Original article

Trend of incompleteness of vital statistics in the neonatal period, State of Rio de Janeiro, Brazil, 1999-2014*

doi: 10.5123/S1679-49742019000200014

Renata Rodrigues Garcia Lino[1] – orcid.org/0000-0002-8353-7860
Sandra Costa Fonseca[1] – orcid.org/0000-0001-5493-494X
Pauline Lorena Kale[2] – orcid.org/0000-0001-5439-9158
Patrícia Viana Guimarães Flores[2] – orcid.org/0000-0001-5074-5113
Rejane Sobrino Pinheiro[2] – orcid.org/0000-0002-3361-3626
Cláudia Medina Coeli[2] – orcid.org/0000-0003-1757-3940

[1]Universidade Federal Fluminense, Instituto de Saúde Coletiva, Niterói, RJ, Brasil
[2]Universidade Federal do Rio de Janeiro, Instituto de Estudos em Saúde Coletiva, Rio de Janeiro, RJ, Brasil

Abstract

Objective: to analyze the incompleteness of variables retrieved from the Live Birth Information System (SINASC) and the Mortality Information System (SIM), in the State of Rio de Janeiro, in the time series 1999 to 2014. **Methods:** this was a time series study of SINASC and SIM; the Romero & Cunha score was used to analyze incompleteness; Joinpoint regression was used to analyze temporal trend; in 2014, deaths were evaluated separately according to investigation. **Results:** regarding SINASC variables, in 1999 marital status (37.40%), number of stillborn children (31.30%), and race/skin color (24.63%) had the highest incompleteness percentages; in 2014, no variable exceeded 5% incompleteness, except occupation; regarding SIM variables, in 1999 five variables exceeded 20%, and missing data for previous stillborn children was over 60%; in 2014, only one variable exceeded 20% incompleteness; investigated deaths had better completeness. **Conclusion:** both systems were found to have improved; the quality of SINASC completeness remained superior to that of SIM.

Keywords: Vital Statistics; Live Birth; Health Information Systems; Time Series Studies.

* This article is derived from the Master's Degree dissertation by Renata Rodrigues Garcia Lino, entitled 'Temporal analysis of information system (SINASC and SIM) completeness: live births and neonatal deaths in the State of Rio de Janeiro, 1999-2014', defended at the Public Health Postgraduate Program, Institute of Public Health, Universidade Federal Fluminense, in February 2018. It is one of the elements of the research project entitled 'Integrated Health Records: application of the chain link model in women's and children's health studies', funded by the National Council for Scientific and Technological Development/ Ministry of Science, Technology, Innovation and Communications (CNPq: 447199/2014-5) and by the Rio de Janeiro State Research Support Foundation (FAPERJ: E-26/203.195/2015). Cláudia Medina Coeli is a CNPq Productivity fellow (305545/2015-9).

Correspondence:
Renata Rodrigues Garcia Lino – Rua João Afonso, No. 11, Apto. 312, Humaitá, Rio de Janeiro, RJ, Brazil. Postcode: 22261-040
E-mail: renatalino2@gmail.com

Introduction

Recording vital events serves to meet both legal requirements and also statistical purposes, as well as enabling research into the health situation.[1,2] In Brazil, the Mortality Information System (SIM) was implanted in 1975, while the Live Birth Information System (SINASC) was implanted in 1990. The aim of both systems was to consolidate information on vital statistics in Brazil as a whole as well as in all its Federative Units.[1] SIM data is provided via Death Certificate (DC) while SINASC data is provided via Live Birth Certificate (LBC).[1]

SINASC and SIM have been evaluated nationally and locally with regard to mother and child health data in relation to several quality indicators, such as coverage, completeness, reliability and timeliness of these systems, as well as consistency and validity of the data they provide.[1-6]

> ***With regard to the SIM system, variable completeness quality is lower than that of SINASC, in particular completeness of sociodemographic variables.***

With regard to completeness, studies of SINASC have shown it to be good to excellent, with few exceptions in the case of some variables.[4,6-9,10-14] The variables identified by these studies as having poorest completeness were: race/skin color, parity, maternal schooling and occupation, liveborn children and stillborn children. It should be noted that more recent studies have found lower levels of LBC incompleteness, mostly below 10%,[10] even for specific at-risk live birth populations.[8] An exception was the state of Acre, where completeness of the majority of variables worsened between 2005 and 2010.[15] National evaluations for the periods 2006-2010 and 1996-2013 found differences between regions of the country and state capital cities.[10,11] The Northeast region, for example, had the highest percentage of incompleteness, mainly for the 'race/skin color' and '5-minute Apgar' variables;[10] the same region also had the poorest LBC incompleteness for 'maternal schooling', although it has improved regarding this variable in recent years; the Southern region had the best performance.[11]

With regard to the SIM system, variable completeness quality is lower than that of SINASC, in particular completeness of sociodemographic variables.[3,12-14,16-19] There is a high proportion of data incompleteness on SIM, above all for the following variables: 'race/skin color', maternal 'schooling', 'liveborn children' and 'stillborn children', maternal 'age' and 'length of pregnancy'. Even birth weight, which is a relevant indicator for women's and children's health, has a regular percentage of incomplete data.[14,17]

In 2011, changes occurred to the way in which some LBC and DC variables are filled in, with the aim of enhancing the information held on both systems, i.e. SINASC and SIM, and also to make them more comparable with the model used by the Brazilian Institute of Geography and Statistics (IBGE).[20,21]

In the case of LBCs,[20] these changes included maternal age, which started to be calculated based on date of birth and no longer by asking the mother her age; and race/skin color, which previously referred to the baby but with effect from 2011 became the mother's self-reported race/skin color. Data on maternal schooling began to include the categories of elementary education I, elementary education II, high school and higher education; as well as school grade, described in years of study. Data on gestational age (GA) changed from clustered categories to the exact number of weeks of pregnancy, as well as requiring a diagnostic approach to GA, either by the date of last menstruation, or by other methods. Another relevant change was made to the number of stillborn children variable, referred to with effect from 2011 as 'fetal losses', by including miscarriages and fetal deaths in a new field on the mother's reproductive history.

