

# **A qualidade dos registros de nascimentos e as condições de vida do Semiárido brasileiro: construção de indicadores sintéticos**

## **Resumo:**

As variáveis da natalidade representam importantes sinalizadores para identificar as condições de saúde das mães, dos recém-nascidos e subsidiam o planejamento e as políticas públicas. No Brasil, a região do Semiárido é marcada historicamente por grandes desigualdades socioeconômicas e com deficiências nas estatísticas de nascimentos. Estas últimas são conhecidamente subenumerados e a qualidade do preenchimento das variáveis nas declarações de nascimentos não é completa. Teve-se por objetivo estudar a relação entre a qualidade dos registros de nascimentos e as condições de vida na região Semiárida brasileira, para o ano 2010, através da criação de indicadores sintéticos. Para tal propósito fez-se uso de técnicas demográficas de estimação para avaliar a qualidade dos registros de nascimentos e de técnicas estatísticas multivariadas para estudar o relacionamento entre qualidade dos dados e as condições de vida. Para o estudo foi contemplado as 137 microrregiões. Os dados foram extraídos do SINASC - Ministério da Saúde e as variáveis referentes aos fatores socioeconômicos foram retirados da PNUD. Os resultados apontaram para níveis elevados da completude de várias variáveis para boa parte das microrregiões do Semiárido e elevada cobertura dos nascimentos em 2010. Foram propostos um Indicador Sintético de relacionamento e um Índice de Qualidade das Informações. Embora haja espaço para melhorias, os resultados sugerem um acelerado processo de universalização das estatísticas de nascimentos. Foram identificados grupos de microrregiões com relacionamentos diferenciados entre a qualidade dos dados e as condições de vida, cuja regionalização pode ser útil para uma melhor definição das prioridades no planejamento das ações públicas para a saúde materno-infantil da região Semiárida brasileira.

**Palavras-chave:** Nascimentos, condições de vida, indicadores sintéticos.

## INTRODUÇÃO

Segundo dados oficiais do Ministério da Integração, a região semiárida brasileira abrange uma área de 969.589,4 km<sup>2</sup>, a maior do mundo em termos de extensão territorial e populacional que corresponde a 12% do território nacional e compreende 1.133 municípios de nove Estados do Brasil: todos os estados do Nordeste, exceto Maranhão e o norte de Minas Gerais (IBGE, 2011).

A qualidade das informações e as condições em que ocorrem os nascimentos têm sido uma questão central e muito importante para a consistência dos dados. A falta de informações confiáveis é apenas um dos motivos, talvez o mais importante, que inviabiliza o conhecimento do padrão de natalidade e fecundidade de determinada região.

Sabe-se que os distintos Estados da Região Semiárida têm apresentado variações no comportamento relacionado à natalidade. Mas, pouco se sabe sobre essas variações, principalmente quanto ao nível ou magnitude. Nesse sentido, apontar um diagnóstico descritivo e analítico das regiões, características das microrregiões e das populações revela-se em uma etapa fundamental que possam auxiliar no planejamento e na avaliação de ações que visem minimizar ou erradicar as situações de pobreza e as desigualdades regionais que ocorrem nas diferentes microrregiões do Semiárido.

É possível que devido à ampla diversidade do Semiárido, as microrregiões guardem entre si relações de similaridade ou padrões de relacionamento dos nascimentos com as condições de vida. Para tal investigação, a técnica multivariada *Cluster Analysis* e a regressão linear, são ferramentas estatísticas, que possibilitam investigar e identificar a existência desses padrões de relacionamentos que permitem reduzir os dados sobre Nascidos Vivos em grupos, não sacrificando boa parte das informações contidas nas variáveis originais. Configura-se, ainda, como problema a ser superado com base nos resultados desta pesquisa a carência de um Indicador Estatístico Sintético que avalie a dialogicidade entre as condições de vida e os nascimentos na região Semiárida brasileira.

Nesse sentido, esse estudo se sustenta pela necessidade de conhecer e avaliar a qualidade das informações geradas pelos sistemas de saúde de uma região pouco estudada. Destaca-se também a necessidade de buscar elementos explicativos que possam associar a qualidade desses dados com as condições de vida da população. Desse modo, espera-se traçar um panorama da situação dessas estatísticas, que possam

subsidiar os gestores de políticas de saúde ao adotar medidas mais efetivas na rede assistencial do Semiárido.

## **METODOLOGIA**

### *BASE E FONTES DE DADOS*

As fontes oficiais de dados contínuos são o Instituto Brasileiro de Geografia e Estatística (IBGE; <http://www.ibge.gov.br>) e o Sistema de Informações sobre Nascidos Vivos (SINASC) do Departamento de Informática do SUS (<http://tabnet.datasus.gov.br>) do Ministério da Saúde. Investigações preliminares apontaram que no ano 2010 o SINASC captou um número maior de registros, além de disponibilizar um número maior de variáveis. Essa fonte desde 2000 já se afirmava como muito útil em informações por permitir estimativas de indicadores da fecundidade e de natalidade.

