

Eliane de Freitas Drumond<sup>I</sup>

Carla Jorge Machado<sup>II</sup>

Elisabeth França<sup>III</sup>

# Underreporting of live births: measurement procedures using the Hospital Information System

---

## ABSTRACT

**OBJECTIVE:** To assess underreporting of live birth records by health information systems.

**METHODS:** Secondary data of the Sistema de Informação Hospitalar (Hospital Information System – SIH) and of the Sistema de Informação de Nascidos Vivos (Information System on Live Birth – SINASC) were used in the state of Minas Gerais, Southeastern Brazil, in 2001. Two procedures were used in the analysis: the comparison between the number of live births per city and the probabilistic record linkage of individual data. For both procedures, indicators of underreporting considered were the proportion of live births presented at SIH system that were not obtained at SINASC. The municipalities were later added into four strips of population size.

**RESULTS:** The probabilistic linkage was able to identify a greater proportion of live births underreported at SINASC, relative to the comparison of live births in the municipalities. The variations of the differences among underreporting percentages per procedures were 9.4% in cities with population lower than 5,000 inhabitants; 9.1% in cities with population ranging from 5,000 and 9,999; and 8.0% in municipalities between 10,000 and 49,999 and over 50,000 inhabitants.

**CONCLUSIONS:** The amount of underreporting was sensitive to the procedures adopted. Probabilistic linkage reinforced the certainty of pairings, and also enabled to identify a greater proportion of cases not recorded at SINASC, also in greater cities. SIH was a strong indicator of underreporting of live births.

**KEY WORDS:** Live birth. Registries. Underregistration. Systems integration. Vital statistics.

---

## INTRODUCTION

The information systems in national health have several data volume to plan and assess public policies. Its use enables that the possible problems coming from the quality of information are known and solved.<sup>12</sup> The importance and the difficulties to obtain the information on the number of live births, whose official source is the *Instituto Brasileiro de Geografia e Estatística* (Brazilian Institute of Geography and Statistics – IBGE) are acknowledged. However, its use is limited especially due to underreporting. To face this issue, in 1990 the Ministry of Health (MS) introduced the Information System on Live births

<sup>I</sup> Gerência de Epidemiologia e Informação. Secretaria Municipal de Saúde. Belo Horizonte, MG, Brasil

<sup>II</sup> Departamento de Demografia. Faculdade de Ciências Econômicas. Universidade Federal de Minas Gerais (UFMG). Belo Horizonte, MG, Brasil

<sup>III</sup> Departamento de Medicina Preventiva e Social. Faculdade de Medicina. UFMG. Belo Horizonte, MG, Brasil

### Correspondence:

Carla Jorge Machado  
Av. Augusto de Lima 1376/ sala 908  
30190-003 Belo Horizonte, MG, Brasil  
E-mail: carla@cedeplar.ufmg.br

Received: 7/13/2007  
Approved: 9/10/2007

(SINASC)<sup>a</sup> which presents important variables for epidemiological analysis such as weight at birth. Even though there were significant improvements,<sup>3,12</sup> there are still some problems in the quality of information of SINASC.<sup>b</sup> Furthermore, underreporting of live births reflects the inability to capture these events by the health system. This reality, seen as a challenge for public health, encourages health services and researchers<sup>12</sup> to explore ways of measuring and decreasing this underreport.

There are local, regional and state differences on the coverage and quality of SINASC,<sup>12</sup> managed by the three levels of government (federal, state and municipality). However, the role of municipal managers in the qualification of SINASC is increasingly important due to the progressive decentralization of the management of this system. Therefore, to estimate and locate the underreporting of live births in the municipal level allow adopting specific and necessary actions for its reduction locally.

One of the techniques to identify underreporting of live births is to link the information at SINASC to the information of other health information systems.<sup>9</sup> With this purpose, we have used, for example, the *Sistema de Informação da Atenção Básica* (Primary Care Information System - SIAB)<sup>11</sup> and the *Sistema de Informações Hospitalares* (Hospital Information System - SIH).<sup>15</sup>

SIH is the only health information system on hospital morbidity and it presents epidemiological variables,<sup>4</sup> in addition to information for accounting purposes. It refers only to records of admissions occurred in hospitals of the National Health System (SUS) or those which have agreements with it. Therefore, SIH does not cover home labors, those paid privately or by insurance plans. Thus, when the records of SINASC – with universal coverage – are compared to those of SIH, we assume that SINASC would always present equal or greater number of reports of live births. If SINASC does not present a volume of reports equal to or greater than that of SIH, the difference between the two could be considered an indicator of underreporting<sup>15</sup> of live births at SINASC.