In the case of DCs, the schooling variable was modified in the same way as LBC.[21] Gestational age, which previously was stratified in age ranges, was now recorded as the number of weeks; and the field on the number of children mothers had already had also took on the nomenclature of 'fetal losses/miscarriages', thus replacing the former term 'stillborn children'.[21]

Evaluation conducted by the Ministry of Health's General Coordination of Epidemiological Information and Analysis for the period 2000-2011, detected both improvements and challenges for Brazil, taking the year 2011 as its reference.[20,21] In short, there continue to be problems with information quality, capable of both affecting the indicators used to inform public policies[2] and also producing bias in epidemiological studies.[18,19]

With regard to socio-economic variables, poor filling-in of data hinders, above all, identification of inequalities in various outcomes of women's and children's health, such as infant death and access to health services, represented by the number of antenatal appointments and type of child delivery.[3,18,19]

It is also important to remember that investigation of infant and fetal deaths, which has been compulsory in Brazil since 2010,[22] is yet another strategy intended to qualify DC information, improving the notification of underlying cause of death and enabling identification of shortcomings in health care provision, in addition to enabling the adoption of health prevention and promotion measures.

SIM and SINASC decentralization has led to these information systems being closer to the sources that capture event data at local level, as well as enabling better monitoring of data collection. The changes to the fields relating to maternal schooling, age and race/skin color have enabled the collection of more precise, detailed and qualified information.[20, 21]

Although SINASC and SIM coverage in the state of Rio de Janeiro is high, being close to 100%;[20,21] scientific production on the quality of vital statistics held on them has been scarce in the last 20 years.[16]

The objective of this study was to analyze the incompleteness of SINASC and SIM variables in the state of Rio de Janeiro for the 1999-2014 time series.

Methods

This is a time series study on the quality of vital statistics for the state of Rio de Janeiro, with regard to the completeness of data available on the SINASC and SIM databases, retrieved using the Rio de Janeiro State Health Department data tabulator – Tabnet – for each system: http://sistemas.saude.rj.gov.br/tabnet/tabcgi.exe?sinasc/nascido.def and http://sistemas.saude.rj.gov.br/tabnet/deftohtm.exe?sim/infantis.def

We used the definition of quality proposed by Lima et al.,[5] according to whom completeness is translated by the proportion of filled-in fields containing values that are not null. In addition to this definition, we also took Romero & Cunha[3] as a reference for null fields: i.e. unknown or unfilled fields, which are evidence of the 'incompleteness' of an information system. The extent to which an information system field is filled in was stratified by these authors as being: excellent, when incompleteness is less than 5%; good, 5%-10%; regular, 10%-20%; poor, 20%-50%; and very poor, when incompleteness is greater than 50%.[7]

This score continued to be one of the most used in completeness studies published between 2013 and 2017.[7,9,13-17,19] Adaptations were made at the cut-off points of the 'good' and 'regular' categories of classification, to avoid overlapping of incompleteness values (good, 5% <10% and regular, 10% <20%).

The population we studied consisted of all live births and neonatal deaths notified in the state of Rio de Janeiro for the period 1999-2014. The followed variables were selected for analysis on SINASC:

a) Variables relating to the newborn baby
- sex;
- birth weight;
- 5-minute Apgar;
- congenital abnormality

b) Maternal sociodemographic variables
- schooling
- usual occupation
- maternal age
- marital status
- race/skin color

c) Reproductive variables, relating to pregnancy and delivery
- number of liveborn children
- number of stillborn children
- length of pregnancy
- antenatal appointments
- type of pregnancy
- type of delivery

With regard to the 'race/skin color' variable, we also took into consideration data applied to newborn babies in the period 1999-2010, before the change in 2011, with the objective of completing the time series. We opted for evaluating only the 5-minute Apgar variable, since this is the variable that is most related to mortality from neonatal asphyxia.[23]

We analyzed the same variables on SIM, except for 'congenital abnormality', 'antenatal appointments', 'marital status' and 'newborn Apgar', as these do not have a corresponding field on the death certificate (DC).

In order to analyze temporal trends we used the Joinpoint Regression application which fits a series of lines and their joinpoints on a logarithmic scale to demonstrate annual trends. We used the Monte Carlo Permutation Method to test for significance. This method adjusts the best line for each segment. Once these segments have been established, their respective annual percentage changes

(APC) are estimated and tested. When direction is inverted or when different trend patterns are seen at a joinpoint, the periods are analyzed separately. In this situation, the final year of one period coincides with the first year of the next period and, for each variable, these periods can be different in relation to the years analyzed, depending on the moment when change is detected; if there is no change, the period is analyzed as a whole.

Incompleteness percentages for SIM variables for the year 2014 were also compared between investigated and uninvestigated neonatal deaths. Fisher's exact test was used to estimate the differences between the percentages.