Dessa forma os dados básicos sobre nascimentos foram coletados através do Banco de Dados do SINASC. Este sistema objetiva não somente levantar o número real dos nascidos vivos, mas também, variáveis importantes para analisar as condições de nascimento que estão presentes na DN. As variáveis pesquisadas da DN se referem a: *anomalia congênita, apgar de primeiro minuto, apgar de quinto minuto, consultas pré-natal, cor/raça, duração da gestação, estado civil da mãe, idade da mãe, instrução da mãe, local de ocorrência, peso ao nascer, sexo, tipo de anomalia congênita, tipo de parto e tipo de gravidez*. Para fins comparativos para cada Estado do Semiárido foram investigados os dados referentes aos anos de 2000 e 2010.

Pelo Sistema IBGE de Recuperação Automática (SIDRA), foram acessados os códigos atribuídos a cada município brasileiro e, com eles, foi construída uma base de dados relativa aos municípios do Semiárido. As bases foram organizadas utilizando o programa Microsoft Excel 2007.

Cada um dos Estados do Semiárido, assim como os demais Estados brasileiros, está dividido em mesorregiões geográficas e, estas, por sua vez, divididas em microrregiões geográficas em um total de 137. Inicialmente, foram coletados os dados referentes a todos os municípios do Semiárido, tendo em vista que o Ministério da Saúde não classifica os dados especificamente para esta área. Após extrair as tabelas do DATASUS referentes a cada variável, tais dados foram devidamente agrupados por microrregiões geográficas.

Os montantes dos nascimentos ignorados foram extraídos para cada variável pesquisada. Através do total de ignorados e do total de nascimentos, foi calculado o percentual de ignorados para cada microrregião, o que permite uma avaliação da variação de incompletitude na década. É importante ressaltar que o número de ignorados não foi registrado em todas as variáveis

pesquisadas pelo SINASC, criando uma discrepância entre os anos pesquisados. Para a classificação das variáveis foi utilizada a seguinte escala: a) Abaixo de 10%, classificado como muito bom; b) de 10% a 19%, classificado como bom; c) de 20% a 29%, classificado como regular; e d) acima de 30%, classificado como ruim.

#### *COBERTURA DOS NASCIDOS VIVOS*

A má qualidade dos dados e o subregistro dos nascimentos fazem com que levantamentos a esse respeito sejam mais difíceis de serem implementados. Para minorar esses problemas, os demógrafos desenvolveram um conjunto de procedimentos para estimação da cobertura dos nascimentos e análise de dados.

Para investigação do sub-registro de nascimentos recorreu-se ao cálculo da cobertura dos nascidos obtido através da relação entre o valor observado e o valor esperado. O número de nascidos vivos esperados foi obtido através do produto do número de mulheres por grupos de idade quinquenais de 15 a 49 anos pelas correspondentes taxas específicas de fecundidade do Estado, a qual cada microrregião pertence. Admitiu-se, assim, que o padrão etário de fecundidade do Estado é similar ao da microrregião a ele pertencente no mesmo ano. Assim, a cobertura dos nascidos vivos é dada pela seguinte fórmula:

$$C_i = \frac{NV_{obs}(i)}{NV_{est}(i)} * 100$$

Em que,

$$NV_{est}(i) = \sum_{j=15}^{49} TEF_{j(i)} * TM_j$$

onde,

$C_i$  = cobertura de nascidos vivos da região  $i$ ;

$NV_{obs}(i)$  = total de nascidos vivos observados na região  $i$ ;

$NV_{est}(i)$  = total de nascidos vivos estimados para região  $i$ ;

$TEF_{j(i)}$  = taxa específica de fecundidade na faixa etária quinquenal  $j$  da região  $i$ ;

$TM_j$  = total de mulheres na faixa etária quinquenal  $j$  da região  $i$ ;

$j$  = faixa etária quinquenal de 15 a 49 anos.

Adotou-se no Quadro 1 a seguinte classificação para categorizar a qualidade da cobertura dos nascidos vivos.

Quadro 1: Classificação das coberturas dos nascidos vivos, segundo o SINASC, 2010.

Cobertura (%)	Classificação
> 90	Muito Boa
81 a 90	Boa
71 a 80	Regular
≤ 70	Ruim

Fonte: Laboratório de Estudos Demográficos (LED), Departamento de Estatística - UFPB.

A qualidade dos dados dos nascidos vivos foi explorada de maneira global ao construir o *Índice de Qualidade das Informações Ignoradas dos Nascimentos*, aqui denominado de *IQII*.

É importante ressaltar que os percentuais de ignorados das variáveis da DN apresentados não correspondem ao número real de ignorados, já que essas informações somente dizem respeito aos nascimentos que foram registrados oficialmente. Para o volume de nascimentos não registrados, ou seja, os sub-registrados, todas as variáveis são evidentemente consideradas como ignoradas. Desta forma, para obter uma ideia mais próxima do montante real dos “ignorados” faz-se necessário encontrar um indicador que capte ambos componentes: os registrados, mas com informações ignoradas, e os não registrados.