Based on this assumption, Almeida & Alencar<sup>1</sup> proposed to use delivery numbers at SIH as a parameter to capture live births at SINASC. Schramm & Szwarcwald<sup>15</sup> used data from SIH also to estimate live birth mortality and neonate mortality in Brazilian states. In both studies, although it is impossible to state that it was the same live birth reported by both systems, the greater number

of delivery records by SIH led to the conclusion that it was a viable source to capture live births in areas with lower coverage by SINASC.

To decrease error probability when we say that a record in a health information system refers to the same individual in another health information system, they have to be linked, at least, through one blocking variable. The probabilistic linkage enables to state that records from different databases are from the same individual, based on agreement and disagreement probabilities of the key variables. Several strategies of probabilistic record linkage have been used<sup>15,8,9</sup> to overcome the absence of a unique identifier. Very frequently, in probabilistic linkage, the names of patients are used as the main blocking variable, associated or not with other variables.<sup>5,9</sup>

Aiming at exploring the idea that using different procedures to encourage underreporting lead to different results, the objective of the present study was to assess underreporting of records of live births at SINASC obtained by two procedures.

## METHODS

Data from SINASC and from SIH were used referring to live births and deliveries of residents of Minas Gerais in 2001. To compare the number of events, we used data from a CD-ROM given by the Information Technology Department of SUS (DATASUS).<sup>c</sup> For the probabilistic linkage, data was obtained in the Health State Secretariat (SES) of Minas Gerais. They had individual blocking variables enabling previous correction of duplicates and errors in the specification of residence municipalities.

Data from SIH referred to the authorized procedure in 2001 and 2002. In the subdirectory MA (Part of AIH), the records which had the variable 54 (LIVE\_BIRTHS) filled in were selected, which is only filled in for deliveries and it enabled to identify admissions for deliveries of live births. Then, the number of live births reported at SINASC and of deliveries of live births reported by SIH through the variable place of residence at SINASC (CODMUNRES) and at SIH (MUNIC\_RES).

The probabilistic linkage was adopted because there was not a blocking variable in the two data base. Of SINASC we have included all the records of live births, living in Minas Gerais. For inclusion of the records of SIH three criteria were used: 1) admission of women; 2) admission in 2001 and 3) field 54 different from zero.

<sup>a</sup> Ministry of Health. Health Information Nascidos Vivos/Notas Técnicas.[accessed on 11/27/ 2007]. Available from <http://tabnet.datasus.gov.br/cgi/sinasc/nvdescr.htm#atdados>

<sup>b</sup> *Indicadores Básicos para Saúde no Brasil* (Health Metrics -IDB 2006). Rede Interagencial de Informações para Saúde [accessed on 6/16/2007]. Available from: <http://tabnet.datasus.gov.br/cgi/idb2006/c01.htm>.

<sup>c</sup> Ministry of Health. Health Surveillance Secretariat. Database of *Sistemas de Informações sobre Mortalidade* (Mortality Information System SIM) and Live Births (SINASC), 1997 to 2003. [CD-ROM]. Brasília; 2005.

Some adjustments in the database have been necessary to perform probabilistic linkage suitably.<sup>7</sup> To that end, the field has been maintained identical in the two files. Dates have been standardized in order to have the same format (string) and a CLSS field has been created coming from the field NOME (name) and DTSTRING matched, where the content of DTSTRING has been added to the end of the field name. This result has been stored at the field CLSS.

After variables had been prepared, the six steps of the linking process were started.<sup>7</sup> Record blocks were created with homologous fields (with information of the same nature) and association of the blocks matching one or more fields, it started with a more restricted key and with a later decrease in restriction (Table 1). Thus, we have tried to minimize the loss of pairs, that is, the occurrence of false-negative. In the first five steps we have maintained one authorization for hospital stay (AIH) for a birth declaration (DN). Only on the sixth step the correlation of an AIH was admitted for one or more DN.