This study is part of the project entitled 'Integrated Health Records: application of the chain link model in women's and children's health studies (RIS-RJ)', approved by the Federal University of Rio de Janeiro Institute of Public Health Studies Research Ethics Committee (CEP/IESC/UFRJ): Certification of Submission for Ethical Appraisal (CAAE) No. 07534512.9.0000.5286, dated October 3rd 2012. It should be noted that this article is based on public domain data whereby there is no identification of participants.

Results

Table 1 shows percentage incompleteness of the selected SINASC variables and their temporal trend. In 1999, percentage incompleteness was highest for the 'marital status', 'number of stillborn children' and baby's 'race/skin color' variables which were all above 20%, followed by 'number of liveborn children' (16.04%), 'congenital abnormality' (13.99%) and mother's 'usual occupation' (10.20%). Incompleteness was under 5% for all other variables. In 2014, degree of incompleteness had fallen for all variables, with the exception of 'length of pregnancy' and 'usual occupation'; none of them were in excess of 5% incompleteness, except mother's 'usual occupation' which rose to 14.07%. The biggest changes occurred in relation to 'marital status' which fell abruptly between 1999 and 2001, and also in relation to the reproductive variables 'number of liveborn children' and 'number of stillborn children', which fell constantly and significantly by more than 20% per annum in the period 2010-2014. 'Race/skin color' fell by 50.4% between 1999 e 2000, and from then on fell constantly by 8.5% per annum and reached under 2% incompleteness in 2014. Other variables, notwithstanding the low percentages found in 1999, continued to reduce incompleteness and achieved extremely low levels – under 1% – in 2014. On the other hand, completeness of five variables worsened: usual occupation, type of pregnancy, length of pregnancy, antenatal appointments and 5-minute Apgar.

Table 2 shows percentage incompleteness of the SIM variables for neonatal deaths recorded in the period 1999-2014. In 1999, incompleteness of the 'number of stillborn children' variable was over 60%, followed by maternal 'schooling' (36.6%), 'usual occupation' (29.6%), 'number of liveborn children' (29.3%), maternal 'age' (28.2%) and 'race/skin color' (27.4%). The other variable fields – sex and weight – were between 10% and 20% incomplete; only 'sex' had an excellent completeness score (0.59%). By 2014, incompleteness of almost all variables had reduced in a statistically significant manner, most of which achieved good or excellent scores; the exceptions were 'number of stillborn children' and 'schooling' which achieved regular scores in 2014. The mother's 'usual occupation' (26.26%) variable showed no significant improvement and maintained a poor level of incompleteness of over 20%.

As for temporal trend, reduction in incompleteness predominated in the period studied, for the majority of variables on both systems, although the reduction was greater on SINASC. At times variables were stable, or had different trend patterns and different reduction intensity. On SINASC, 'race/skin color', 'sex' and 'marital status' stood out by having reductions equal to or greater than 50%; on SIM, the biggest reductions occurred in relation to 'race/skin color', 'type of delivery' and 'type of pregnancy'. The exception was 'length of pregnancy' which, following a reduction in incompleteness (-17.8%), increased again between 2004 and 2014 (8.0%). Despite the trend variation seen, percentage incompleteness of this variable was lower in 2014 when compared to 1999.

Considering the two information systems, incomplete data was found to be higher on SIM both in 1999 and in 2014 for the majority of variables. In 2014, with the exception of 'usual occupation', no SINASC variable exceeded 5% incompleteness. However, on SIM in 2014 two variables exceeded 10% incompleteness, namely 'number of stillborn children' and 'schooling', while the 'usual occupation' variable continued to be more than 20% incomplete.

In 2006 only 16% of total neonatal deaths in the state of Rio de Janeiro were investigated, whereas in

Table 1 – Live Birth Information System (SINASC) data incompleteness trend, state of Rio de Janeiro, 1999-2014

Variables	1999 (%) N=268,213	2014 (%) N=233,641	Period[d]	Annual percentage change (95%CI[a])	Trend
Neonatal variables					
Sex	0.37 (excellent)[c]	0.01 (excellent)	1999-2004 2004-2007 2007-2014	2.0 (-14.5;12.3) -66.3[c] (83.3;32.1) 1.1 (-6.6;9.5)	Stable Decrease Stable
Birth weight	0.41 (excellent)	<0.01 (excellent)	1999-2014	-36.9[c] (-42.0;-31.3)	Decrease
5-minute Apgar	2.64 (excellent)	2.07 (excellent)	1999-2010 2010-2014	-2.6[c] (-4.8;-0.5) 12.5[c] (0.4;26.1)	Decrease Increase
Congenital abnormality	13.99 (regular)	2.74 (excellent)	1999-2014	-12.6[c] (-14.9;-12.3)	Decrease
Maternal sociodemographic variables					
Maternal age	0.63 (excellent)	<0.01 (excellent)	1999-2014	-35.4[c] (-40.2;-30.4)	Decrease
Schooling	3.93 (excellent)	1.31 (excellent)	1999-2014	-5.4[c] (-8.3;-2.4)	Decrease
Usual occupation[b]	10.20 (regular)	14.07 (regular)	2001-2014	4.9 (-1.7;12.0)	Stable
Marital status	37.40 (poor)	0.84 (excellent)	1999-2001 2001-2014	-79.2[c] (-91.1;-51.3) -2.5 (-7.5; 1.3)	Decrease Stable
Race/skin color (baby/mother)	24.63 (poor)	1.67 (excellent)	1999-2000 2000-2014	-50.40[c] (-80.5;-26.4) -8.5[c] (-12.1;-4.7)	Decrease Decrease
Reproductive/pregnancy/delivery variables					
Number of liveborn children	16.04 (regular)	3.03 (excellent)	1999-2010 2010-2014	-7.5[c] (-11.0;-3.8) -23.5[c] (-33.6;-11.9)	Decrease Decrease
Number of stillborn children	31.30 (poor)	4.98 (excellent)	1999-2010 2010-2014	-8.5[c] (-12.2;-4.7) -25.0[c] (-35.7;-12.5)	Decrease Decrease
Length of pregnancy	0.90 (excellent)	2.41 (excellent)	1999-2009 2009-2014	-11.2[c] (-19.7;-1.7) 62.8[c] (21.2;118.8)	Decrease Increase
Antenatal appointments	4.48 (excellent)	2.94 (excellent)	1999-2004 2004-2009 2009-2014	-1.4 (-11.3;9.3) -18.4[c] (-29.5;-5.6) 14.0[c] (2.5;26.7)	Stable Decrease Increase
Type of pregnancy	0.14 (excellent)	0.13 excellent)	1999-2010 2010-2014	-6.7[c] (-9.8;-3.4) 25.2[c] (5.6;48.8)	Decrease Increase
Type of delivery	0.19 (excellent)	0.11 (excellent)	1999-2014	-5.0[c] (-8.3;-1.7)	Decrease

