

Concordância e completude dos dados sobre nascidos vivos e óbitos infantis

Agreement and completeness of data on live births and infant deaths
Concordancia y completitud de los datos sobre nacidos vivos y defunciones infantiles

Amanda de Ataídes Romaguera  <https://orcid.org/0000-0001-8617-5618>^{1,2}

Aline Luzia Sampaio Guimarães  <https://orcid.org/0000-0001-5752-8373>³

Conceição Maria de Oliveira  <https://orcid.org/0000-0002-2220-5782>^{3,4}

Mirian Domingos Cardoso  <https://orcid.org/0000-0002-2256-8874>⁴

Cristine Vieira do Bonfim  <https://orcid.org/0000-0002-4495-9673>^{1,5}

Como citar:

Romaguera AA, Guimarães AL, Oliveira CM, Cardoso MD, Bonfim CV. Concordância e completude dos dados sobre nascidos vivos e óbitos infantis. Acta Paul Enferm. 2020;33:e-APE20180309

DOI

<http://dx.doi.org/10.37689/acta-ape/2020A00309>



Descritores

Estatísticas vitais; Sistemas de informação em saúde; Mortalidade infantil; Registro de nascimento; Registros de mortalidade

Keywords

Vital statistics; Health information systems; Infant mortality; Birth registration; Mortality registries

Descriptores

Estadísticas vitales; Sistemas de información en salud; Mortalidad infantil; Registro de nacimiento; Registros de mortalidad

Submetido

31 de Dezembro, 2018

Aceito

26 de Agosto, 2019

Autor correspondente

Cristine Vieira do Bonfim
cristine.bonfim@uol.com.br

Resumo

Objetivo: Avaliar a qualidade dos dados (concordância e completude) dos óbitos infantis no Sistema de Informações de Mortalidade (SIM) e no Sistema de Informações sobre Nascidos Vivos (Sinasc), Recife, Pernambuco, Brasil.

Métodos: Estudo transversal com dados de óbitos infantis capturados no Sinasc e no SIM. Para os óbitos, foi utilizado o período 2013-2016 e para os nascidos vivos, o de 2012-2016. Foi utilizado o *linkage* determinístico. Calculou-se a porcentagem de incompletude de 10 variáveis comuns a ambas as bases, antes e após a vinculação das bases de dados. A concordância foi avaliada pelo índice Kappa para variáveis qualitativas, e pelo coeficiente de correlação intraclasse (ICC do inglês *intraclass correlation coefficient*) para variáveis quantitativas.

Resultados: Foi possível relacionar 96,64% dos óbitos às respectivas declarações de nascidos vivos. Todas as variáveis analisadas foram classificadas como excelentes (menos de 5% de incompletude), antes e após a vinculação das bases de dados. No Sinasc, a maior incompletude ocorreu na variável duração da gestação (1,55%), e no SIM, no número de natimortos (2,89%). A concordância foi classificada como quase perfeita para todas as variáveis qualitativas (Kappa entre 0,8 e 1). Todas as variáveis quantitativas foram classificadas como excelentes (ICC maior que 0,75).

Conclusão: Apesar dos avanços na qualidade do SIM e do Sinasc, ainda houve problemas de completude das variáveis, principalmente no SIM. A vinculação das bases de dados contribuiu para aprimorar as informações para a análise de óbitos infantis pelos serviços de saúde e para pesquisa. O *linkage* uma técnica de fácil acesso e baixo custo operacional, que pode ser incluída na rotina de vigilância da mortalidade infantil para a melhoria contínua das informações.

Abstract

Objective: To assess the quality of data (agreement and completeness) on infant deaths in the Mortality Information System (SIM) and in the Information System on Live Births (Sinasc), Recife, Pernambuco, Brazil.

Methods: Cross-sectional study with data on infant deaths captured in Sinasc and SIM. For the deaths, the period 2013-2016 and the live births of 2012-2016 were used. The deterministic linkage was used. The percentage of incompleteness of 10 variables common to both bases pre- and post-linkage was calculated. The agreement was assessed by the Kappa index for qualitative variables and by the intraclass correlation coefficient (ICC) for the quantitative variables.

¹Fundação Joaquim Nabuco, Recife, PE, Brasil.

²Faculdade de Enfermagem Nossa Senhora das Graças, Universidade de Pernambuco, Recife, PE, Brasil.

³Secretaria de Saúde, Paulista, PE, Brasil.

⁴Secretaria de Saúde, Recife, PE, Brasil.

⁵Programa de Pós-Graduação em Saúde Coletiva, Universidade Federal de Pernambuco, Recife, PE, Brasil.

Conflitos de interesse: nenhum conflito de interesse a declarar.

Results: It was possible to relate 96.64% of the deaths to their respective declaration of live birth. All analyzed variables were classified as excellent (less than 5% incompleteness), pre- and post-linkage. In Sinasc, the greatest incompleteness was in the variable length of pregnancy (1.55%) and in the SIM, the number of stillbirths (2.89%). The agreement was classified as almost perfect for all qualitative variables (Kappa between 0.8 and 1). All quantitative variables were excellent (ICC greater than 0.75).

Conclusion: Despite advances in the quality of SIM and Sinasc, there were still problems of completeness of variables, especially in SIM. The linkage contributed to the improvement of information for the analysis of infant deaths by health services and for research. It is a technique that is easy to access and low operational cost, which can be included in the routine of infant mortality surveillance for the continuous improvement of information.

Resumen

Objetivo: Evaluar la calidad de los datos (concordancia y completitud) de defunciones infantiles en el Sistema de Información de Mortalidad (SIM) y en el Sistema de Información sobre Nacidos Vivos (Sinasc), Recife, estado de Pernambuco, Brasil.