Este índice consiste na adição do percentual do sub-registro dos nascimentos com o produto da cobertura dos nascimentos pela média dos percentuais de ignorados das variáveis estudadas (cujos percentuais foram os mais elevados): *anomalia congênita, apgar de primeiro minuto, apgar de quinto minuto, consultas pré-natal, cor/raça, estado civil da mãe e instrução da mãe*. Dessa maneira, o índice é dado pela seguinte fórmula:

$$IQII = (100 - C) + (C \cdot I_M)/100$$

*IQII* : Percentual do *Índice de Qualidade das Informações Ignoradas*

C : Percentual de cobertura dos óbitos

$I_M$  : Média do Percentual de Ignorados das variáveis da Declaração de Nascimentos

#### *RELAÇÃO ENTRE O IDH E A QUALIDADE DOS DADOS*

A técnica multivariada *Cluster Analysis* foi utilizada para identificar grupos de microrregiões, tendo como variáveis o IDH, a cobertura dos nascimentos e as proporções de ignorados das variáveis selecionadas. O relacionamento entre estas variáveis foi testado admitindo-se um nível de significância  $p < 0,05$ . Para a aplicação da técnica, a análise foi iniciada com a transformação das variáveis para a escala normal. Em seguida, foi elaborado o *Dendograma*, estabelecendo-se o critério de distância euclidiana de 5 pontos para a

formação/identificação dos grupos. Foram traçados gráficos de ajuste da reta para verificar a normalidade das variáveis. Após a identificação dos pontos destoantes (outliers) dos gráficos para os grupos formados, eles foram eliminados. Em seguida, procedeu-se ao ajuste da reta de regressão para cada grupo tendo como intercepto o IDH, a cobertura e as proporções de ignorados das variáveis.

### *INDICADOR SINTÉTICO*

Com a formação final do agrupamento das microrregiões, criou-se um Indicador Síntese para expressar a relação entre a qualidade dos dados de nascimento e o IDH. O uso de técnicas multivariadas fez-se necessário para o cálculo deste indicador que correspondente aos grupos formados por microrregiões, as quais foram divididas em dois estratos e hierarquizadas internamente.

O *Indicador sintético* é entendido aqui como um valor derivado de outros indicadores, que indica ou descreve o grau de relacionamento entre o IDH e a qualidade dos dados de nascimento para as microrregiões. A metodologia para a criação do indicador sintético foi baseada na que foi utilizada para o cálculo do IDH, que leva em consideração três elementos, renda, longevidade e educação. Para cada um desses elementos é calculado um índice, utilizando-se a média geométrica de seus valores. Por fim, calcula-se a média geométrica desses índices e é obtido o IDH. A qualidade dos registros de nascimentos é representada aqui pela média dos percentuais de variáveis ignoradas e média da cobertura dos nascidos vivos das microrregiões de cada estrato que foi identificado pela metodologia explicitada no item anterior.

Para classificar o índice criado, foi utilizada a mesma categorização utilizada para o IDH, como mostra o Quadro 2.

Quadro 2 – Classificação do IDH segundo o Atlas de Desenvolvimento Humano, 2013.

<b>IDH</b>	<b>Classificação</b>
0 a 0,499	Muito Baixo
0,500 a 0,599	Baixo
0,600 a 0,699	Médio
0,700 a 0,799	Alto
0,800 a 1	Muito Alto

Fonte: Atlas do Desenvolvimento Humano, 2013.

As etapas realizadas para construção do Indicador de interesse são descritas a seguir.

#### *i) Agrupamento das microrregiões*

As microrregiões contidas no Semiárido possuem características distintas entre si. Essas microrregiões foram identificadas e agrupadas em conjuntos distintos para em seguida serem analisadas separadamente. A ferramenta usada foi a técnica de *cluster analysis*, muito utilizada quando se deseja agrupar variáveis que estão fortemente correlacionadas.

ii) *Peso das variáveis nos fatores*

O *Indicador Sintético* foi construído através de uma média ponderada dos indicadores que o compõe, onde estes são calculados pela média aritmética ponderada, dada pela fórmula:

$$\text{Indicador Sintético} = \frac{(1 \times M_{IDH}) + (3 \times M_{IDN}) + (6 \times M_{CO})}{10}$$

Em que,

$M_{IDH}$  = média do IDH para os grupos

$M_{IDN}$  = média das variáveis ignoradas da DN para os grupos

$M_{CO}$  = média das coberturas dos grupos

Considerando a cobertura dos nascimentos ( $M_{CO}$ ) como o indicador de grau mais importante, foi atribuído a ele um peso 6, o dobro do indicador  $M_{IDN}$ , com peso 3, dessa forma o IDH ficou com peso 1.