a) 95%CI: 5% confidence interval.
b) With effect from 2001.
c) Change calculated using logarithmic regression, according to joinpoints for each period of change; p-value <0.05 (Monte Carlo Permutation Method).
d) In the time series analysis of the joinpoint, when the trend pattern changes, the periods are separated ans the final year of a period coincides with the beginning year of the next period.
Note: Romero & Cunha's classification for incomplete filling in of an information system data field: excellent, under 5% incompleteness; good, 5% <10%; regular, 10% <20%; poor, 20%-50%; and very poor, 50% or over. The cut-off points of the 'good' and 'regular' categories were defined to avoid overlap between the incompleteness values.

2014 almost 60% were investigated (57.6%). In Table 3 it can be seen that in 2014 investigated deaths had lower incompleteness percentages than uninvestigated deaths, with statistically significant differences, except for the 'race/skin color' and 'sex' variables.

Discussion

This study analyzed the temporal trend of SINASC and SIM variable completeness in the state of Rio de Janeiro over a fifteen-year period, encompassing the period of changes to these systems with effect from 2011. It identified improvements on both information systems. This finding reflects a national trend,[20,21] resulting from investment in changes to the systems in keeping with the valuing of epidemiological information in health policies.[24] Decentralization of these information systems has enabled them to get closer to their respective data collection sources, thus facilitating monitoring and increasing data quality.[20,21,24]

Table 2 – Mortality Information System (SIM), neonatal deaths, data incompleteness trend, state of Rio de Janeiro, 1999-2014

Variables	1999 (%) N=4,008	2014 (%) N=2,050	Period[d]	Annual percentage change (95%CI[a])	Trend
Neonatal variables					
Sex	0.59 (excellent)[c]	0.58 (excellent)	1999-2014	-4.0 c (-7.4;-0.5)	Decrease
Race/skin color	27.42 (poor)	9.13 (good)	1999-2005 2005-2010 2010-2014	0.2 (-3.4;3.8) -21.3 c (-26.7;-15.5) 3.4 (-3.8;11.1)	Stable Decrease Stable
Birth weight	11.15 (regular)	5.46 (good)	1999-2003 2003-2007 2007-2014	-15.4 c (-20.0;-10.5) 10.6 c (0.8;21.5) -5.0 c (-7.0;-2.9)	Decrease Increase Decrease
Maternal sociodemographic variables					
Maternal age	28.21 (poor)	6.64 (good)	1999-2004 2004-2007 2007-2014	-18.2 c (-24.7;-11.1) 18.0 (-28.2;94.1) -7.9 c (-12.2;-3.5)	Decrease Stable Decrease
Schooling	36.60 (poor)	14.68 (regular)	1999-2004 2004-2007 2007-2014	-14.5 c (-18.0;-10.8) -16.1 c (-11.2;-51.8) -5.0 c (-7.2;-2.7)	Decrease Decrease Decrease
Usual occupation[b]	29.56 (poor)	26.26 (poor)	2001-2014	1.36 (-1.9;4.7)	Stable
Reproductive/pregnancy/delivery variables					
Number of liveborn children	29.29 (poor)	8.25 (good)	1999-2003 2003-2007 2007-2014	-9.8 c (-14.1;-5.4) 5.8 c (-2.3;14.6) -14.5 c (-16.1;-12.8)	Decrease Stable Decrease
Number of stillborn children	62.62 (very poor)	15.23 (regular)	1999-2004 2004-2007 2007-2014	-10.2 c (-12.9;-7.4) 5.6 c (-11.4;-25.8) -13.9 c (-15.4;-12.4)	Decrease Stable Decrease
Length of pregnancy	16.81 (regular)	8.74 (good)	1999-2004 2004-2014	-17.8 c (-30.8;-2.3) 8.0 c (1.9;14.4)	Decrease Increase
Type of pregnancy	15.11 (regular)	3.85 (excellent)	1999-2001 2001-2014	-36.3 c (-72.7;48.5) -1.0 (-3.6;1.8)	Stable Stable
Type of delivery	14.52 (regular)	4.29 (excellent)	1999-2001 2001-2014	-17.9[c] (-26.7; -8.0) -6.3[c] (-12.2; -0.1)	Decrease Decrease

a) 95%CI: 5% confidence interval.
b) With effect from 2001.
c) Change calculated using logarithmic regression, according to joinpoints for each period of change; p-value <0.05 (Monte Carlo Permutation Method).
d) In the time series analysis of the joinpoint, when the trend pattern changes, the periods are separated ans the final year of a period coincides with the beginning year of the next period.
Note: Romero & Cunha's classification for incomplete filling in of an information system data field: excellent, under 5% incompleteness; good, 5% <10%; regular, 10% <20%; poor, 20%-50%; and very poor, 50% or over. The cut-off points of the 'good' and 'regular' categories were defined to avoid overlap between the incompleteness values.

