

Avaliação da qualidade dos dados do Sistema de Informações sobre Nascidos Vivos em 1997-1998

Evaluation of data quality from the Information System on Live Births in 1997-1998

Antônio Augusto Moura da Silva^a, Valdinar Sousa Ribeiro^b, Antônio Freitas Borba Júnior^a, Liberata Campos Coimbra^c e Raimundo Antonio da Silva^a

^aDepartamento de Saúde Pública da Universidade Federal do Maranhão (UFMA). São Luís, MA, Brasil. ^bDepartamento de Medicina III da UFMA. São Luís, MA, Brasil. ^cDepartamento de Enfermagem da UFMA. São Luís, MA, Brasil

Descritores

Estatísticas vitais.[#] Sistemas de informação.[#] Natalidade.[#] Registro de estatísticas vitais.[#] Recém-nascido de baixo peso. Prematuro.

Resumo

Objetivo

Verificar a concordância entre as informações constantes no Sistema de Informação sobre Nascidos Vivos (Sinasc) referentes a partos hospitalares e aquelas obtidas por inquérito seccional (padrão-ouro).

Métodos

Realizou-se inquérito por amostragem em 2.831 partos hospitalares ocorridos em dez maternidades do município de São Luís, MA, abrangendo cerca de 98% dos nascimentos hospitalares da cidade. O arquivo do Sinasc foi comparado com os dados do inquérito perinatal após a realização de *linkage* por programa de computador. Foram analisados: a cobertura estimada do Sinasc em relação aos nascimentos hospitalares verificados, o percentual de informação ignorada ou não preenchida e o grau de concordância entre as informações constantes no arquivo do Sinasc com o arquivo perinatal. Na medida da concordância, foram utilizados o indicador "kappa", no caso de variáveis qualitativas, e a correlação intraclass, em se tratando de variáveis quantitativas.

Resultados

A cobertura estimada do Sinasc foi de 75,8% (IC95%; 73,3%-78,2%). Os campos do Sinasc que apresentaram boa confiabilidade foram: peso ao nascer, sexo, hospital de nascimento, tipo de parto e idade materna. As estimativas do baixo peso ao nascer foram semelhantes nas duas pesquisas e a concordância foi alta (kappa=0,94). Entretanto, a prematuridade foi mais alta no inquérito (11,2%) do que no Sinasc (1,7%), e a concordância foi baixa (kappa=0,09).

Conclusão

A cobertura estimada do Sinasc foi baixa. A análise do arquivo do Sinasc mostrou que a cobertura estimada foi baixa; a taxa de baixo peso ao nascer parece ter boa validade e reprodutibilidade; e a taxa de prematuridade está subestimada.

Correspondência para/Correspondence to:

Antônio Augusto Moura da Silva
Departamento de Saúde Pública da UFMA
Rua Barão de Itapary, 155
65020-070 São Luís, MA, Brasil
E-mail: aasilva@elo.com.br

Trabalho realizado no Núcleo de Pesquisa em Saúde Materno-Infantil da UFMA. Financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - Processo nº 520664/98).

Recebido em 31/7/2001. Reapresentado em 21/8/2001. Aprovado em 12/9/2001.

Keywords

Vital statistics.[#] Information systems.[#]
Birth rates.[#] Vital statistics registry.[#]
Infant. Low birth weight. Preterm.

Abstract

Objectives

To verify the concordance between data routinely collected by the Sinasc (Information System on Live Births) and data obtained by a cross-sectional survey (gold standard).

Methods

The survey was performed in a sample of 2,831 hospital deliveries in ten maternity hospitals of the municipality of São Luís, Brazil. The sample frame represented about 98% of hospital births. Sinasc's data was compared with the survey data using a computerized linkage program. Sinasc's hospital births coverage, percentage of missing or unrecorded data and consistency between Sinasc's and the survey's data were analyzed. Concordance was measured using the kappa indicator for qualitative variables and the intraclass correlation coefficient for quantitative variables.

Results

The estimated Sinasc coverage was 75.8% (95% CI: 73.3%-78.2%). Five Sinasc's data fields showed good reliability: birth weight, newborn sex, hospital of birth, type of delivery and maternal age. Low birth weight rates estimated by the two data sources were similar and the concordance was high (kappa=0.94). However, preterm birth rates were higher in the survey (11.2%) than according to Sinasc's data (1.7%) and the concordance was low (kappa=0.09).