The following functions have been defined to use in the blocking keys:

1. Soundex of the mother's name: created to reduce loss of true pairs due to problems coming from mistakes and/or differences in spelling;
2. Difdata: created to return in days the difference between the date of admission and birth, since in SIH the application is for admission and there may be several admissions for the same individual in the same year; this function aimed at minimizing the possibility of false-positive, unreal pairs;
3. Municnasc: created to compare codes among coincident birth municipalities, returning equal for coincident cases or different for non-coincident cases, and aimed especially at reducing the formation of unreal pairs, which is possible by the presence of hospitals with the same name in different cities;
4. Municres: created to compare codes of municipalities where mothers live, showing equal or different; aimed at rescuing true pairs that were not formed in previous steps;
5. Estab: function created to compare establishments where births have occurred, showing equal or different.

Linkages were performed using a sequence in SQL language called Datalink,<sup>7</sup> which makes the direct comparison between Soundex according to first name in the two databases. The single field employed for linkage in all blocking keys was CLSS, submitted to

the function of Levenshtein distance. This result underwent another function, called percentage of similarity, which determined the percentage of similarity among the strings compared, ranging from zero to 100%.<sup>7</sup> This similarity percentage was stored in a variable called SCORE.

Between each automated step, a manual review of the formed pair was conducted so they could be admitted as true pairs to check suitability of the median of the SCORE field as the cut-off point adopted. When manual review of the pairs in absolute agreement was conducted, we have seen that intervals among admissions of mothers and births were, mostly, lower than four days. However, aiming at including a greater number of true pairs, longer intervals have been admitted.

Based on data from DATASUS (comparison of the number of events) and from SES (probabilistic linkage), the total number of live births were calculated together with the ratio between live births at SIH and SINASC according to group of population of the residence municipalities.<sup>8</sup> Although criteria to classify municipalities according to group of population was not been found, the 853 municipalities were classified according to four groups of population: municipalities with population smaller than 5,000 inhabitants (small municipalities); population among 5,000 and 9,999 inhabitants (middle-size municipalities level 1); population among 10,000 and 49,999 inhabitants (middle-size municipalities level 2) and population greater than or equal to 50,000 inhabitants (big cities). We decided to define a group of small municipalities because of the great number of municipalities included in this category (N=246). Another 547 municipalities were subdivided into two similar groups: one with 272 municipalities and population from 5,000 to 9,999 inhabitants, and another with 275 municipalities with 10,000 to 49,999 inhabitants. And for the 60 municipalities with over 50,000 inhabitants, the classification (big city) proposed by Andrade & Szwarcwald was adopted.<sup>3</sup>

To process data analysis, the program SPSS 12.0 was used. The study was approved by the Ethical Research Committee of the Universidade Federal de Minas Gerais (ETIC N. 095/04).

## RESULTS

In the data bases of DATASUS and SES/MG 298,515 and 293,213 live births were respectively identified, and at SIH 237,441 and 223,443 admissions for deliveries of live births were respectively identified at SINASC.

In the small municipalities, 11,715 live births were found reported at SINASC and 12,258 deliveries of live births reported at SIH. Thus, 543 records from

<sup>8</sup> Ministry of Health. DATASUS. Health Information – demographic and socioeconomic [accessed on 7/3/2007]. Available from: <http://tabnet.datasus.gov.br/cgi/deftohtm.exe?ibge/cnv/poptmg.htm>

**Table 1.** Steps for linkage in data base of the Hospital Information System (SIH) and Information System for Live Births (SINASC). State of Minas Gerais, Southeastern Brazil, 2001.

Step	Linkage key
1	Soundex (first name) + Difdata + Municnasc + Estab
2	Soundex (first name) + Difdata + Municnasc
3	Soundex (first name) + Difdata + Municres
4	Soundex (first name) + Difdata
5	Soundex (first name)
6 (twins)	Soundex (first name)

SIH were not present at SINASC. We have assumed that after correction had been performed adding the live births captured only by SIH, there would be a total of 12,258 live births in this group of municipalities. Consequently, the percentage of live births underreported at SINASC and known in the records of SIH was 4.43%, that is, 543 compared to a total corrected at SINASC of 12,258 (Table 2).

In level 1 middle-size municipalities, 28,213 live births reported at SINASC, and 28,912 deliveries of live births at SIH were recorded. Thus, at least for 699 live births there was no report at SINASC, only at SIH. Likewise, after correction, it was considered that there were 28,912 live births in these municipalities and that the percentage of cases known by SIH was 2.4%, that is, 699 in relation to a total corrected at SINASC of 28,912 live births (Table 2).