In 1999, SINASC in the state of Rio de Janeiro had three variables with poor completeness and three with regular completeness; completeness of the other variables was excellent; in 2014, completeness of the majority of the variables showed improvement and, as per Romero & Cunha's[4] classification model, all variables achieved excellence, with the exception of maternal 'usual occupation'.

It is noteworthy that although it had low incompleteness in 2014 (2.4%), data on length of pregnancy worsened, this being a result also found by the report about changes to SINASC:[20] increased incompleteness in 2011, and an increase unknown or blank answers from 0.6% to almost 4%. Another negative finding was the 'antenatal appointments' field which worsened in the period 2009-2014 in the state of Rio de Janeiro, perhaps because of the change in the option from a variable with intervals to a continuous variable.

When comparing data completeness in the state of Rio de Janeiro with that of other Brazilian locations, regions and Federative Units,[9,10,13,14] an increase in completeness can be seen nationwide. The only exception was the state of Acre,[15] where the incompleteness of 11 out of 15 variables analyzed increased between 2005 and 2010. It should be noted that the majority of the studies referred to assessed data prior to 2011 and this makes comparison difficult.

The reproductive variables still need to be improved. In the case of the 'number of stillborn children' variable, the way this field is filled in changed in 2011 to include previous miscarriages. This alteration may explain the unusual increase in the percentage of women having one or more stillbirths in the period studied, from 2010 to 2014 (from 9.5% to 18.2%), this being in disagreement with the falling trend in fetal deaths in Southeast Brazil and in the country as a whole, as found by a study on low birth weight.[25]

Similarly to SINASC data, SIM data completeness with regard to neonatal deaths in the state of Rio de Janeiro showed statistically significant improvement in the variables analyzed. However, there continue to be shortcomings with some variables, in particular maternal 'usual occupation', albeit with poor completeness. The evaluation of SIM completeness for the Brazilian state capitals in 2011 and 2012[19] found that maternal occupation was the most poorly filled in variable: the corresponding field not only had problems related to poor completeness, but also problems of consistency, with variation in classification between different sources, as per Romero & Cunha.[4] These results are in agreement with those of evaluations of other systems, such as the Notifiable Diseases Information System (SINAN).[26] It is probable that the fact of this variable requiring searches on the Brazilian Classification of Occupations, as well as difficulties in interpreting the instructions contained in the manuals, favor poor completeness.[4]

Number of stillborn children was the variable with the highest percentage of incompleteness according to our study and also according to other studies.[14,16,18] In the state of Rio de Janeiro, the rate of this variable remained high in 2014 (15.3%), whereas in Recife in

Table 3 – Mortality Information System (SIM) data percentage incompleteness, investigated and uninvestigated deaths, state of do Rio de Janeiro, 2014

Variables	Investigated deaths % (N=1,180)	Uninvestigated deaths % (N=712)	P value[a]
Neonatal variables			
Sex	0.51 (excellent)	0.56 (excellent)	0.557
Race/skin color	8.64 (good)	9.97 (good)	0.186
Birth weight	2.29 (excellent)	10.53 (regular)	<0.001
Maternal sociodemographic variables			
Maternal age	3.98 (excellent)	10.25 (regular)	<0.001
Schooling	12.10 (regular)	18.80 (regular)	<0.001
Usual occupation	20.90 (poor)	31.18 (poor)	<0.001
Reproductive/pregnancy/delivery variables			
Number of liveborn children	5.08 (good)	13.20 (regular)	<0.001
Number of stillborn children	11.44 (regular)	22.19 (poor)	<0.001
Length of pregnancy	5.51 (good)	13.62 (regular)	<0.001
Type of pregnancy	2.12 (excellent)	6.32 (good)	<0.001
Type of delivery	2.29 (excellent)	6.74 (good)	<0.001

a) Fisher's Chi-square test.
Note: Romero & Cunha's classification for incomplete filling in of an information system data field: excellent, under 5% incompleteness; good, 5% <10%; regular, 10% <20%; poor, 20%-50%; and very poor, 50% or over. The cut-off points of the 'good' and 'regular' categories were defined to avoid overlap between the incompleteness values.

the period 2010-2012 it was only 4.3%, demonstrating potential for improved completeness.[14]

The falling trend of incompleteness, together with improvement in the quality of data found on SIM, is probably related to the work of the Infant Death Prevention Committees, given the difference found regarding investigated infant deaths in almost all variables; this fact was also reported in Ribeirão Preto and Pernambuco for infant deaths.[12,17] Although the main objective of the death investigation committees is to determine correct underlying cause of death and potential death avoidability, correcting vital information records is also an objective stipulated in the Infant and Fetal Death Surveillance Manual.[27] In our study the 'birth weight' and maternal 'age' variables fell into the excellent completeness category for investigated deaths, while they only had regular completeness scores in the case of uninvestigated deaths.