Métodos: Estudio transversal con datos de defunciones infantiles registrados en el Sinasc y en el SIM. Para las defunciones, se utilizó el período 2013-2016 y para los nacidos vivos, 2012-2016. Fue utilizada la vinculación determinística. Se calculó el porcentaje de incompletitud de 10 variables comunes de ambas bases, antes y después de su vinculación. La concordancia fue evaluada por el índice Kappa en las variables cualitativas y por el coeficiente de correlación intraclass (ICC, por sus siglas en inglés *intraclass correlation coefficient*) en las variables cuantitativas.

Resultados: Fue posible relacionar 96,64% de las defunciones con las respectivas declaraciones de nacidos vivos. Todas las variables analizadas fueron clasificadas como excelentes (menos de 5% de incompletitud), antes y después de la vinculación de las bases de datos. En el Sinasc, la mayor incompletitud ocurrió en la variable duración de la gestación (1,55%) y, en el SIM, en el número de mortinatos (2,89%). La concordancia fue clasificada como casi perfecta en todas las variables cualitativas (Kappa entre 0,8 y 1). Todas las variables cuantitativas fueron clasificadas como excelentes (ICC mayor a 0,75).

Conclusión: A pesar de los avances en la calidad del SIM y del Sinasc, aún hay problemas de completitud de las variables, principalmente en el SIM. La vinculación de las bases de datos contribuyó en la mejora de la información para el análisis de defunciones infantiles por parte de los servicios de salud y para estudios. La vinculación es una técnica de fácil acceso y bajo costo operativo, que puede incluirse en la rutina de la vigilancia de la mortalidad infantil para la mejora continua de la información.

Introdução

A taxa de mortalidade infantil estima o risco de uma criança morrer durante o primeiro ano de vida, reflète as desigualdades e é considerada um dos indicadores mais sensíveis da condição de vida e de saúde de uma população.^(1,2) A definição de taxa de mortalidade infantil consiste no número de mortes por 1.000 nascidos vivos. Sua redução está na agenda global dos países desenvolvidos e em desenvolvimento e constitui um desafio para os sistemas de saúde. O quarto Objetivo de Desenvolvimento do Milênio era reduzir a mortalidade infantil em 75% no período de 1990 a 2015. O Brasil conseguiu cumprir com antecedência, reduzindo de 29,7 mortes por mil nascidos vivos em 2000 para 15,7 mortes por mil nascidos vivos em 2011. A taxa de mortalidade infantil diminuiu 53,1%, de 64,7 mortes por 1.000 nascidos vivos em 1990 para 30,3 em 2016. No Brasil, a taxa de mortalidade infantil em 1990 era de 52,6 mortes por 1.000 nascidos vivos, e diminuiu para 14,6 em 2016, mas o país apresenta ainda uma grande variabilidade nas taxas de mortalidade infantil, que é maior nas regiões Norte e Nordeste.⁽³⁾ Em 2016, a taxa de mortalidade infantil no estado de Pernambuco foi de 13,94 por 1.000 nascidos vivos. No Recife, a taxa de mortalidade infantil é de 12,09 por 1.000 nascidos vivos.⁽⁴⁾ Para monitorar e avaliar

o cumprimento das metas de redução da mortalidade infantil, é essencial que os sistemas de informações estatísticas vitais apresentem dados confiáveis.^(5,6)

As principais informações registradas pelo setor de saúde sobre as ocorrências e circunstâncias de nascimentos e mortes devem ser usadas pelo registro civil para aumentar a notificação e o registro oficial.⁽⁷⁾ O monitoramento dos Objetivos de Desenvolvimento Sustentável (ODS) dependerá da disponibilidade de estatísticas vitais contínuas, detalhadas e específicas. Nesse sentido, dois objetivos dos ODS estão relacionados à melhoria do registro civil e das estatísticas vitais.⁽⁸⁾

Para o monitoramento adequado do progresso na redução da mortalidade infantil, é fundamental ter informações sobre as estatísticas vitais confiáveis e completas.^(5,9) Cada vez mais, sistemas de estatísticas vitais estão sendo reconhecidos como condutores para direitos humanos e programas de saúde e desenvolvimento, especialmente para mulheres e crianças.⁽⁹⁾ Nos países de baixa renda, onde os sistemas de estatísticas vitais são precários, o ônus das mortes maternas e infantis também é alto.⁽⁵⁾ Portanto, eles precisam de estimativas atualizadas e informações de qualidade para monitorar os níveis e tendências da mortalidade materna e infantil.⁽⁶⁾ Essas informações são necessárias para desenvolver os indicadores de saúde e influenciam o planeja-

mento de políticas, o financiamento de pesquisas e programas, e as medidas de assistência à saúde.⁽¹⁰⁾

Sistemas de estatísticas vitais com funcionamento adequado beneficiam a formulação de políticas públicas e a saúde da população.⁽¹¹⁾ A manutenção desses sistemas é necessária para garantir a confiabilidade dos dados, já que eles fornecem informações sobre nascimentos e mortes; e para subsidiar ações que melhorem a saúde pública.⁽¹²⁾

No Brasil, em 1975, o Ministério da Saúde implantou o Sistema de Informações sobre Mortalidade (SIM) com o objetivo de incorporar a sistematização dos dados sobre óbitos no cotidiano das secretarias de saúde.⁽¹³⁾ O sistema é alimentado pela Declaração de Óbito (DO), que possui informações essenciais para a vigilância epidemiológica.⁽¹⁴⁾ Além disso, o SIM facilita a investigação de fatos sobre óbitos e oferece subsídios para investigar óbitos infantis, através de dados, como: a causa básica da morte.⁽¹⁵⁾