In level 2 middle-size municipalities, 2,343 unreported cases at SINASC, reports at SIH have been found. The total of live births, after correction, was 94,831 (92,488 of SINASC in addition to another 2,343 identified at

SIH). In this group, the percentage of live births known at SIH and unknown at SINASC was 2.5%.

In big cities, underreporting at SINASC has not been identified. On the contrary, fewer reports at SIH (101,440) than at SINASC (166,099) have been found.

Thus, 3,585 unreported deliveries of live births at SIH among the 302,100 live births at SINASC have been obtained by the comparison of the number of events. By the definition of underreporting adopted, the percentage of live births underreported at SINASC and known only at SIH was 1.2%. These underreporting, occurred in 189 municipalities, located especially in the North of the State and in Jequitinhonha.

Using probabilistic linkage, 193,259 univocal pairs of live births have been identified – link of a DN to one AIH. We have also obtained 6,316 twin pairs, in which 3,145 AIH were linked to 6,316 DN, resulting, on average, in 2.01 DN for each AIH.

Table 3 indicates the distribution of pairs and non-pairs for reports of SIH and of SINASC, according to population group. Regarding SINASC, of a total of 293,213 DN, at least one pair was obtained for 199,565 reports, resulting in 68.0% of DN paired to one AIH. For SIH, of the 223,443 AIH, it was possible to obtain pairs for 87.9% of them (N=196,404). Thus, 27,039 records of AIH and of SIH were not paired to records of SINASC and, according to the underreporting criteria adopted, they were considered as underreported of live births at SINASC.

We have observed in small municipalities that there were 9,995 deliveries of live births at SIH linked to 8,111 reports among the 11,277 live births at SINASC. That is, in these municipalities, 1,884 reports of deliveries at SIH did not pair with any DN and were

**Table 2.** Absolute distribution and percentage of records at SINASC and SIH, underreporting at SINASC and live births corrected at SINASC according to the procedure to compare the number of events per group of population of the cities of the State of Minas Gerais, Southeastern Brazil, 2001.

Variable	Group of population	Group of population				Total
		< 5,000	5,000 - 9,999	10,000 – 49,999	>50,000	
Sinasc	n (1)	11,715	28,213	92,488	166,099	298,515
	%	4.1	9.6	31.4	55.0	100
SIH	n (2)	12,258	28,912	94,831	101,440	237,441
	%	5.2	12.2	39.9	42.7	100
Underreporting at Sinasc	n (3)	543	699	2,343	0	3,585
	%	15.2	19.5	65.4	0	100
Corrected Sinasc	n (1+3)	12,258	28,912	94,831	166,099	302,100
	%	3.9	9.5	31.0	100	100

SIH: Hospital Information System

SINASC: Information System on Live Births

Source: Information Department of SUS – DATASUS

**Table 3.** Absolute Distribution and percentage of pairs and non-pairs of the records at SINASC and SIH and of live births corrected at SINASC according to probabilistic linkage per population group of the municipalities of the State of Minas Gerais, Southeastern Brazil, 2001.

Variable		Population group				Total
		< 5,000	5,000 - 9,999	10,000 - 49,999	>50,000	
SINASC pairs	n (1)	8,253	19,711	65,101	106,510	199,404
	%	4.1	9.9	33.0	53.0	100
SIH pairs	n (2)	8,111	19,404	64,119	104,770	196,404
	%	4.1	9.9	32.6	53.3	100
Sinasc non-pairs	n (3)	3,024	7,499	25,092	58,023	93,638
	%	3.2	8.0	26.8	62.0	100
SIH non-pairs	n (4)	1,884	3,596	10,689	10,870	27,039
	%	7.0	13.3	39.5	40.2	100
Corrected SINASC	n (1+3+4)	13,161	30,806	100,882	175,403	320,252
	%	4.1	9.6	31.5	54.8	100

Source: Health State Secretariat of Minas Gerais SES/MG

**Table 4.** Distribution of absolute number and percentages of live births reported at SINASC and of live births estimated at SINASC after correction according to population group in the cities of the State of Minas Gerais, Southeastern Brazil, 2001.