The persistence of incompleteness of sociodemographic variables, such as race/skin color, maternal occupation and schooling, hinders analysis of several mother and child health outcomes, in particular neonatal mortality.[3,19] A recent study showed that the completeness of maternal schooling data on SINASC improved for Brazil as a whole between 1996 and 2013, although it still requires efforts to achieve ideal completeness;[11] its authors stress the importance of maternal schooling in studies on infant morbidity and mortality in the context of health inequalities.

As SIM incompleteness percentages were higher than those of SINASC for all variables, this corroborates the superior quality of live birth data.[12-14,16,19] The procedure for linking the databases of both systems should be encouraged in order to enhance studies on infant mortality in Brazil. Doing this, whenever possible, will result in SINASC data completing or correcting SIM data.[18,19]

Another aspect to be discussed, although it cannot be identified on the information systems, is the training of health professionals responsible for filling in LBC and DC.[28-30] Only doctors are authorized to issue Death Certificates, whereas other health professionals can issue Live Birth Certificates. Studies have demonstrated that doctors do not receive sufficient training and/or do not keep up to date on adequate filling in of DCs.[28,30] Health professionals responsible for filling in LBCs, usually nursing or administrative staff, generally incorporate this routine into their daily work in a more consistent manner.[29] Collaboration between different health professionals working in the same service can improve the quality of these documents.[30]

One study has also described a possible difficulty felt by health professionals in recording events of a negative nature.[30] Notwithstanding, problems were found both with the filling in of unfavorable variables, i.e. stillborn children, and with the filling in of favorable variables, i.e. liveborn children and even of neutral variables, e.g. maternal occupation. This is a relevant aspect that should be considered by future studies.

Despite the length of implementation time of the mortality and live birth information systems being different, which may interfere with information quality and limit comparison of the completeness of variables common to both systems, ever since it was implemented SINASC data fields have been filled in better when compared to those of SIM.[3,4,12]

It is appropriate to comment that the study conducted by Oliveira et al.[10] about SINASC quality in the period 2006-2010 used a different classification for completeness, taking a 90% cut-off point to define adequate classification.[10] Notwithstanding, the vast majority of studies continue using the classifications and cut-off points proposed by Romero & Cunha.[7,9,13-17,19] This is the reason why our study opted to use this score in order to facilitate result comparison.

Temporal evolution of data completeness, both for SINASC and for SIM, showed favorable changes, thus reinforcing the use of these databases in epidemiological studies. However, changes in the format of some variables in 2011, such as 'length of pregnancy' and 'number of stillborn children/fetal losses', may have had repercussions on other dimensions of data quality.

The results we present will contribute to greater visibility of vital statistic information quality in the state of Rio de Janeiro. To complement this, new studies should explore the reliability and validity of this information, with the aim of making information systems, which were the object of our study, even more robust and reliable for clinical and research use. Undoubtedly the work of the death investigation committees has been fundamental for qualifying information on mortality. Our final recommendation is that efforts and procedures aimed at building the capacity of health professionals – in this case the adequate filling in of data available on SINASC and SIM – should not only be continued but also enhanced, with the aim of ensuring greater data completeness.

Authors' contributions

Lino RRG and Fonseca SC were responsible for the conception of the study, data collection and analysis.Lino RRG, Fonseca SC, Kale PL, Flores PVG, Pinheiro RS and Coeli CM contributed to data analysis and interpretation, drafting the preliminary versions and critically revising the manuscript. All the authors have approved the final version of the manuscript and declare that they are responsible for all aspects of the study, ensuring its accuracy and integrity.