Outra fonte importante de informação para o monitoramento da mortalidade infantil é o Sistema de Informações sobre Nascidos Vivos (Sinasc), que é alimentado pela Declaração de Nascido Vivo (DNV) e possui dados sobre a mãe e o recém-nascido, desde o pré-natal até o nascimento.⁽¹⁶⁾ Implementado em 1990 pelo Ministério da Saúde, o Sinasc foi criado para eliminar sub-registros de nascimentos e a necessidade de coletar dados sobre a saúde de recém-nascidos e características maternas.⁽¹⁷⁾

O Ministério da Saúde fez investimentos maciços com o objetivo de aperfeiçoar a qualidade do SIM e do Sinasc.⁽¹⁸⁾ Apesar das diferenças existentes entre as regiões do país, não há dúvida sobre a melhoria na cobertura, consistência e completude das informações nesses sistemas.⁽¹⁴⁾ Alguns aspectos precisam ser investigados para medir a qualidade das estatísticas vitais, como completude e concordância de variáveis.⁽¹⁹⁾ O *linkage* de base de dados possibilita o aumento de informações e identificação de possíveis erros capazes de modificar indicadores de saúde, como preenchimento incorreto dos campos nos formulários.^(20,21)

Um conjunto de estratégias pode ser utilizado para a adequação das estatísticas vitais, em especial: a incorporação da busca ativa de eventos na rotina dos municípios, progresso da vigilância do óbito, autópsia verbal em locais distantes, integração com comitês de

prevenção de morte materna, fetal e infantil e treinamento dos envolvidos na produção de informações.⁽²²⁾

Com o *linkage*, duas ou mais bases de dados com variáveis em comum se relacionam, o que permite a identificação de registros do mesmo indivíduo em duas ou mais fontes de dados.⁽²³⁾ Essa técnica é cada vez mais utilizada em estudos em saúde, pois aprimora a qualidade da informação, tem baixo custo operacional e é de fácil execução.⁽²⁴⁾ O *linkage* entre bancos de dados permite sua estruturação mais consistente e completa, com informações mais confiáveis. Consequentemente, as informações sobre nascidos vivos e óbitos infantis são melhores, contribuindo para o planejamento e desenvolvimento de ações públicas mais assertivas.

O *linkage* das bases de dados de nascimentos e óbitos infantis permite utilizar muitas variáveis sobre as características maternas e infantis para realizar análises mais detalhadas da mortalidade infantil, que são úteis para entender o padrão básico de ocorrência e suas relações. Estatísticas confiáveis e oportunas de nascimentos e óbitos infantis são essenciais para a formulação de políticas, intervenções apropriadas e alocação de recursos para enfrentar os desafios prioritários da saúde materna e infantil. Além disso, o *linkage* entre os bancos do Sinasc e do SIM oferece uma oportunidade para medir as exposições que ocorrem antes do nascimento nos resultados fetais e infantis. O objetivo deste estudo foi avaliar a qualidade dos dados (concordância e completude) sobre óbitos infantis no Sistema de Informações sobre Mortalidade e no Sistema de Informações sobre Nascidos Vivos, Recife, Pernambuco, Brasil.

Métodos

Estudo descritivo, transversal, no qual foram avaliados os registros de óbitos infantis de residentes no Recife (PE) nos bancos de dados do SIM e do Sinasc. A cidade do Recife possui extensão territorial de 219.423 km², totalmente urbana. Administrativamente, está dividida em 94 bairros e oito distritos sanitários. No ano de 2016, a população estimada na cidade era de 1.625.584 habitantes.⁽²⁵⁾

Foram incluídos no estudo os óbitos infantis registrados de 1 de janeiro de 2013 a 31 de dezem-

bro de 2016, e nascidos vivos de 1 de janeiro de 2012 (devido a possibilidade de resgatar todas as DNV de óbitos ocorridos em 2013) a 31 de dezembro de 2016. O programa R Project for Statistical Computing versão 3.5.0 foi utilizado para a vinculação dos bancos de dados do SIM e do Sinasc.

Antes do *linkage*, foram realizadas a padronização e a “limpeza” nos bancos de dados. O termo “recém-nascido de” (RN de) foi excluído do campo, assim como a correção das abreviações existentes para nomes e verificação de espaços duplos. O nome da mãe também foi padronizado. Para evitar dúvidas sobre o par verdadeiro, foram verificadas algumas variáveis que poderiam ser fatores de confusão, como recém-nascidos gêmeos, registros duplicados, erros no preenchimento de campos essenciais como: nome da mãe, sexo, data de nascimento. As informações contidas na DNV foram consideradas padrão-ouro quando havia divergências ou falta de informações na DO.

A técnica de *linkage* determinístico foi aplicada e o número da DNV foi utilizado como variável chave, pois é único para cada nascimento e possui um campo específico para o seu preenchimento na DO. Após o *linkage*, as variáveis “nome da mãe”, “sexo” e “data de nascimento” foram verificadas para confirmar os pares verdadeiros, duvidosos e não-pares. Foi realizada uma revisão manual das DOs não relacionadas ao nascimento, e aquelas duplamente verificadas quanto aos pares foram adicionadas, usando como critério o nome da mãe, sexo e data de nascimento. A data de nascimento é fundamental, principalmente em estudos mais longos, quando há possibilidade de mais de um nascimento no período estudado. A relação manual do óbito com as variáveis dos nascidos vivos permitiu preencher as variáveis ausentes para descrever as características dos óbitos.