Conclusions

Estimated Sinasc coverage was low. Low birth weight rates from Sinasc's data seem to be valid and reliable. Preterm birth rates from Sinasc's data are underestimated.

INTRODUÇÃO

Os dados de nascidos vivos e de mortalidade são importantes para o planejamento e a avaliação das ações de saúde da criança no Sistema Único de Saúde (SUS), pois são usados no cálculo de vários indicadores de saúde, entre os quais os coeficientes de mortalidade infantil e materna. O Ministério da Saúde possui duas bases de dados que divulgam rotineiramente as informações coletadas e divulgadas sobre nascidos vivos (Sinasc) e mortalidade (SIM). Os dados de mortalidade são utilizados para o cálculo de algumas estatísticas vitais, e se pretende que o Sinasc, por ser baseado em informações geralmente coletadas no hospital, forneça estimativa mais confiável de nascidos vivos para o cálculo do coeficiente de mortalidade infantil, substituindo as estatísticas de nascidos vivos do registro civil que, em algumas regiões brasileiras, apresentam elevado grau de sub-registro.⁸ Pesquisas atuais sobre cobertura e grau de sub-registro do Sinasc são importantes para o seu aprimoramento.

O Sinasc é um sistema novo, implantado em 1990 no Brasil e em 1993 no Maranhão. A preocupação com a avaliação do Sinasc tem surgido em vários artigos recentes na literatura epidemiológica.^{6,8,10} É conhecido o problema de variáveis incompletas, o que compromete a fidedignidade dos dados e a elaboração de estatísticas confiáveis.⁹

A avaliação de qualidade do Sinasc tem pretendido calcular a cobertura obtida pelo sistema, quantificar o sub-registro, o grau de preenchimento dos campos da Declaração de Nascido Vivo (DN) e o percentual de informação ignorada.^{7,12} Mello Jorge et al⁸ (1996) mostraram que a cobertura do Sinasc no município de São Luís foi de 84,4%, em 1993; a percentagem de informação ignorada segundo variáveis constantes da DN foi menor com referência ao sexo (1,8%) e maior para filhos nascidos mortos (65,8%); dos 11 campos avaliados, apenas cinco tiveram informação considerada excelente, e quatro, má informação.

Recentemente, foi realizado em São Luís o projeto integrado de pesquisa "Saúde perinatal e atenção à saúde da mãe e da criança no município de São Luís", um inquérito seccional por amostragem abrangendo 2.831 partos em dez hospitais.¹¹ Desse modo, tem-se uma oportunidade de comparar os dados da DN com aqueles obtidos no presente inquérito por entrevistadores bem treinados, avaliando o grau de concordância entre a DN e a pesquisa perinatal, ou seja, de realizar uma avaliação da qualidade dos dados coletados pelo Sinasc.

No presente estudo, avaliou-se o grau de concordância entre as informações da DN e os dados obtidos por inquérito epidemiológico perinatal realizado em 1997-1998, utilizando-se a técnica de *linkage* de

banco de dados.¹ Estimou-se a cobertura global do Sinasc e por unidade de saúde, comparando-se os nascimentos com DN preenchida àqueles ocorridos nas dez maternidades da pesquisa perinatal. Comparou-se o percentual de campos com informação não preenchida ou ignorada nos dois arquivos (Sinasc e perinatal). Avaliou-se o grau de concordância entre as informações da DN e as informações da pesquisa perinatal, visando a identificar os campos da DN em que há problemas na qualidade do preenchimento ou da informação, para melhorar a qualidade do sistema de informação, de importância fundamental para o planejamento e a avaliação das ações de saúde da criança.