Ao serem obtidos os valores dos *Indicadores Sintético das Condições de Vida e da qualidade dos dados de nascimentos* foi possível ordenar as Microrregiões, sendo estas dispostas de acordo com os níveis dos escores em ordem decrescente. As Microrregiões foram agrupadas em dois grupos, correspondentes às seguintes categorias ou estratos de condição de vida associada à qualidade dos dados de nascimento: (I) IDH Baixo com Alta qualidade dos dados; (II) IDH Baixo com Baixa qualidade dos dados.

Com a estratificação das Microrregiões do Semiárido brasileiro por meio do Indicador Sintético das Condições de Vida e da qualidade dos dados de nascimentos, tornou-se possível identificar as regiões com melhores e piores condições de vida associada às qualidades dos dados.

## RESULTADOS E DISCUSSÃO

Nos últimos anos, tem-se observado uma mudança no perfil de nascimentos no Brasil. Segundo Berquó e Cavenaghi (2006), a educação e o aumento da renda das mulheres afetam negativamente a fecundidade. Ou seja, a melhora no grau de instrução e da inserção das

mulheres no mercado de trabalho, fatores que colaboram para sua independência, foi fundamental na redução do número de nascimentos observado na última década no Brasil.

Apesar do Semiárido se constituir em uma das regiões mais pobres e menos desenvolvidas do país, a magnitude das TBN do Semiárido para 2000 (24,6) e 2010 (17,2) foi um pouco acima da TBN do Brasil, respectivamente com 20,3 e 15,8 por 1000 indivíduos (RIPSA, 2014). Ainda segundo a RIPSA, a comparação com o Nordeste mostra que estes valores foram: 22,6 e 17,3 por 1000 indivíduos, respectivamente para 2000 e 2010. Em 2000 a TBN do Semiárido era superior à do Nordeste, já em 2010, o diferencial foi praticamente anulado. Os diferenciais entre esta taxa para o Semiárido e o Nordeste ao ser comparada com a do Brasil diminuíram na década, sinalizando para uma convergência.

A Tabela 1 mostra os dados brutos de nascimentos e as TBN para o espaço Semiárido dos Estados. A evolução temporal mostra um importante decréscimo tanto no volume de nascimentos como nas TBN na década para todos os espaços semiáridos dos Estados, exceto a Bahia que se manteve no mesmo patamar, mesmo havendo uma redução numérica do número de nascimentos. No total, a redução na década foi da ordem de 40% na TBN, alcançando em 2010 uma taxa de 17,2 nascimentos para cada 1000 indivíduos.

Tabela 1: Números de Nascimentos e Taxa Bruta de Nascimentos do espaço Semiárido, por Estado, 2000 e 2010.

Estados	Número de Nascimentos		TBN (‰)	
	2000	2010	2000	2010
Piauí	16764	14045	24,9	13,4
Ceará	85186	73292	24,5	15,5
Rio Grande do Norte	49743	25708	22,9	14,5
Paraíba	33635	31952	22,7	15,2
Pernambuco	70522	61388	20,6	16,7
Alagoas	22021	16544	28,6	18,3
Sergipe	8476	6905	25,2	15,6
Bahia	167204	102448	23,0	23,1
Minas Gerais	17295	16913	19,4	13,7
TOTAL	470846	349195	24,6	17,2

Fonte: Laboratório de Estudos Demográficos, Departamento de Estatística – UFPB.

## COBERTURA DOS NASCIMENTOS

Tendo em vista a necessidade de obtenção de parâmetros para avaliar a cobertura e a qualidade dos dados gerados pelos sistemas de informações no Brasil, propuseram-se indicadores de acompanhamento de sua evolução. No decorrer dos anos, porém, com a melhoria dos sistemas de informações, as fontes de dados utilizadas para definir os padrões de comparação mostraram-se insuficientes.

A importância da proposição de novas formas de avaliar a cobertura dos nascimentos encontra, neste estudo de caso, uma estratégia privilegiada. A partir da identificação do número observado de nascimentos ocorridos em 2010 no Semiárido brasileiro, foi possível o estabelecimento de uma comparação entre a cobertura dos nascimentos entre a estimativa realizada pelo DATASUS e a encontrada neste trabalho para 2010. As estimativas do DATASUS se referem ao Estado como um todo. As coberturas dos nascidos vivos estimadas para as áreas geográficas dos estados do Semiárido são ilustradas na Tabela 2.

Tabela 2- Cobertura dos Nascidos Vivos por estado da área geográfica do semiárido brasileiro comparada com a do DATASUS, 2010.