O percentual de óbitos vinculados foi calculado após o *linkage*, de acordo com o número de pares corretamente vinculados ao respectivo nascimento. Para analisar a incompletude das variáveis, a porcentagem de campos não preenchidos foi definida pelos campos em branco e os campos com o código 9 (ignorado). A incompletude avalia o número de campos não preenchidos em cada variável (ignorada + em branco) e foi analisada antes e após o *linkage*. Os seguintes critérios foram utilizados para classificar a incompletude: exce-

lente (<5%); bom (5 a 9,9%); regular (10 a 19,9%); ruim (20 a 49,9%) e muito ruim ($\geq 50\%$).⁽²⁶⁾

Por meio da concordância, é verificado se o preenchimento de determinada variável foi feito de forma idêntica nos dois bancos de dados. Para esta análise, as variáveis foram divididas em dois tipos: as qualitativas (escolaridade da mãe, duração da gestação, tipo de gravidez, tipo de parto, sexo da criança, data de nascimento da criança, número de filhos vivos e número de natimortos) foram analisadas pelo índice Kappa; e as variáveis quantitativas (idade da mãe e peso ao nascer) pelo coeficiente de correlação intraclasse (ICC do inglês *intraclass correlation coefficient*). O índice Kappa é um método estatístico para avaliar o nível de concordância ou reprodutibilidade entre dois conjuntos de dados. O ICC é um índice de confiabilidade amplamente utilizado nas análises de confiabilidade teste-reteste, intra-examinador e inter-indicador.

Os parâmetros utilizados como referência para classificar o índice Kappa e o ICC foram: concordância excelente (0,80 a 1,00), substancial (0,60 a 0,79), moderada (0,40 a 0,59), razoável (0,20 a 0,39), ruim (0 a 0,19) e sem concordância (<0). O nível de significância foi de 5%, portanto, os intervalos de confiança foram construídos com 95% de confiança.⁽²⁷⁾ Os cálculos foram realizados no programa R versão 3.5.0.

A pesquisa foi aprovada pelo Comitê de Ética em Pesquisa da Fundação Joaquim Nabuco (CAAE: 90160818.7.0000.5619).

Resultados

Entre 2012 e 2016, houve 114.424 nascidos vivos no Sinasc e entre 2013 e 2016, foram 1.071 óbitos infantis no SIM. Na primeira etapa do *linkage*, foi possível combinar 978 (91,31%) DOs com suas respectivas DNVs. Posteriormente, uma pesquisa manual foi realizada nas DOs não vinculadas e 57 pares (5,32%) foram recuperados. Assim, foram encontrados 1.035 pares perfeitos (DO/DNV), resultando em uma taxa de vinculação de base de dados de 96,64% (Figura 1).

As 10 variáveis estudadas apresentaram um percentual de incompletude inferior a 5% no SIM e no Sinasc, e foram classificadas como excelentes (antes e após o *linkage*). Na análise antes do *linkage*, no banco SIM, a

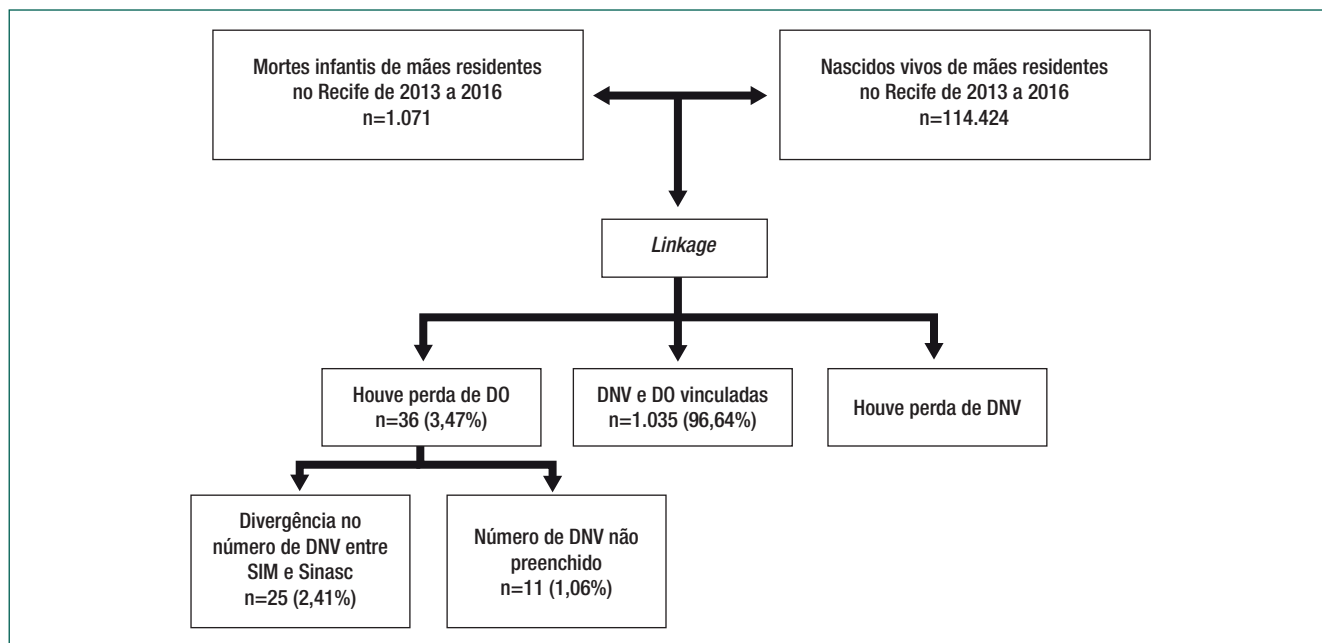


Figura 1. Fluxograma do *linkage* entre o SIM e o Sinasc

variável “sexo” foi a única com 100% de preenchimento; a variável com maior percentual de incompletude foi “número de natimortos” (2,89%), seguida de “duração da gestação” (2,67%). No Sinasc, as variáveis “data de nascimento”, “sexo” e “idade da mãe” estavam 100% completas; a variável “duração da gestação” apresentou a maior incompletude (1,55%), seguida pela “escolaridade da mãe” (1,15%) (Tabela 1).