References

1. Mello Jorge MHP, Ruy L, Gotlieb SLD. Análise da qualidade das estatísticas vitais brasileiras: a experiência de implantação do SIM e do Sinasc. Ciên Saúde Coletiva [Internet]. 2007 maio-jun [citado 2019 mar 19];12(3):643-54. Disponível em: http://www.scielo.br/pdf/csc/v12n3/14.pdf. doi: 10.1590/S1413-81232007000300014
2. Frias PG, Szwarcwald CL, Lira PI. Avaliação dos sistemas de informações sobre nascidos vivos e óbitos no Brasil na década de 2000. Cad Saúde Pública [Internet]. 2014 out [citado 2019 mar 19];30(10):2068-280. Disponível em: http://www.scielo.br/pdf/csp/v30n10/0102-311X-csp-30-10-2068.pdf. doi: 10.1590/0102-311X00196113
3. Romero DE, Cunha CB. Avaliação da qualidade das variáveis socioeconômicas e demográficas dos óbitos de crianças menores de um ano registrados no Sistema de Informações sobre Mortalidade do Brasil (1996/2001). Cad Saúde Pública [Internet]. 2006 mar [citado 2019 mar 19];22(3):673-84. Disponível em: http://www.scielo.br/pdf/csp/v22n3/22.pdf. doi: 10.1590/S0102-311X2006000300022
4. Romero DE, Cunha CB. Avaliação da qualidade das variáveis epidemiológicas e demográficas do Sistema de Informações sobre Nascidos Vivos, 2002. Cad Saúde Pública [Internet]. 2007 mar [citado 2019 mar 19];23(3):701-14. Disponível em: http://www.scielo.br/pdf/csp/v23n3/28.pdf. doi: 10.1590/S0102-311X2007000300028
5. Lima CRA, Schramm JMA, Coeli CM, Silva MEM. Revisão das dimensões de qualidade dos dados e métodos aplicados na avaliação dos sistemas de informação em saúde. Cad Saúde Pública [Internet]. 2009 out [citado 2019 mar 19];25(10):2095-109. Disponível em: http://www.scielo.br/pdf/csp/v25n10/02.pdf. doi: 10.1590/S0102-311X2009001000002
6. 6.Paiva NS, Coeli CM, Moreno AB, Guimaraes RM, Camargo KR. Sistema de informações sobre nascidos vivos: um estudo de revisão. Ciên Saúde Coletiva [Internet]. 2011 [citado 2019 mar 19];16 Supl 1:1211-20. Disponível em: http://www.scielo.br/pdf/csc/v16s1/a53v16s1.pdf. doi: 10.1590/S1413-81232011000700053
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 [Internet]. 2013 abr-jun [citado 2019 mar 19];22(2):347-52. Disponível em: http://scielo.iec.gov.br/pdf/ess/v22n2/v22n2a16.pdf. doi: 10.5123/S1679-49742013000200016
8. Guimarães PV, Coeli CM, Cardoso RCA, Medronho RA, Fonseca SC, Pinheiro RS. Reliability of data from a very low birth weight population in the Live Birth Information System 2005-2006. Rev Bras Epidemiol [Internet]. 2012 Dec [citado 2019 Mar 19];15(4):694-704. Disponível em: http://www.scielo.br/pdf/rbepid/v15n4/02.pdf. doi: 10.1590/S1415-790X2012000400002
9. Gabriel GP, Chiquetto L, Morcillo AM, Ferreira MC, Bazan IGM, Daolio LD, et al. Avaliação das informações das Declarações de Nascidos Vivos do Sistema de Informação sobre Nascidos Vivos (Sinasc) em Campinas, São Paulo, 2009. Rev Paul Pediatr [Internet]. 2014 set [citado 2019 mar 19];32(3):183-8. Disponível em: http://www.scielo.br/pdf/rpp/v32n3/0103-0582-rpp-32-03-0183.pdf. doi: 10.1590/0103-0582201432306
10. Oliveira MM, Andrade SSCA, Dimech GS, Oliveira JCG, Malta DC, Rabello Neto DL, et al. Avaliação do Sistema de Informações sobre Nascidos Vivos. Brasil, 2006 a 2010. Epidemiol Serv Saúde [Internet]. 2015 out-dez;24(4):629-40. Disponível em: http://www.scielo.br/pdf/ress/v24n4/2237-9622-ress-24-04-00629.pdf. doi: 10.5123/S1679-49742015000400005
11. Silvestrin S, Buriol VCS, Silva CH, Goldani MZ. Avaliação da incompletude da variável escolaridade materna nos registros das Declarações de Nascidos Vivos nas capitais brasileiras - 1996 a 2013.Cad Saúde Pública [Internet]. 2018 [citado 2019 mar 19];34(2):e00039217. Disponível em: http://www.scielo.br/pdf/csp/v34n2/1678-4464-csp-34-02-e00039217.pdf. doi: 10.1590/0102-311x00039217
12. Barbuscia DM, Rodrigues Júnior AL. Completude da informação nas declarações de nascido vivo e nas declarações de óbito, neonatal precoce e fetal, da

região de Ribeirão Preto, São Paulo, Brasil, 2000-2007. Cad Saúde Pública [Internet]. 2011 jun [citado 2019 mar 19];27(6):1192-200. Disponível em: http://www.scielo.br/pdf/csp/v27n6/16.pdf. doi: 10.1590/S0102-311X2011000600016

13. Silva LP, Moreira CMM, Amorim MHC, Castro DS, Zandonade E. 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. Ciên Saúde Coletiva [Internet]. 2014 jul [citado 2019 mar 19];19(7):2011-20. Disponível em: http://www.scielo.br/pdf/csc/v19n7/1413-8123-csc-19-07-02011.pdf. doi: 10.1590/1413-81232014197.08922013

14. 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 [Internet]. 2016 out-dez [citado 2019 mar 19];25(4):849-54. Disponível em: http://www.scielo.br/pdf/ress/v25n4/2237-9622-ress-25-04-00849.pdf. doi: 10.5123/s1679-49742016000400019

15. Dombrowski JG, Ataíde R, Marchesini P, de Souza RM, Marinho CR. Effectiveness of the Live Births Information System in the Far-Western Brazilian Amazon. Ciên Saúde Coletiva [Internet]. 2015 Apr [citado 2019 Mar 19];20(4):1245-54. Disponível em: http://www.scielo.br/pdf/csc/v20n4/1413-8123-csc-20-04-01245.pdf. doi: 10.1590/1413-81232015204.00792014

16. Barreto IC, Vieira MG, Teixeira GP, Fonseca SC. Morte neonatal: incompletude das estatísticas vitais. Rev Bras Pesq Saúde [Internet]. 2017 abr-jun [citado 2019 mar 19];19(2):64-72. Disponível em: http://periodicos.ufes.br/RBPS/article/view/18863/12840. doi: 10.21722/rbps.v19i2.18863

17. Ramalho MO, Frias PG, Vanderlei LC, Macêdo VC, Lira PI. Avaliação da incompletude da declaração de óbitos de menores de um ano em Pernambuco, Brasil, 1999-2011. Cien Saúde Coletiva [Internet]. 2015 set [citado 2019 mar 19];20(9):2891-8. Disponível em: http://www.scielo.br/pdf/csc/v20n9/1413-8123-csc-20-09-2891.pdf. doi: 10.1590/1413-81232015209.09492014