MÉTODOS

A pesquisa perinatal foi realizada em dez maternidades de São Luís, de primeiro de março de 1997 a 28 de fevereiro de 1998, representando cerca de 94% dos nascimentos do município. Considerando-se apenas os nascimentos hospitalares nessas dez unidades, o inquérito representou cerca de 98% do universo, incluindo-se nascidos vivos e mortos. A amostragem foi estratificada por hospital, com partilha proporcional ao número de partos, e baseou-se em todas as possíveis fontes de obtenção de registro de nascimentos no hospital. A partir de uma listagem contendo os nascimentos ocorridos em cada unidade, ordenados por hora do parto, foi realizada amostragem sistemática em cada estabelecimento com intervalo de amostragem sete. Para essa análise, todas as 2.831 mães de nascidos vivos foram entrevistadas logo após o parto, e foram coletados, entre outras informações, os dados também obtidos na DN. As perdas representaram 5,8%. Detalhes da metodologia podem ser obtidos em publicação anterior.¹¹

Foi construído um arquivo a partir dos dados do Sinasc, abrangendo nascimentos ocorridos nos mesmos meses de 1997 e 1998, nas dez maternidades da pesquisa perinatal, e excluídos dos arquivos os meses em que não houve a pesquisa perinatal. Pelo Sinasc e pela pesquisa perinatal, foram coletados dados dos nascimentos ocorridos em São Luís. Na presente análise, foram considerados apenas os nascidos vivos de famílias residentes no município de São Luís. No arquivo da pesquisa perinatal, foram excluídos os dados de nascidos mortos. Os dados dos campos a ser comparados foram recodificados para ficar exatamente iguais aos do arquivo do Sinasc.

A variável nome da mãe só constava em 69,1% dos registros do Sinasc, no período. Avaliando-se os registros mensais, observou-se que, nos meses de junho a setembro de 1997 e janeiro e fevereiro de 1998, constava o nome da mãe em 98,3% dos registros. Dessa

forma, optou-se por trabalhar com esse período de seis meses, no qual 9.228 nascimentos constavam do Sinasc, e 1.214 foram estudados pela pesquisa perinatal nas dez maternidades abrangidas pelo projeto.

Os campos utilizados no *linkage*, feito por programa de computador especialmente desenvolvido para esse fim, foram: data de nascimento, hospital, sexo e peso de nascimento, nessa ordem. Partindo-se dos 1.214 registros da pesquisa perinatal, buscava-se no arquivo do Sinasc registros com a mesma data de nascimento, mesmo hospital, mesmo sexo e mesmo peso de nascimento. Se houvesse discordância em pelo menos uma dessas variáveis e dois ou mais casos de crianças que nasceram na mesma data, no mesmo hospital, do mesmo sexo e com o mesmo peso de nascimento, o *linkage* não era realizado. Após esse processo, todos os *linkages* foram checados a partir de listagens dos dois arquivos, comparando-se o nome da mãe que constava nos arquivos da pesquisa perinatal e no do Sinasc. Os campos comparados foram: hospital de nascimento, sexo, peso ao nascer, idade gestacional, idade materna, número de fetos, tipo de parto, número de consultas pré-natais, escolaridade materna e número de nascidos vivos, nascidos mortos e abortos. O índice de Apgar não foi comparado, pois essa informação não foi obtida na pesquisa perinatal.

A cobertura do Sinasc era considerada muito boa quando atingia mais de 90% de cobertura, boa entre 81 e 90%, regular entre 71% e 80% e ruim quando igual ou menor a 70%.

Na análise do percentual de informações ignoradas ou em branco, foi utilizada a seguinte classificação (Mello Jorge et al,⁸ 1996): excelente, quando o percentual de ignorado/branco era menor que 10%; bom, entre 10% e 29,9%; e mau, quando igual ou superior a 30%.

Na análise da concordância, foram utilizados o indicador kappa para variáveis qualitativas e a correlação intraclass para variáveis quantitativas (peso ao nascer, idade materna e número de nascidos vivos, nascidos mortos e abortos).³ Em algumas comparações, foram descartados os casos sem preenchimento ou com informação ignorada nas duas bases de dados. Diferenças entre duas ou mais proporções foram testadas pelo qui-quadrado, assumindo-se nível de significância de 0,05.

Na pesquisa perinatal, a idade gestacional foi calculada a partir da data da última menstruação referida pela mãe. Quando o mês da última menstruação era ignorado, a data era considerada ignorada, e o dia

15 foi imputado para todos os casos nos quais apenas o dia da última menstruação era ignorado. Devido aos problemas descritos de erros na data referida da última menstruação, que tendem a provocar superestimativa da taxa de prematuridade,⁵ os casos de peso incompatível com a data da última menstruação referida que estivessem acima do percentil 99 da curva inglesa, foram recodificados como ignorados.² O mesmo procedimento foi empregado para os casos com idade gestacional improvável (menos de 20 ou mais de 50 semanas). Foi classificado com pré-termo o recém-nascido com idade gestacional menor que 37 semanas.