Após o *linkage*, as variáveis “data de nascimento”, “sexo”, “idade da mãe” e “peso ao nascer” apresentaram 100% de preenchimento; as variáveis “escolaridade da mãe” e “duração da gestação” tiveram o maior percentual de incompletude (0,48% cada) (Tabela 1).

Tabela 1. Incompletude (%) das variáveis de óbitos infantis no Sinasc e SIM antes e após o *linkage*

Variáveis	Antes do <i>linkage</i>				Após o <i>linkage</i>	
	SIM		Sinasc		SIM+Sinasc	
	n(%)	Escore	n(%)	Escore	n(%)	Escore
Data de nascimento	8(0,77)	E	-(-)	E	-(-)	E
Variáveis qualitativas						
Escolaridade da mãe	18(1,73)	E	12(1,15)	E	5(0,48)	E
Duração da gestação	27(2,67)	E	16(1,55)	E	5(0,48)	E
Tipo de gestação	15(1,44)	E	1(0,09)	E	1(0,09)	E
Sexo	-(-)	E	-(-)	E	-(-)	E
Tipo de parto	17(1,64)	E	1(0,09)	E	1(0,09)	E
Variáveis quantitativas						
Idade da mãe	13(1,25)	E	-(-)	E	-(-)	E
Peso ao nascer	13(1,25)	E	-(-)	E	-(-)	E
Número de natimortos	30(2,89)	E	1(0,09)	E	1(0,09)	E
Número de filhos nascidos vivos	22(2,12)	E	1(0,09)	E	1(0,09)	E

*E - Excelente (<5%); B - bom (5 a 9%); RE-regular (10 a 19,9%); Ruim (20 to 49,9%) e Muito ruim (≥ 50%)

Quanto à concordância das variáveis, todas tiveram classificação quase perfeita; a maior concordância foi a variável “data de nascimento” (ICC = 0,988). Todas as variáveis quantitativas apresentaram concordância excelente, e “peso ao nascer” foi a variável com maior concordância (ICC = 0,983), seguida de “idade da mãe” (ICC = 0,981) (Tabela 2).

Tabela 2. Análise de concordância entre as variáveis em comum no SIM e no Sinasc

Variáveis	Coefficiente	L.I.	L.S.
Variáveis qualitativas¹			
Escolaridade da mãe	0,866	0,838	0,893
Duração da gestação	0,813	0,785	0,842
Tipo de gestação	0,951	0,917	0,985
Tipo de parto	0,967	0,951	0,982
Sexo	0,970	0,955	0,984
Data de nascimento	0,988	0,982	0,995
Variáveis quantitativas²			
Número de filhos nascidos vivos	0,875	0,777	0,953
Número de nascidos mortos	0,934	0,873	0,978
Idade da mãe	0,981	0,971	0,989
Peso ao nascer	0,983	0,980	0,986

¹ Coeficiente Kappa; ² Coeficiente de correlação intraclass; L.I. – Limite inferior do intervalo de confiança de 95%; L.S. – Limite superior do intervalo de confiança de 95%

Discussão

O *linkage* das bases de dados do SIM e do Sinasc obteve um percentual de óbitos infantis vinculados de 96,64%. Todas as variáveis estudadas apresentaram percentual de incompletude excelente no SIM e no

Sinasc antes e depois do *linkage*. Todas as variáveis analisadas apresentaram concordância quase perfeita. Isso sugere uma boa qualidade (completude e concordância) das informações do SIM e do Sinasc no Recife.

Estudos recentes mostraram que a taxa de *linkage* das base de dados no Recife permaneceu acima de 95%, o que indica boa qualidade dos sistemas.^(23,24) A comparação do percentual de óbitos infantis relacionados ao Sinasc entre as capitais do Brasil, mostrou que no Recife, mais de 97% dos registros foram vinculados, o que coloca a cidade em posição de destaque entre os maiores percentuais de sucesso de *linkage* entre o SIM e o Sinasc no país.⁽²⁴⁾

Neste estudo, optou-se pelo uso do *linkage* determinístico, considerando a presença do identificador único (número da DNV) encontrado nas duas bases de dados e a qualidade dos dados analisados. Teoricamente, o *linkage* probabilístico fornece uma captura mais verdadeira em relação ao determinístico.⁽²⁸⁾ No entanto, o *linkage* probabilístico envolve um algoritmo mais complicado e demorado que o determinístico.⁽²⁹⁾ Pesquisas comparando métodos probabilísticos e determinísticos descobriram que o probabilístico era mais preciso com dados de pior qualidade, enquanto o determinístico foi igualmente válido e mais veloz em dados de alta qualidade.⁽²⁸⁾ No Brasil, em um estudo que comparou a porcentagem de registros associados entre SIM e Sinasc, foi encontrada predominância do método determinístico em 22 capitais do país.⁽²⁴⁾