18. Maia LTS, Souza WV, Mendes ACG. A contribuição do linkage entre o SIM e SINASC para amelhoria das informações da mortalidade infantil em cinco cidades brasileiras. Rev Bras Saúde Mater Infant [Internet]. 2015 jan-mar [citado 2019 mar 19];15(1):57-66. Disponível em: http://www.scielo.br/pdf/rbsmi/v15n1/1519-3829-rbsmi-15-01-0057.pdf. doi: 10.1590/S1519-38292015000100005

19. Maia LTS, Souza WV, Mendes ADCG, Silva AGSD. Uso do linkage para a melhoria da completude do SIM e do Sinasc nas capitais brasileiras. Rev Saúde Pública [Internet]. 2017 [citado 2019 mar 19];51:112. Disponível em: https://www.revistas.usp.br/rsp/article/view/141551. doi: 10.11606/S1518-8787.2017051000431

20. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Coordenação Geral de Informações e Análise Epidemiológica. Consolidação do Sistema de Informações sobre Nascidos Vivos – 2011 [Internet]. Brasília: Ministério da Saúde; 2013 [citado 20 mar 2017]. Disponível em: http://tabnet.datasus.gov.br/cgi/sinasc/Consolida_Sinasc_2011

21. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Coordenação Geral de Informações e Análise Epidemiológica. Consolidação do Sistema de Informações sobre Mortalidade – 2011 [Internet]. Brasília: Ministério da Saúde; 2013 [citado 20 mar 2017]. Disponível em: http://tabnet.datasus.gov.br/cgi/sim/Consolida_Sim_2011.pdf

22. Brasil. Ministério da Saúde. Portaria MS/GM nº 72, de 11 de janeiro de 2010. Estabelece que a vigilância do óbito infantil e fetal é obrigatória nos serviços de saúde (públicos e privados) que integram o Sistema Único de Saúde (SUS) [Internet]. Diário Oficial União, Brasília (DF), 2010 jan 12 [citado 2019 mar 19]; Seção 1:29. Disponível em: http://bvsms.saude.gov.br.saudelegis/gm/2010/prt0072_11_01_2010.html

23. Daripa M, Caldas HM, Flores LP, Waldvogel BC, Guinsburg R, Almeida MF. Asfixia perinatal associada à mortalidade neonatal precoce: estudo populacional dos óbitos evitáveis. Rev Paul Pediatr [Internet]. 2013 mar [citado 2019 mar 19];31(1):37-45. Disponível em: http://www.scielo.br/pdf/rpp/v31n1/07.pdf. doi: 10.1590/S0103-05822013000100007

24. Morais RM, Costa AL, Gomes EJ. Os sistemas de informação do SUS: uma perspectiva histórica e as políticas de informação e informática. Nucleus [Internet]. 2014 abr [citado 2019 mar 19];11(1):287-04. Disponível em: http://www.nucleus.feituverava.com.br/index.php/nucleus/article/view/998/1427. doi: 10.3738/1982.2278.998

25. Veloso HJ, Silva AA, Barbieri MA, Goldani MZ, Lamy Filho F, Simões VM, et al. Secular trends in the rate of low birth weight in Brazilian State Capitals in the period 1996 to 2010. Cad Saúde Pública [Internet]. 2013 Jan [citado 2019 Mar 19];29(1):91-101. Disponível em: http://www.scielo.br/pdf/csp/v29n1/11.pdf. doi: 10.1590/S0102-311X2013000100011

26. Cordeiro TMSC, D'Oliveira Júnior A. Qualidade dos dados das notificações de hepatites virais por acidentes de trabalho, Brasil. Rev Bras Epidemiol [Internet]. 2018 [citado 2019 mar 19];21:e180006. Disponível em: http://www.scielo.br/pdf/rbepid/v21/1980-5497-rbepid-21-e180006.pdf. doi: 10.1590/1980-549720180006

27. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Secretaria de Atenção à Saúde. Manual de vigilância do óbito infantil e fetal e do comitê de prevenção do óbito infantil e fetal [Internet]. 2. ed. Brasília: Ministério da Saúde; 2009 [citado 2019 mar 19]. 96 p. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/manual_obito_infantil_fetal_2ed.pdf

28. Mendonça FM, Drumond E, Cardoso AMP. Problemas no preenchimento da Declaração de Óbito: estudo exploratório. Rev Bras Estud Popul [Internet]. 2010 jul-dez [citado 2019 mar 19];27(2):285-95. Disponível em: http://www.scielo.br/pdf/rbepop/v27n2/04.pdf. doi: 10.1590/S0102-30982010000200004

29. Schoeps D, Almeida MF, Raspantini PR, Novaes HMD, Silva ZP, Lefevre F. SIM e SINASC: representação social de enfermeiros e profissionais de setores administrativos que atuam em hospitais no município de São Paulo. Ciên Saúde Coletiva [Internet]. 2013 maio [citado 2019 mar 19];18(5):1483-92. Disponível em: http://www.scielo.br/pdf/csc/v18n5/34.pdf. doi: 10.1590/S1413-81232013000500034

30. Schoeps D, Lefevre F, Silva ZP, Novaes HMD, Raspantini PRM, Almeida MF. Representações sociais de médicos obstetras e neonatologistas sobre declaração de óbito fetal e neonatal precoce no município de São Paulo. Rev Bras Epidemiol [Internet]. 2014 jan-mar [citado 2019 mar 19];17(1):105-18. Disponível em: http://www.scielo.br/pdf/rbepid/v17n1/pt_1415-790X-rbepid-17-01-00105.pdf. doi: 10.1590/1415-790X201400010009ENG

Received on 03/06/2018
Approved on 17/02/2019

Associate editor: Doroteia Aparecida Höfelmann – orcid.org/0000-0003-1046-3319