A análise estatística foi realizada nos programas Epi Info⁴ e Stata.¹³

RESULTADOS

Foi possível vincular corretamente pelo programa 72,4% dos registros. Pelo processo manual, 1,2% dos registros foram identificados como vinculados incorretamente, e mais de 4,6% foram vinculados pelo processo manual, totalizando sucesso no *linkage* de 920 registros, representando 75,8% (Intervalo de Confiança 95%; 73,3% a 78,2%). A cobertura estimada do

Sinasc apresentou ampla variação segundo a unidade hospitalar, de 60,1% a 95,8% (Tabela 1).

Seis campos no Sinasc tiveram elevado percentual de informação ignorada: idade gestacional, escolaridade materna, número de consultas pré-natais, número de nascidos vivos, mortos e abortos. Na pesquisa perinatal, apenas o campo idade gestacional teve percentual de informação ignorada acima de 5%. A concordância foi considerada boa para hospital de nascimento, sexo, tipo de parto, idade materna e peso ao nascer (Tabela 2). Para todas as variáveis, a concordância foi superior à que seria obtida ao acaso ($p < 0,001$).

Avaliou-se também a concordância em relação ao baixo peso ao nascer e à prematuridade entre a pesquisa perinatal e o Sinasc. No caso do baixo peso ao nascer, a concordância ajustada foi 0,94, e no da prematuridade, 0,15 (Tabela 3). Incluindo-se os casos com informação ignorada, a concordância para a prematuridade foi 0,08. Dos 14 casos de prematuridade identificados pelo Sinasc, oito eram verdadeiro-positivos, e seis, falso-positivos. Dos 97 prematuros catalogados no inquérito, o Sinasc não foi capaz de identificar 69% deles, resultando em 67 casos falso-negativos e 22 casos ignorados (Tabela 4).

Tabela 1 - Cobertura estimada do Sistema de Informação sobre Nascidos Vivos (Sinasc) por unidade hospitalar, São Luís, 1997-1998.

Hospital	n	%	Total	IC95%
Hospital Universitário	199	60,1	331	54,6-65,4
Benedito Leite	300	83,8	358	79,6-87,5
Santa Casa	174	85,3	204	79,7-89,9
Maria do Amparo	80	69,6	115	60,3-77,8
Luíza Coelho	53	86,9	61	75,8-94,2
São Marcos	46	95,8	48	85,7-99,5
São José	10	66,7	15	38,4-88,2
Unidade de Diagnóstico e Imagem	9	81,8	11	48,2-97,7
Nossa Sra da Penha	43	69,4	62	56,4-80,4
Estiva	6	66,7	9	29,9-92,5
Total	920	75,8	1.214	73,3-78,2

Tabela 2 - Percentual de campos com informação ignorada ou não preenchidos na pesquisa perinatal e no Sinasc e estimativas de concordância e correlação, São Luís, 1997-1998.

Variável	Campos com informação ignorada ou não preenchidos				Kappa	IC 95% *
	Inquérito perinatal		Sinasc			
	n	%	n	%		
Hospital	0	0,0	6	0,7	0,9795**	0,9444-1,0000
Sexo	0	0,0	4	0,4	0,9177**	0,8536-0,9818
Tipo de parto	0	0,0	5	0,5	0,9047**	0,8406-0,9687
Idade gestacional	57	6,2	114	12,4	0,0544*	0,0174-0,0913
Escolaridade materna	1	0,1	179	19,5	0,3000**	0,2667-0,3334
Número de consultas pré-natais	5	0,5	253	27,5	0,1281**	0,0907-0,1655
Número de fetos	0	0,0	12	1,3	0,5051**	0,4632-0,5469
					Correlação Intraclasse	IC95%*
Idade materna	0	0,0	31	3,4	0,8543**	0,7680-0,9406
Peso ao nascer	1	0,1	5	0,5	0,9544**	0,9402-0,9687
Número de nascidos mortos	0	0,0	679	73,8	0,7028**	0,2110-1,0000
Número de nascidos vivos	0	0,0	370	40,2	0,5306**	0,1470-0,9143
Número de abortos	0	0,0	629	68,4	0,6585**	0,1548-1,0000
Total	920	100,0	920	100,0	-	-

* $p < 0,01$ ** $p < 0,001$

cobertura muito boa (maior que 90%), quatro apresentaram cobertura boa (entre 81 e 90%), e as demais cinco, baixa cobertura (menor ou igual a 70%).