Na análise de incompletude, todas as variáveis SIM e Sinasc foram classificadas como excelentes (menos de 5% de incompletude). A avaliação da cobertura, regularidade e qualidade das informações do SIM e do Sinasc demonstrou melhora em todo o país, especialmente nas regiões Norte e Nordeste.⁽²²⁾ O Sinasc avançou consideravelmente, a cobertura é superior a 90% no país e a completude de variáveis acima de 97% para partos hospitalares, isso demonstra a qualidade e a importância das informações desse sistema para as políticas de saúde materna e infantil.⁽³⁰⁾ Da mesma forma, o SIM vem evoluindo significativamente na qualidade das informações, apesar da persistente variabilidade entre regiões e problemas na completude das variáveis.^(24,31)

No SIM, as variáveis com maior percentual de incompletude foram “número de natimortos” e “duração da gestação”, no entanto, foram as que obtiveram maior

ganho de informações após a vinculação da base de dados. A variável “número de natimortos” geralmente apresenta uma alta proporção de incompletude, mas quando preenchida, apresenta alta confiabilidade. A falta de clareza dessa variável nos formulários e nas instruções de preenchimento suscita dúvidas, como: inclusão de abortos ou não, e a contagem de filhos anteriores. Isso se reflete em erros no preenchimento e possivelmente contribui para a sua incompletude.⁽³²⁾ Um estudo que avaliou a incompletude das mortes neonatais, identificou problemas no preenchimento das variáveis relacionadas à história reprodutiva e repetição sistemática de erros, como a inclusão da natimorto atual como um nascimento vivo ou falecido anterior.⁽³³⁾

A duração da gestação é um dos fatores mais importantes para a sobrevivência e subsequente saúde da criança, além de ser essencial para a classificação da morte fetal.⁽³⁴⁾ A variável “duração da gestação” foi apontada em outro estudo como a de pior preenchimento entre as variáveis relacionadas com a gravidez e o parto, o que pode mascarar fatos sobre os cuidados prestados à gestante e ao recém-nascido, deixando de informar variáveis preditivas do risco de morte infantil.⁽³⁵⁾

No Sinasc, as variáveis “duração da gestação” e “escolaridade da mãe” apresentaram o maior percentual de incompletude. Uma recente avaliação do preenchimento da variável “escolaridade da mãe” nas declarações de nascidos vivos nas capitais do Brasil identificou uma tendência de melhora no preenchimento, que foi atribuída ao treinamento de profissionais responsáveis pelo preenchimento e processamento de dados e melhor acesso aos sistemas de informação.⁽³⁶⁾ Especificamente para o Recife, o coeficiente de incompletude melhorou e foi de 33,4 (1996) para 2,8 por mil nascidos vivos (2011), uma redução de 91,6%, demonstrando um claro progresso na completude da variável.⁽³⁶⁾

A incompletude da variável “duração da gestação” encontrada nos dois sistemas pode estar relacionada a alguns fatores, entre eles, a falta de informações do acompanhante da mulher e a falta de informações nos prontuários.⁽³⁴⁾

Após o *linkage*, as variáveis “data de nascimento”, “idade da mãe” e “peso ao nascer” foram completamente preenchidas, o que mostra o ganho de campos anteriormente ignorados/em branco por meio do Sinasc para o SIM. Todas as variáveis da DO e da DNV

são obrigatórias, e não é possível inserir uma DNV no Sinasc sem preencher as variáveis “sexo”, “peso”, “nome” e “idade da mãe”. Esse fato contribuiu para que essas variáveis tivessem 100% de preenchimento. A disponibilidade de informações no prontuário hospitalar ou no cartão da gestante favorece a completude das variáveis na DNV, pois o preenchimento ocorre próximo ao parto e facilita a recuperação de informações desconhecidas pelo acompanhante.⁽³⁶⁾

Na análise de concordância, as variáveis qualitativas apresentaram classificação quase perfeita e as quantitativas, classificação excelente. Alguns fatores contribuíram para a evolução da qualidade das variáveis do SIM e do Sinasc no Recife, a saber: investimentos em treinamento profissional para os envolvidos na produção de informações, treinamento para preenchimento de instrumentos e fortalecimento da vigilância do óbito.^(23,37) Com mais de 10 anos de implementação no Recife, um dos propósitos da estratégia de vigilância do óbito infantil é melhorar a qualidade dos sistemas de informação e tem se mostrado um diferencial na qualificação das estatísticas vitais.⁽³⁷⁾ Uma pesquisa que analisou a concordância da causa básica e da evitabilidade de mortes infantis após a investigação da vigilância do óbito mostrou uma redefinição da maioria das causas básicas, indicando a contribuição dessa estratégia para a melhoria das informações.⁽³⁸⁾

As limitações do estudo consistem em não relacionar todos os registros devido a problemas já identificados por outros estudos, como divergência no nome da mãe entre DNV e DO e ausência de número de DNV na DO.⁽³⁹⁻⁴¹⁾ Outras limitações do estudo envolvem potenciais fatores de confusão que poderiam comprometer a identificação de pares verdadeiros, como duplicidade de registro, gêmeos, erros de ortografia no nome da mãe. No entanto, mesmo com problemas no registro das variáveis, o *linkage* do SIM e do Sinasc fornece informações relevantes para a análise do perfil de mortalidade infantil e pode ser usada para pesquisas e pelos gestores no planejamento de ações de cuidados maternos e infantis. É importante entender a qualidade do *linkage* e o potencial viés que pode ser introduzido nos resultados das análises dos dados vinculados. Neste estudo, a taxa de vinculação de base de dados foi comparável com a literatura anterior. Deve-se enfatizar que a avaliação da qualidade das informações

sobre óbitos infantis é fundamental para a melhoria contínua do SIM e do Sinasc e para propor medidas de redução da mortalidade infantil.