Variable	Population group				Total
	<5,000	5,000 - 9,999	10,000 - 49,999	>50,000	
Unreported LB					
Comparison of Events	543	699	2,343	0	3,585
Probabilistic	1,884	3,695	10,689	10,870	27,039
Reported LB					
Comparison of Events	11,715	28,213	92,488	166,099	298,515
Probabilistic	11,277	27,210	90,193	164,533	293,213
Total after correction					
Comparison of Events	12,258	28,912	94,831	166,099	302,100
Probabilistic	13,161	30,305	100,882	175,403	320,252
Percentage of LB in comparison to the total corrected					
Comparison of Events	4.4	2.4	2.5	0	1.2
Probabilistic	14.3	11.7	10.6	6.2	8.4

Source: SES/MG and DATASUS

LB: live birth

considered as underreporting of SINASC. In level 1 middle-size municipalities, of the 23,000 deliveries of live births at SIH, 19,404 paired with 19,711 reports at SINASC. Thus, in this group of municipalities, for 3,596 reports of SIH there were no pairs among the records of SINASC, and they were also considered as underreporting of SINASC. For level 2 middle-size municipalities, there were no pairs and they were considered as underreporting of SINASC in 10,689 records of deliveries of live births at SIH. In big cities, 10,870 records of SIH were underreported at SINASC.

The results of the probabilistic linkage according to resident population group presented similarities in the

distribution of pairs of SINASC and SIH, but differences between the percentages of non-pairs. Regarding SINASC, there was greater concentration of unpaired DN in big cities. And regarding SIH, greater concentration of unpaired AIH was observed especially in small and middle-size municipalities (Table 3).

Based on the unpaired records at SIH, the expected number of notification of live births at SINASC, per group of population was calculated. Thus, a corrected total of 13,161 live births at SINASC has been obtained in small municipalities, using probabilistic linkage (Table 3).

Differences between the proportions of live births using SIH have been observed according to the procedure

adopted. Probabilistic linkage identified 8.4% and the comparison of events per municipality has identified 1.2% of underreporting at SINASC. In the two procedures, underreporting of live births was lower in big cities, and was identified only by probabilistic linkage (Table 4).

## DISCUSSION

The main interest to perform the present study was the need to explore techniques to directly measure underreporting of live births per increasingly separated geographical unities, such as the municipalities. Obtaining good quality information on live births enables, for example, to calculate directly infant mortality rate. This is an essential step to adequate public health policies aiming at decreasing these deaths.<sup>a</sup>

Among the results obtained in the present study, we can highlight: 1) the little explored potentiality of SIH as a source to capture live births at SINASC, verified by both procedures performed here. This result matches the studies that have used linkage strategies of the data from SIH and other health information systems to assess underreporting of events and diseases<sup>d</sup> and 2) the differences in results obtained, according to the procedure adopted. Thus, regardless of the population group considered, the probabilistic linkage identified greater number of live births underreported at SINASC, when it obtained greater proportion of live births known only at SIH. A possible explanation for the difference obtained between the two procedures is due to the type of error that each procedure may be subject to. Comparison of the number of events would be more prone to false-positive errors. When only the variable municipality of residence is used, it would be possible to mistakenly consider two records as belonging to the same individual. On the other hand, probabilistic linkage would be relatively more prone to false-negative errors when six variables are used; that is, it would be possible to consider two records that do not belong to the same individual. Thus, a greater number of records found at SIH may not necessarily mean greater proportion of underreporting of live births at SINASC.

The quality of information recorded in the two systems may also affect the formation of true pairs in the probabilistic linkage, since the poor quality increases the probability of homonymous errors.<sup>e</sup> Some pairs may be classified as true when in fact records refer to different individuals. Thus, as in the present study in cases where the data base in which the pair is searched for (SINASC) presents a high number of reports, rules to minimize the number of false-positive pairs are

adopted. This procedure, on the other hand, may lead to an increase in the number of false-negative results. Therefore, it seems reasonable to suppose, that the real magnitude of the proportion of underreports is higher than that observed in the comparison of events, however, lower than that observed using probabilistic linkage. Thus, the two methods, if used together, may supply one inferior and one higher limit for the real magnitude of the underreporting.