Dos 12 campos do Sinasc avaliados, seis tiveram um percentual considerado ótimo (menor que 10%) de informações ignoradas ou em branco: hospital de nascimento, sexo do recém-nascido, tipo de parto, número de fetos, idade materna e peso ao nascer. Três variáveis tiveram percentual entre 10% e 29,9%, considerado bom: idade gestacional, idade materna e número de consultas pré-natais. Para as demais variáveis, o percentual esteve acima de 30%, considerado ruim: número de nascidos vivos, mortos e abortos.

O grau de concordância ajustado medido pelo indicador kappa foi considerado ótimo (entre 0,81 e 0,99) para três variáveis: hospital, sexo e tipo de parto. Foi considerado regular (entre 0,41 e 0,60) para número de fetos, sofrível (entre 0,21 e 0,40) para escolaridade materna e fraco (menor ou igual a 0,20) para as demais variáveis: idade gestacional e número de consultas pré-natais.

No caso das variáveis quantitativas, a correlação intraclasse mostrou valor ótimo para o peso ao nascer e a idade materna. Para as demais variáveis, a correlação esteve entre 0,50 e 0,80: número de nascidos vivos, mortos e abortos. Nesses casos, porém, o percentual de informação ignorada para essas variáveis compromete o seu uso, pois a concordância foi medida descartando-se os casos ignorados.

A concordância ajustada em relação ao baixo peso ao nascer entre a pesquisa perinatal e o Sinasc foi 0,94, considerada excelente, indicando que as estimativas de baixo peso ao nascer obtidas a partir do Sinasc são de alta confiabilidade. No caso da prematuridade, a concordância foi ruim, de 0,09, indicando que o Sinasc não pode ser utilizado para o cálculo da taxa de prematuridade, pois há uma subestimação muito grande. Na amostra estudada, apenas 8,3% dos nascimentos prematuros foram corretamente identificados pelo Sinasc. Desse modo, há muitos casos de falso-negativos no Sinasc, que identificou 14 casos de prematuros, ao passo que a pesquisa perinatal identificou 97 casos, ou seja, cerca de sete vezes mais.

No caso da taxa de prematuridade, utilizou-se um procedimento na pesquisa perinatal para corrigir a superenumeração de prematuros, ou seja, descartar os casos com idade gestacional implausível ou incompatível com o peso de nascimento que foram reclassificados como ignorados. Tal critério reduziu o número de falso-positivos na identificação dos pré-termos, mas, entretanto, elevou o número de casos ignorados, que chegaram a 6,2% na pesquisa perinatal.

Não foram detectadas diferenças nas taxas de baixo peso ao nascer e prematuridade calculadas no arquivo completo da pesquisa perinatal e no arquivo de seis meses ($p < 0,05$), indicando que o procedimento de utilizar o período de seis meses não levou a viés de seleção.

Em conclusão, a cobertura estimada do Sinasc foi baixa: 75,8%. Os campos do Sinasc que apresentaram alta concordância com o inquérito e que, portanto, podem ser utilizados foram: peso ao nascer, sexo, hospital de nascimento, tipo de parto e idade materna. A taxa de baixo peso ao nascer estimada pelos dados do Sinasc possui alta confiabilidade. Entretanto, a taxa de prematuridade calculada a partir do sistema está subestimada. Novas investigações serão necessárias para avaliar se a subenumeração de nascimentos pré-termos observada no Sinasc em uma capital de Estado do Nordeste estaria também ocorrendo em outras localidades do Brasil.