Conclusão

Apesar dos avanços na qualidade do SIM e do Sinasc, ainda houve problemas de completude das variáveis, principalmente no SIM. O *linkage* contribuiu para aprimorar as informações para análise de óbitos infantis pelos serviços de saúde e para pesquisa. É uma técnica de fácil acesso e baixo custo operacional, que pode ser incluída na rotina de vigilância da mortalidade infantil para a melhoria contínua das informações.

Colaborações

Romanguera AA e Bonfim CV contribuíram com a concepção do estudo, análise e interpretação dos dados, redação do manuscrito, revisão crítica de seu conteúdo intelectual e aprovação final da versão a ser publicada. Guimarães ALS, Oliveira CM e Cardoso MD contribuíram com a análise e interpretação dos dados, redação do manuscrito e aprovação final da versão a ser publicada.

Referências

1. Kim D, Saada A. The social determinants of infant mortality and birth outcomes in Western developed nations: a cross-country systematic review. *Int J Environ Res Public Health*. 2013;10(6):2296–335.
2. Son M, An SJ, Kim YJ. Trends of social inequalities in the specific causes of Infant mortality in a nationwide birth cohort in Korea, 1995–2009. *J Korean Med Sci*. 2017;32(9):1401–14.
3. The World Bank [Internet]. Mortality rate, infant (per 1,000 live births). [cited 2019 Aug 12]. Available from: <http://data.worldbank.org/indicator/SP.DYN.IMRT.IN>
4. Pernambuco. Secretaria de Saúde. Óbitos - Pernambuco. In: Sistema de Informação Sobre Mortalidade - Recife - 2016. [citado 2019 Ago 12]. Disponível em: <<http://tabnet.saude.pe.gov.br/cgi-bin/dh?tab/tabsim/obito.def>>
5. Mikkelsen L, Phillips DE, AbouZahr C, Setel PW, de Savigny D, Lozano R, et al. A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *Lancet*. 2015;386(10001):1395–406.
6. Oomman N, Mehl G, Berg M, Silverman R. Modernising vital registration systems: why now? *Lancet*. 2013;381(9875):1336–7.
7. AbouZahr C, Bratschi MW, Muñoz DC, Santon R, Richards N, Riley I, et al. How can we accelerate progress on civil registration and vital statistics? *Bull World Health Organ*. 2018;96(4):226–226A.