<b>Estado</b>	<b>DATASUS (%) *</b>	<b>Região semiárida (%)</b>
Piauí	93,7	92,1
Ceará	93,5	91,9
Rio Grande do Norte	93,6	92,9
Paraíba	94,2	93,0
Pernambuco	94,9	94,5
Alagoas	92,5	91,5
Sergipe	95,4	94,8
Bahia	92,7	90,3
Minas Gerais	96,5	95,5

Fonte: DATASUS (2010)

\* Refere-se ao Estado como um todo

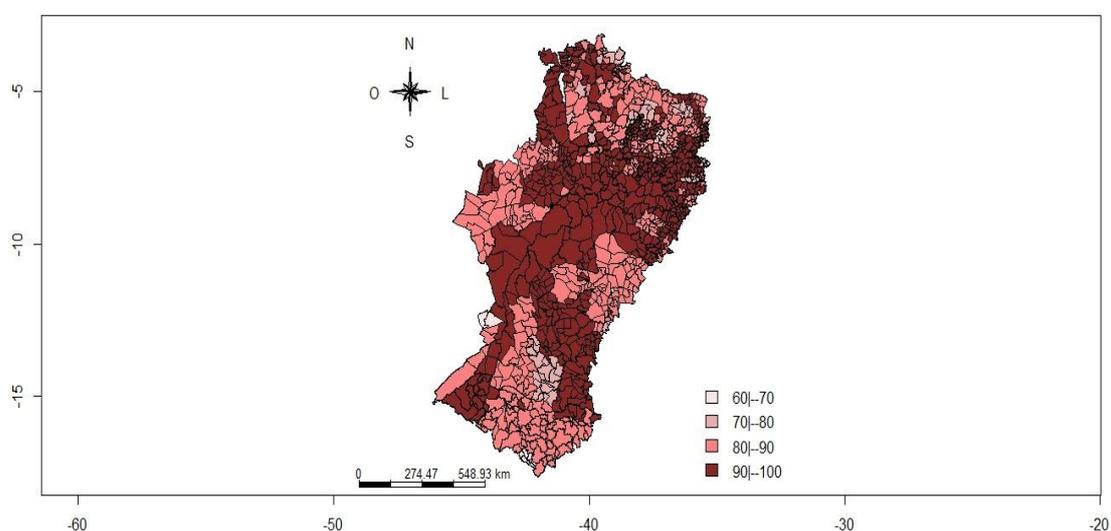
De modo geral, as coberturas das duas estimativas foram muito próximas, com valores um pouco maiores para o DATASUS. As diferenças entre os resultados produzidos aqui e a do Ministério da Saúde não ultrapassaram 2,7%. Os Estados de Pernambuco, Sergipe e Rio Grande do Norte, apresentaram o mesmo patamar de cobertura em relação à estimativa do DATASUS e a encontrada neste trabalho para o espaço geográfico do Semiárido.

Segundo a Organização Mundial de Saúde pode-se considerar como de boa qualidade as estatísticas vitais que capturem acima de 90% dos eventos analisados. Deste modo pode-se dizer que as coberturas para as regiões Semiáridas e para o Estado como um todo pode ser classificado como de boa qualidade, entretanto quando comparadas a

de Estados da região Sudeste, como por exemplo, os Estados de Espírito Santo, Rio de Janeiro e São Paulo, onde a cobertura dos nascimentos já atingiu o patamar de 100%, em 2010, nota-se que o Nordeste e principalmente o Semiárido, ainda tem espaço para avançar até atingir a universalização dos registros dos nascimentos ocorridos.

A distribuição espacial da cobertura dos nascimentos para os municípios do Semiárido brasileiro referente ao ano de 2010 foi mapeada conforme mostra a Figura 1. Ressalta-se que para construção do mapa, os municípios receberam a mesma cobertura das microrregiões. No entanto, quando este espaço é desagregado ao nível municipal, as diferenças ficam mais evidentes.

Figura 1 - Distribuição espacial da Cobertura dos Nascidos Vivos por municípios do Semiárido brasileiro, 2010.



Fonte: Laboratório de Estudos Demográficos – LED – DE/UFPA

Observa-se que a maior parte dos municípios teve uma cobertura acima de 90%, sendo classificada como *ótima*. Entretanto para algumas áreas nota-se que a cobertura foi maior que 90% (faixa de cor mais escura). Essas áreas perpassam as áreas mais centrais do Semiárido e atingem algumas microrregiões mais à direita, ao Nordeste do Semiárido. Compreendem as microrregiões que estão mais próximas das capitais, como por exemplo, a microrregião de Sertão de Cratús, no Estado do Ceará com 94,5%, a microrregião de Petrolina, no Estado de Pernambuco com 94,7% e a microrregião de Grão Mogol, no Estado de Minas Gerais com 94,7%.

#### *QUALIDADE DO PREENCHIMENTO DA DN*

O estudo da completude, expresso pelo percentual de ignorados para cada variável, permite aprofundar os conhecimentos sobre a qualidade dos dados de nascidos vivos. Com o passar dos anos espera-se que as proporções de ignorados diminuam, ou seja, esperam-se melhoras na qualidade do preenchimento das variáveis dos registros de nascimentos.

A análise do SINASC, no período compreendido entre 2000 a 2020, para a região semiárida, permitiu identificar a ocorrência de percentuais de variáveis ignoradas contidas na DN.

