascidos Vivos**. 2023.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise Epidemiológica e Vigilância de Doenças Não Transmissíveis. **Declaração de Nascido Vivo: manual de instruções para preenchimento**. 4ª ed. Brasília: Ministério da Saúde, 2022.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. **Portaria n.º 116, de 11 de fevereiro de 2009**. Regulamenta a coleta de dados, fluxo e periodicidade de envio das informações sobre óbitos e nascidos vivos para os Sistemas de Informações em Saúde sob gestão da Secretaria de Vigilância em Saúde. Brasília, Distrito Federal: Ministério da Saúde, 2009.

COSTA, J. M. B. S.; FRIAS, P. G. Avaliação da completitude das variáveis da Declaração de Nascido Vivo de residentes em Pernambuco, Brasil, 1996 a 2005. **Cad. Saúde Pública**, v. 25, n. 3, p. 613-624, 2009.

GIRODO, A. M. *et al.* Cobertura do Sistema de Informações sobre Nascidos Vivos e potenciais fontes de informação em municípios de pequeno porte em Minas Gerais, Brasil. **Rev. Bras. Saúde Matern. Infant.**, v. 15, n. 3, p. 317-324, 2015.

MARQUES, D. H. F.; SIMÃO, A. B.; SOUZA, L. M. Qualidade dos dados de nascimento do estado de Minas Gerais e mesorregiões: uma análise comparativa. **Caminhos da Geografia**, v. 20, n. 72, p. 107-117, 2019.

LIMA, C. R. A. *et al.* Revisão das dimensões de qualidade dos dados e métodos aplicados na avaliação dos sistemas de informação em saúde. **Cad. Saúde Pública**, v. 25, n. 10, p. 2095-2109, 2009.

MELLO JORGE, M. H. P.; LAURENTI, R.; GOTLIEB, S. L. D. Análise da qualidade das estatísticas vitais brasileiras: a experiência da implantação do SIM e do SINASC. **Ciência & Saúde Coletiva**, v. 12, n. 3, p. 643-654, 2007.

PEDRAZA, D. F. Qualidade do Sistema de Informações sobre Nascidos Vivos (Sinasc): análise crítica da literatura. **Ciência & Saúde Coletiva**, v. 17, n. 10, p. 2729-2737, 2012.

ROMERO, D. E.; CUNHA, C. B. Disponibilidade e qualidade da informação epidemiológica e demográfica no Sistema de Informações sobre Nascidos Vivos (SINASC) e no Sistema de Informação de Mortalidade (SIM) para menores de um ano. *In: XV Encontro Nacional de Estudos Populacionais*, 15., 2006, Caxambu-MG.

SILVA, L. P. *et al.* Avaliação da qualidade dos dados do Sistema de Informações sobre Nascidos Vivos e do Sistema de Informações sobre Mortalidade no período neonatal, Espírito Santo, Brasil, de 2007 a 2009. **Ciência & Saúde Coletiva**, v. 19, n. 7, p. 2011-2020, 2014.

STEVANATO, J. M.; GAÍVA, M. A. M.; MATHIAS, T. A. F. Análise da Qualidade do Sistema de Informações sobre Nascidos Vivos. **Cienc. Cuid. Saúde**, v. 16, n. 2, p. 265-274, 2017.