With the comparison of the number of events, we could identify small proportions of underreporting of live births at SINASC, lower than 5%, only in 189 middle-size and small municipalities. The absence of underreporting at SINASC in big cities seems unlike.<sup>14</sup> But, the procedure presents advantages such as the facility to obtain data, fast performance and methodological easiness. Therefore, it is possible even when there are few technical resources. Its main disadvantages are: impossibility to state that individuals recorded are the same in each of the information systems and inability to identify underreporting of live births in municipalities with poorer coverage of SIH. Thus, based on the results obtained, it is possible to infer that the comparison of the number of events may have a great performance in the cities where deliveries covered by the National Health System (SUS) is close to the total number of births of the municipality. The follow-up and surveillance of systems of vital statistics, especially necessary in this group of municipalities, may benefit from the calculation and use of this parameter.

Using probabilistic linkage, we could identify live births at SIH that were not reported at SINASC, in 852 of the 853 municipalities of Minas Gerais, including the capital city. The number of live births identified only at SIH was seven times greater than that obtained by the comparison of number of events. The 12.1% percentage of live births underreported at SINASC obtained by probabilistic linkage matches the one estimated by the Ministry of Health (13.7%)<sup>b</sup> for Minas Gerais and it corroborates the suitability of the results obtained using probabilistic linkage. It also corroborates this suitability of greater proportion of unpaired records at SINASC compared to those of SIH in big cities, where the population with better socioeconomic conditions is concentrated and that do not use the National Health System, and where a greater percentage of deliveries is paid by insurance plans. In Minas Gerais, specifically, for the year 2006, estimated population coverage of insurance plans was 18.9%; 45.5% was in the capital city, and 31.6% in the metropolitan region.<sup>c</sup> Although deliveries occurring at home or in the traffic, more common among the socially less favored groups, are

<sup>a</sup> Organização Mundial de Saúde. Objectifs du Millénaire pour le développement [accessed on 7/26/2007]. Available from: <http://www.un.org/french/millenniumgoals/index.html>

<sup>b</sup> Ministry of Health. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde. Saúde Brasil 2005: uma análise de saúde no Brasil. [accessed on 11/27/2007]. Available from: [http://portal.saude.gov.br/portal/arquivos/pdf/saude\\_brasil\\_2005.pdf](http://portal.saude.gov.br/portal/arquivos/pdf/saude_brasil_2005.pdf)

<sup>c</sup> Ministry of Health. Agência Nacional de Saúde Suplementar. Caderno de Informação de Saúde Suplementar: beneficiários, operadoras e planos. [accessed on 6/16/2007]. Available from: [http://www.ans.gov.br/portal/upload/informacoess/caderno\\_informaca\\_12\\_2006.pdf](http://www.ans.gov.br/portal/upload/informacoess/caderno_informaca_12_2006.pdf)

excluded from SIH, the search for underreporting of live births at SINASC through SIH sheds light on the population that uses SUS, among which those more socially vulnerable are.

Among the main advantages of probabilistic linkage is the considerable decrease in uncertainty of pairing, enabling detection of underreporting of live births at SINASC, even in a city with small coverage of deliveries by SUS. The importance of SUS hospitals as reporting sources of live births in all municipalities is confirmed, regardless of the population size. But the probabilistic linkage also presents difficulties, among them the problems regarding completing the variables and using different codes for each system. To identify the health institution where delivery took place, for example, each system uses its own way. The *Cadastro Nacional de Estabelecimentos de Saúde* (National Registry of Health Establishments – CNES), that is being implemented, has a great facilitator potential to link data from the occurrence.

The variable occurrence municipality presented limitations related to the confidence and completeness.<sup>7</sup> However, we have chosen to use it because of the so called “ordeal of parturients”,<sup>13</sup> they feared they would not be taken care of in a municipality where they did not live, so they do not say correctly where they live. Thus, there could be a greater tendency to error if the variable municipality of residence was the only one used and it justifies why municipality codes were in different blocking keys. Linkage strategy of manually

reviewing all the pairs formed aimed at reducing probable false-negatives. However, this procedure demands clearly defined criteria and its performance is time consuming.