Avaliando-se o percentual de ignorados para cada variável pesquisada na DN, foi possível identificar aquelas com menor completude. De uma maneira geral, os dados relacionados ao *período gestacional* e ao *trabalho de parto*, possuíram maior cobertura, possivelmente devido a sua importância na assistência ao parto e na atenção ao neonato no período pós-parto imediato. Excluídas essas duas variáveis de estudo, as variáveis avaliadas foram: *cor/raça*, *instrução da mãe*, *apgar de primeiro minuto*, *apgar de quinto minuto*, *número de consultas pré-natais* e *estado civil da mãe*.

Observa-se na Tabela 3 a evolução da completude dos registros na região semiárida na década. Em 2000, duas das variáveis foram classificadas como *regular*, duas como *bom* e duas como *muito bom*. Já em 2010, todas estas variáveis tiveram *boa* completude, com um percentual de ignorados inferior a 10%.

Tabela 3 - Percentual de ignorados das variáveis selecionadas da Declaração de Nascidos Vivos no espaço geográfico do semiárido brasileiro, 2000.

Estado	Variáveis relacionadas ao recém-nascido (%)				Variável relacionada à mãe				Variáveis Sociais			
	Apgar 1º minuto		Apgar 5º minuto		Consultas Pré-natais		Cor/Raça		Estado Civil da Mãe		Instrução da Mãe	
	2000	2010	2000	2010	2000	2010	2000	2010	2000	2010	2000	2010
Piauí	49	10	54	10	2	1	5	0	2	1	6	1
Ceará	12	1	17	1	2	0	7	4	2	1	5	1
Rio G. Norte	1	0	1	0	1	1	2	1	1	2	2	1
Paraíba	20	3	26	4	5	1	15	1	4	1	9	3
Pernambuco	17	3	18	1	51	1	27	1	34	1	45	1
Alagoas	40	18	41	21	1	2	1	2	1	1	2	2
Sergipe	12	11	17	12	9	1	18	8	8	6	13	1
Bahia	32	13	33	14	8	2	18	12	5	3	9	5
Minas Gerais	14	14	16	14	5	1	15	6	5	5	7	3
Total	22	8	25	9	9	1	12	4	7	2	11	2

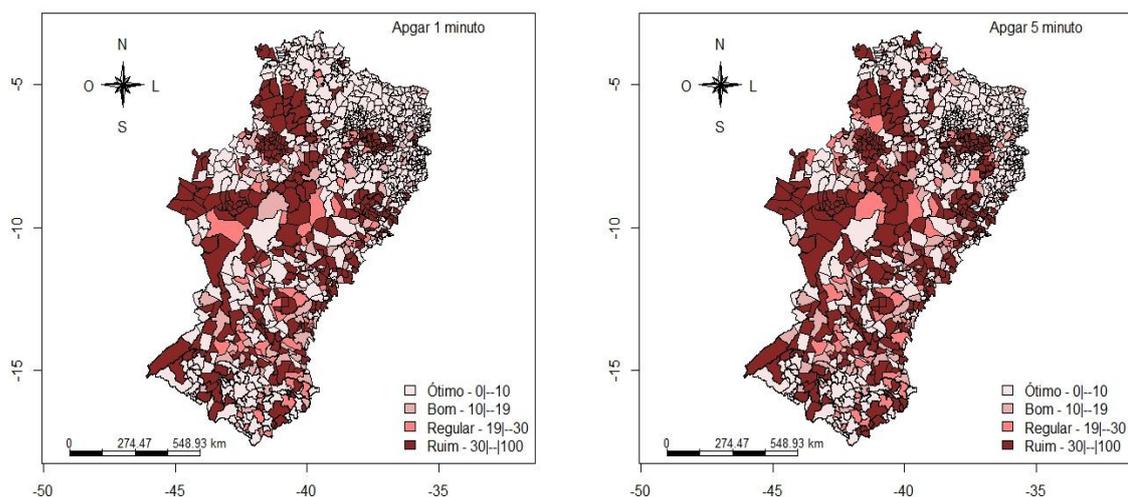
Fonte: Laboratório de Estudos Demográficos, Departamento de Estatística – UFPB.

Classificação: abaixo de 10%: muito bom; entre 10 e 19%: bom; entre 20 e 30%: regular; acima de 30%: ruim.

Ao se observar o percentual de variáveis ignoradas, para os Estados, nota-se que os valores não superam 21%, sendo classificada como *regular*. De modo geral, o percentual de ignorados do Semiárido apresentou classificação como *boa* para o ano de 2010, por apresentar em sua maioria os percentuais entre 10% a 19%, assim mostrando uma melhoria em relação aos resultados para o ano 2000, já para as variáveis: estado civil da mãe, instrução da mãe, consulta pré-natal, apgar de 1º minuto e apgar de 5º minuto, a classificação foi considerada *ruim*.