8. Mills SL, Abouzahr C, Kim JH, Rassekh BM, Sarpong D. Civil registration and vital statistics (CRVS) for monitoring the sustainable development goals (SDGS). Washington (DC): World Bank Group; 2017.
9. Abouzahr C, de Savigny D, Mikkelsen L, Setel PW, Lozano R, Nichols E, et al. Civil registration and vital statistics: progress in the data revolution for counting and accountability. *Lancet*. 2015;386(10001):1373–85.
10. ACOG Committee Opinion No. ACOG Committee Opinion No. 748: The Importance of Vital Records and Statistics for the Obstetrician-Gynecologist. *Obstet Gynecol*. 2018;132(2):e78–81.
11. Phillips DE, Abouzahr C, Lopez AD, Mikkelsen L, de Savigny D, Lozano R, et al. Are well functioning civil registration and vital statistics systems associated with better health outcomes? *Lancet*. 2015;386(10001):1386–94.
12. Naidoo H, Avenant T, Goga A. Completeness of the Road-to-Health Booklet and Road-to-Health Card: results of cross-sectional surveillance at a provincial tertiary hospital. *South Afr J HIV Med*. 2018;19(1):765.
13. Frias PG, Szwarcwald CL, Morais OL, Leal MD, Cortez-Escalante JJ, Souza PR, et al. [Use of vital data to estimate mortality indicators in Brazil: from the active search for events to the development of methods]. *Cad Saude Publica*. 2017;33(3):e00206015. Portuguese.
14. Caetano SF, Vanderlei CM, Frias PG. [Evaluation of completeness of Instruments for Research on Child Death in the city of Arapiraca, Alagoas]. *Cad Saude Colet*. 2013;3(21):309–17. Portuguese.
15. Ishitani LH, Teixeira RA, Abreu DM, Paixão LM, França EB. [Quality of mortality statistics' information: garbage codes as causes of death in Belo Horizonte, 2011-2013]. *Rev Bras Epidemiol*. 2017;20(20 Suppl 1):34–45. Portuguese.
16. da Silva LP, Moreira CM, Amorim MH, de Castro DS, Zandonade E. [Evaluation of the quality of data in the Live Birth Information System and the Information System on Mortality during the neonatal period in the state of Espírito Santo, Brazil, between 2007 and 2009]. *Cien Saude Colet*. 2014;19(7):2011–20. Portuguese.
17. Pereira, et al. [Evaluation of the Live Birth Information System (SINASC) in the Brazilian State of Pernambuco]. *Rev Bras Saude Mater Infant*. 2013;13(1):39–49. Portuguese.
18. Szwarcwald CL, de Frias PG, Júnior PR, da Silva de Almeida W, Neto OL. Correction of vital statistics based on a proactive search of deaths and live births: evidence from a study of the North and Northeast regions of Brazil. *Popul Health Metr*. 2014;12(1):16.
19. Bonilha EA, Vico ES, Freitas M, Barbuscia DM, Galleguillos TG, Okamura MN, et al. [Coverage, completeness and reliability of the data in the Information System on Live Births in public maternity wards in the municipality in São Paulo, Brazil, 2011]. *Epidemiol Serv Saude*. 2018;27(1):e201712811. Portuguese.
20. Maia LT, Souza WV, Mendes AC. [The contribution of the linkage between the SIM and SINASC to improving information on infant mortality in five Brazilian cities]. *Rev Bras Saude Mater Infant*. 2015;15(1):57–66. Portuguese.
21. Deb-Rinker P, León JA, Gilbert NL, Rouleau J, Andersen AM, Bjarnadóttir RI, et al.; Canadian Perinatal Surveillance System Public Health Agency of Canada. Differences in perinatal and infant mortality in high-income countries: artifacts of birth registration or evidence of true differences? *BMC Pediatr*. 2015;15(1):112.
22. Frias PG, Szwarcwald CL, Lira PI. [Evaluation of information systems on live births and mortality in Brazil in the 2000s]. *Cad Saude Publica*. 2014;30(10):2068–280. Portuguese.
23. Marques LJ, Oliveira CM, Bonfim CV. [Assessing the completeness and agreement of variables of the Information Systems on Live Births and on Mortality in Recife-PE, Brazil, 2010-2012]. *Epidemiol Serv Saude*. 2016;25(4):849–54. Portuguese.
24. Maia LT, Souza WV, Mendes AD, Silva AG. Use of linkage to improve the completeness of the SIM and SINASC in the Brazilian capitals. *Rev Saude Publica*. 2017;51:112.
25. Instituto Brasileiro de Geografia e Estatística (IBGE). Portal on-line 2016. [citado 2019 Ago 12]. Disponível em: <http://cidades.ibge.gov.br/xtras/perfil.php?codmun=261160>
26. Romero DE, Cunha CB. Avaliação da qualidade das variáveis epidemiológicas e demográficas do Sistema de Informações sobre Nascidos Vivos, 2002. *Cad Saude Publica*. 2007;23(3):701–14.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–74.
28. Zhu Y, Matsuyama Y, Ohashi Y, Setoguchi S. When to conduct probabilistic linkage vs. deterministic linkage? A simulation study. *J Biomed Inform*. 2015;56:80–6.
29. Ferrante A, Boyd J. A transparent and transportable methodology for evaluating Data Linkage software. *J Biomed Inform*. 2012;45(1):165–72.
30. Oliveira MM, Andrade SS, Dimech GS, Oliveira JC, Malta DC, Rabello Neto DL, et al. Evaluation of the National Information System on Live Births in Brazil, 2006-2010. *Epidemiol Serv Saude*. 2015;24(4):629–40.
31. Queiroz BL, Freire FH, Gonzaga MR, Lima EE. Completeness of death-count coverage and adult mortality (45q15) for Brazilian states from 1980 to 2010. *Rev Bras Epidemiol*. 2017;20(20 Suppl 01):21–33.
32. Gabriel GP, Chiquetto L, Morcillo AM, Ferreira MC, Bazan IG, Daolio LD, et al. Evaluation of data on live birth certificates from the Information System on Live Births (SINASC) in Campinas, São Paulo, 2009. *Rev Paul Pediatr*. 2014;32(3):183–8.
33. Barreto IC, Vieira MG, Teixeira GP, Fonseca SC. Neonatal death: incompleteness of vital statistics. *Rev Bras Pesq Saude*. 2017;19(2):64–72.
34. Matthews TJ, MacDorman MF, Thoma ME. Infant mortality statistics from the 2013 period linked birth/infant death data set. *Natl Vital Stat Rep*. 2015;64(9):1–30.
35. Ramalho MO, de Frias PG, Vanderlei LC, de Macêdo VC, de Lira PI. [Evaluation of the incompleteness in the filling out of death certificates of children under one year of age in the State of Pernambuco, Brazil, 1999-2011]. *Cien Saude Col*. 2015; 20(9):2891-8. Portuguese.
36. Silvestrin S, Buriol VCS, Silva CH, Goldani MZ. [Assessment of the incompleteness of the maternal schooling variable in Live Birth Certificate databases in Brazilian state capitals, 1996-2013]. *Cad Saude Pub*. 2018; 34(2):1-11. Portuguese.
37. Oliveira CM, Bonfim CV, Guimarães MJ, Frias PG, Antonino VC, Medeiros ZM. Infant mortality surveillance in Recife, Pernambuco, Brazil: operationalization, strengths and limitations. *Epidemiol Serv Saude*. 2017;26(2):413–9.
38. Marques LJ, Pimentel DD, Oliveira CM, Vilela MB, Frias PG, Bonfim CV. Agreement between underlying cause and preventability of infant deaths before and after the investigation in Recife, Pernambuco State, Brazil, 2014. *Epidemiol Serv Saude*. 2018;27(1):e20170557.
39. Cardoso AR, Araújo MA, Andrade RF, Saraceni V, Miranda AE, Dourado MI. Underreporting of Congenital Syphilis as a Cause of Fetal and Infant Deaths in Northeastern Brazil. *PLoS One*. 2016;11(12):e0167255.
40. Lima JC, Mingarelli AM, Segri NJ, Zavala AA, Takano OA. Population-based study on infant mortality. *Cien Saude Colet*. 2017;22(3):931–9.
41. Moura BL, Alencar GP, Silva ZP, Almeida MF. [Hospitalizations due to complications of pregnancy and maternal and perinatal outcomes in a cohort of pregnant women in the Brazilian Unified National Health System in São Paulo, Brazil]. *Cad Saude Publica*. 2018;34(1):e00188016.