Parece haver tendência a preencher o campo idade gestacional atribuindo-se a categoria de 37 a 41 semanas (termo) a crianças aparentemente “de termo”. Dessa forma estariam sendo consideradas prematuras apenas aqueles recém-nascidos indiscutivelmente “prematuros”, o que levaria a uma subestimação dos “prematuros limítrofes”, com 35 e 36 semanas de idade gestacional. De fato, quanto maior a idade gestacional entre os prematuros, maior o percentual de prematuros classificados como “a termo” ($p < 0,001$). Talvez a substituição no Sinasc da pergunta idade gestacional por data da última menstruação fosse mais apropriada, pois, apesar de problemas de validade desse indicador, trabalhos mostram que estimativas da idade gestacional baseadas na data da última menstruação têm boa reprodutibilidade.⁵ Ademais, a obtenção de uma estimativa da idade gestacional em semanas, e não em intervalos, permitiria aos pesquisadores a utilização de novas metodologias de correção de erros na data da última menstruação e um melhor ajuste para confundimento em modelos multivariáveis. Dessa forma, poderia ser obtida estimativa mais confiável da taxa de prematuridade no Brasil a partir das estatísticas de rotina.

Os resultados da presente investigação sugerem que mais atenção deve ser dispensada ao aprimoramento da coleta de dados de nascidos vivos nos hospitais. Alguns campos têm alta percentagem de informação ignorada ou não informada. Para outros campos, a confiabilidade foi baixa, o que pode ser devido a problemas relativos aos quesitos do instrumento (DN) ou à pouca importância que tem sido dispensada à informação em saúde no contexto da atenção à saúde e/ou à insuficiência das atividades de supervisão e controle de qualidade da coleta de dados do Sinasc.

REFERÊNCIAS

1. Almeida MF, Mello Jorge MHP. O uso da técnica de "linkage" de sistemas de informação em estudos de coorte sobre mortalidade neonatal. *Rev Saúde Pública* 1996;30:141-7.
2. Altman DG, Coles EG. Nomograms for precise determination of birthweight for dates. *Br J Obstet Gynecol* 1980;87:81-6.
3. Armitage P, Berry G. *Statistical methods in medical research*. Oxford: Blackwell; 1994.
4. Dean AG et al. *Epi Info, Version 6: a word processing, database, and statistics program for epidemiology on micro-computers*. Atlanta, Georgia: Centers of Disease Control and Prevention; 1994.
5. Kramer MS, McLean FH, Boyd ME, Usher RH. The validity of gestational age estimation by menstrual dating in term, preterm and postterm gestations. *JAMA* 1988;260:3306-8.
6. Maia MAC. Caracterização dos nascidos vivos hospitalares no primeiro ano de implantação do Subsistema de Informações sobre Nascidos vivos, em município de Minas Gerais, Brasil, 1996. *Rev Saúde Pública* 1997;31:581-5.
7. Mello Jorge MHP, Gotlieb SLD, Soboll MLMS, Almeida MF, Latorre MRDO. Avaliação do Sistema de Informações sobre Nascidos Vivos e o uso de seus dados em epidemiologia e estatísticas de saúde. *Rev Saúde Pública* 1993;27Supl 6:1-46.
8. Mello Jorge MHP, Gotlieb SLD, Oliveira H. O Sistema de Informações sobre Nascidos Vivos: primeira avaliação dos dados brasileiros. *Inf Epidemiol SUS* 1996;5:15-48.
9. Oliveira H, Pereira IPA. Estatísticas de mortalidade e nascidos vivos: considerações sobre principais problemas. *Inf Epidemiol SUS* 1997;6:15-9.
10. Rodrigues CS, Magalhães Júnior HM, Evangelista PA, Ladeira RM, Laudares S. Perfil dos nascidos vivos no município de Belo Horizonte, 1992-1994. *Cad Saúde Pública* 1997;13:53-7.
11. Silva AAM, Coimbra LC, Silva RA, Alves MTSSB, Lamy Filho F, Lamy ZC, et al. Perinatal health and mother-child health care in the municipality of São Luís, Maranhão, Brazil. *Cad Saúde Pública* 2001;17 (in press).
12. Silva RI, Theme Filha MM, Noronha CP. Sistema de informação sobre nascidos vivos na cidade do Rio de Janeiro, 1993/1996. *Inf Epidemiol SUS* 1997;6:33-48.
13. Stata Corporation. *Stata reference manual release 5*. Texas: Stata Corporation; 1997.
14. Tonial SR, Silva AAM, organizadores. *Saúde, nutrição e mortalidade infantil no Maranhão*. São Luís: UFMA/Secretaria de Estado da Saúde/Unicef; 1997.