Analysis of variables possibly associated with underreporting of live births at SINASC identified by the probabilistic linkage is not the purpose of the present study. However, among the aspects related to underreporting, the problems of understanding concepts such as abortion, fetal death and live birth could be mentioned. Mello-Jorge et al (1993)<sup>10</sup> observed the increase in underreporting of live births in the cases considered as little viable (low weight, for example) and of stillborn that die in the first minutes after delivery (suffocated, for example). Another aspect refers to delivery evasions that could lead to failure in feedback data among the municipalities.<sup>7</sup>

Finally, in studies based on secondary data it should be considered the possibility of information bias. In the present study, the records were obtained at SINASC and SIH without performing independent validation. Regarding SINASC, assessment studies of their coverage and completeness available since the beginning of the 90's,<sup>10</sup> how their progressive qualification.<sup>2,3,12</sup> Veras & Martins<sup>16</sup> verified that confidence of data from SIH was higher than considered – including for the variable live birth. New analysis of the validity and confidence of data from SIH will help this system to be more often used as a trustable and useful source of data for research, planning and assessment in health.

## REFERENCES

1. Almeida MF, Alencar GP. Informações em saúde: necessidade de introdução de mecanismos de gerenciamento dos sistemas. *Inf Epidemiol Sus.* 2000;9(4):241-9.
2. Almeida MF, Alencar GP, Novaes HMD, Ortiz LP. Sistemas de informação e mortalidade perinatal: conceitos e condições de uso em estudos epidemiológicos. *Rev Bras Epidemiol.* 2006;9(1):56-68.
3. Andrade CL, Szwarcwald CL. Desigualdades sócio-espaciais da adequação das informações de nascimentos e óbitos do Ministério da Saúde, Brasil, 2000-2002. *Cad Saude Publica.* 2007;23(5):1207-16.
4. Bittencourt SA, Camacho LAB, Leal MC. O Sistema de Informação Hospitalar e sua aplicação na saúde coletiva. *Cad Saude Publica.* 2006;22(1):19-30.
5. Camargo Jr KR, Coeli CM. *Reclink*: aplicativo para o relacionamento de bases de dados, implementando o método probabilístico record linkage. *Cad Saude Publica.* 2000;16(2):439-47.
6. Coutinho ESF, Coeli, CM. Acurácia da metodologia de relacionamento probabilístico de registros para identificação de óbitos em estudos de sobrevivência. *Cad Saude Publica.* 2006;22(10):2249-52.
7. Drumond EF, Machado CJ, França E. SIHSUS e Sinasc: utilização do método probabilístico para relacionamento de dados. *Cad Saude Coletiva.* 2006;14(2):251-64.
8. Machado CJ. Procedimentos para relacionamento de registros: revisão bibliográfica com enfoque na saúde infantil. *Cad Saude Publica.* 2004;20(2):362-71.
9. Machado CJ, Hill K. Relacionamento probabilístico de dados e um procedimento automático para minimizar o problema da incerteza no pareamento de registros. *Cad Saude Publica.* 2004;20(4):915-25.
10. Mello Jorge MHP, Gotlieb SLD, Soboll MLMS, Almeida MF, Latorre MRDO. Avaliação do sistema de informação sobre nascidos vivos e o uso de seus dados em epidemiologia e estatísticas de saúde. *Rev Saude Publica.* 1993;27(Supl ):1-46.
11. Mello-Jorge MHP, Gotlieb SLD. O Sistema de Informação de Atenção Básica como fonte de dados para os Sistemas de Informações sobre Mortalidade e sobre Nascidos Vivos. *Inf Epidemiol Sus.* 2001;10(10):7-18.
12. Mello-Jorge MHP, Laurenti R, Gottlieb SLD. Análise da qualidade das estatísticas vitais brasileiras: a experiência de implantação do SIM e do Sinasc. *Cienc Saude Coletiva.* 2007;12(3):643-54.
13. Menezes DCS, Leite IC, Schramm JMA, Leal MC. Avaliação da peregrinação anteparto numa amostra de puérperas no Município do Rio de Janeiro, Brasil, 1999/2001. *Cad Saude Publica.* 2006;22(3):553-9.
14. Perpétuo IHO, França-Mendonça E. Avaliação das estatísticas de nascimentos em Belo Horizonte 1974-1994. In: Anais do Encontro Nacional de Estudos Populacionais; 1996; Belo Horizonte, Brasil. Belo Horizonte: ABEP, 1996.
15. Schramm JMA, Szwarcwald CL. Sistema hospitalar como fonte de informações para estimar a mortalidade neonatal e natimortalidade. *Rev Saude Publica.* 2000;34(3):272-9.
16. Veras CMT, Martins MS. A confiabilidade dos dados nos formulários de Autorização de Internação Hospitalar (AIH), Rio de Janeiro, Brasil. *Cad Saude Publica.* 1994;10(3):339-55.