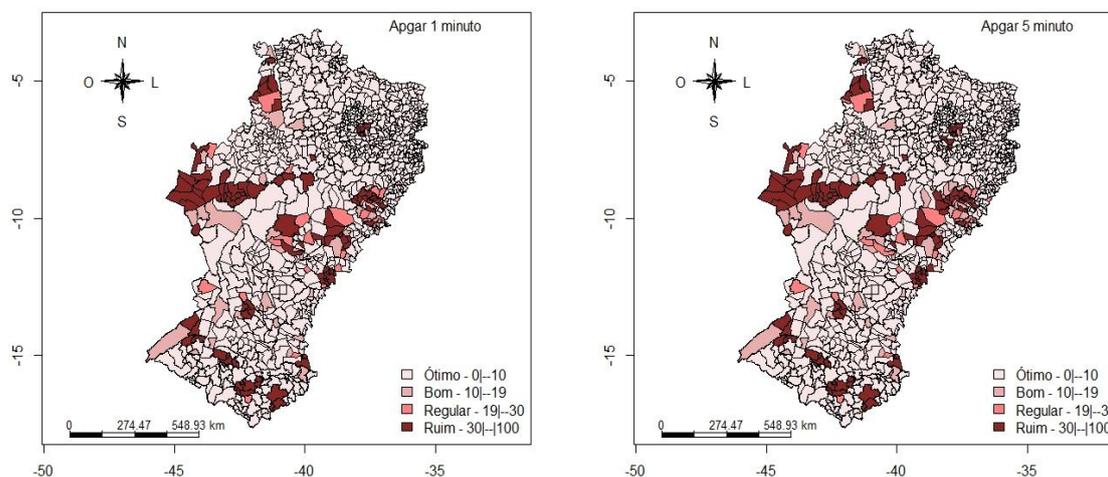
Das variáveis selecionadas as que apresentaram maior percentual de ignorados foram *apgar de primeiro minuto* e *apgar de quinto minuto*, mesmo assim apresentaram uma diminuição na década, conforme mostrado nos mapas da Figura 2 e a Figura 3.

Figura 2- Variação do percentual de ignorados das variáveis *apgar de 1º minuto* e *apgar de 5º minuto* para a década de 2000 no Semiárido Brasileiro.



Fonte: Laboratório de Estudos Demográficos – LED – DE/UFPB

Figura 3- Variação do percentual de ignorados das variáveis *apgar de 1º minuto* e *apgar de 5º minuto* para a década de 2010 no Semiárido Brasileiro.



Fonte: Laboratório de Estudos Demográficos – LED – DE/UFPB

Dentre os Estados do semiárido, o Rio Grande do Norte foi o que apresentou os menores percentuais de ignorados, tanto no ano de 2000, como em 2010. O Estado de Pernambuco, por outro lado, apresentou os piores percentuais no ano de 2000 e apresentou aumento considerável na completude no ano de 2010.

#### *INDICE DE QUALIDADE DE INFORMAÇÕES IGNORADAS DA DN*

O Indicador de Qualidade de Registro de Nascimento foi calculado com o intuito de se obter uma estimativa mais próxima possível da qualidade das informações sobre os nascimentos. O Quadro 3 mostra que o IQII foi maior para o Grupo 1, onde o subregistro encontrado foi superior a 19%. Com isso nota-se que ainda o erro referente aos nascimentos pode ser considerado ainda preocupante, pois em ambos os Grupos o escore do valor real de subregistro foi maior que 10%; isso significa que apresentou uma classificação considerada *regular*. Também foi calculado o número de nascimentos estimados para os grupos em relação ao novo indicador encontrado, e como resultado obteve-se valores maiores de nascimentos para os grupos uma vez que o número do subregistro encontrado nesse trabalho foi maior que o apresentado pelo DATASUS.

Quadro 3- Indicador de Qualidade de Informações Ignoradas das microrregiões do Semiárido brasileiro, 2010.

Grupos	Número de microrregiões	Cobertura Média	1 - Média dos Ignorados	<i>IQII</i>	NV observados	NV estimados
1	89	0,9145	0,9761	0,1976	224341	302567
2	13	0,9226	0,9563	0,1256	53294	78987
3	35	0,9087	0,9876	-	73241	-

Fonte: Laboratório de Estudos Demográficos - LED - DE/UFPB

A média percentual de subregistro no Brasil foi de 8%. Sendo assim, os grupos formados pelas Microrregiões do Semiárido ainda se encontra em situação inferior com relação às regiões mais desenvolvidas, contudo, vale salientar que houve uma grande melhora em relação aos anos anteriores.

#### *INDICADOR SINTÉTICO DA QUALIDADE DA DN E CONDIÇÕES DE VIDA*

Após a aplicação da técnica de Cluster Analysis, foram formados dois grupos de microrregiões (Quadro 4) cujo relacionamento do IDH com a qualidade dos dados foi significativa ( $p < 0,05$ ). A qualidade dos dados foi representada pela cobertura dos nascimentos e pelo percentual das variáveis selecionadas com informações ignoradas. O primeiro grupo reuniu 89 microrregiões e o segundo, 13. Para as demais microrregiões não foram obtidos resultados significativos, ou seja, 31 microrregiões ficaram excluídas da formação de algum grupo, por não evidenciarem uma relação significativa entre as variáveis estudadas. Os resultados são apresentados em termos de grupos de microrregiões, conforme mostra o Quadro 3.

Quadro 4 - Indicador síntese ligado ao relacionamento entre o IDH e a qualidade dos dados de nascimentos das microrregiões do Semiárido brasileiro, 2010.

Grupos	Número de microrregiões	IDH Médio	Cobertura Média	1 - Média dos Ignorados	P- valor	R <sup>2</sup>	<b>Indicador Síntese</b>
1	89	0,5999	0,9145	0,9761	0,0008*	0,1917	0,6843
2	13	0,5542	0,9226	0,9563	0,0001*	0,33	0,6912
3	35	0,5098	0,9087	0,9876	0,17896	0,0047	-

Fonte: Laboratório de Estudos Demográficos - LED - DE/UFPB

\*Valores significativos ao considerar o nível de significância  $\alpha = 5\%$ .

Percebe-se, claramente, que o número de informações ignoradas foi, na realidade, muito maior do que aquele inicialmente calculado. Para o primeiro grupo com 89 microrregiões, a média de ignorados das variáveis estudadas era de 2%. Após ser incorporado o sub-registro, este percentual subiu aproximadamente para 13%, um aumento muito mais próximo da magnitude

real de informações ignoradas da DN para as seis variáveis estudadas aqui. Do mesmo modo, os seguintes grupos, passaram, respectivamente de 4% e 1% para cerca de 11% e 16%, magnitudes muito superiores que as calculadas anteriormente.

## CONCLUSÃO

Levando-se em consideração a distribuição das microrregiões do Semiárido brasileiro segundo as diferentes variáveis empregadas com relação às condições das mães e das crianças geradas, observou-se, em 2010, uma dicotomia sintetizada pelas "condições favoráveis" e "condições desfavoráveis" ao parto.

É preciso considerar, no entanto, que como qualquer estimativa, deve-se atentar para os erros inerentes as estimativas pontuais.

O comportamento geral das coberturas para as microrregiões do Semiárido revela que o Sergipe e o Piauí foram às microrregiões dominadas por baixas coberturas, enquanto que em Minas Gerais predominaram coberturas acima dos 85%. Nos estados com coberturas intermediárias como Pernambuco e Rio Grande do Norte, a composição das regiões foi predominante com valores acima dos 80%. Isolando-se as microrregiões do extremo sul da Bahia, de forma genérica, a distribuição regional da cobertura dos nascimentos sugere um aumento à medida que os estados se posicionam em direção ao sul do Nordeste.

Admite-se que as coberturas obtidas neste trabalho reflitam valores aproximados e plausíveis dos valores verdadeiros.

A qualidade do preenchimento das variáveis cumpriu plenamente o requisito estabelecido como limite de aceitação mínimo de completude, 70%, à exceção da variável Apgar de 1º minuto e Apgar de 5º minuto para o ano de 2000. No entanto, tomando como base o referencial adotado por instituições e pesquisadores importantes no cenário mundial (50% a 60% de completude), pode-se considerar que o sistema de registros dos nascimentos no Semiárido produziu uma qualidade virtualmente satisfatória para se traçar perfis representativos das variáveis investigadas.

A identificação de grupos similares de microrregiões é fundamental para o planejamento de políticas públicas. Para possibilitar essas ações, torna-se necessária a busca de um maior entendimento do contexto onde os erros acontecem e a identificação de áreas onde essas situações compartilham uma dinâmica particular. Só a partir desse conhecimento, será possível realizar o planejamento de ações específicas.

Os resultados encontrados neste trabalho através da busca de um Índice Sintético que resulta da interação entre contexto social e a qualidade dos dados de nascimentos, mostrou que não foi possível diferenciar estatisticamente as microrregiões do Semiárido por uma única variável, nem por um conjunto escasso delas. No caso do Índice de Qualidade dos Registros de Nascimentos, houve uma diferenciação com relação aos grupos criados.

Espera-se que este trabalho possa contribuir para a definição de estratégias que melhorem os registros de nascimentos, e que possa subsidiar o planejamento e delimitar políticas públicas nas áreas de saúde materna e infantil nos estados e suas respectivas microrregiões do Semiárido.

## REFERÊNCIAS

BERQUÓ, ELZA; CAVENAGHI, SUZANA; Fecundidade em declínio: breve nota sobre a redução do número médio de filhos por mulher no Brasil. *CEBRAP* No.74 São Paulo, mar, 2006.

INSTITUTO NACIONAL DO SEMIÁRIDO. *Desertificação e Mudanças Climáticas no Semiárido Brasileiro*. Campina Grande, 2011.

PAES, NEIR ANTUNES; SANTOS, CARLOS SÉRGIO ARAÚJO; As estatísticas de nascimento e os fatores maternos e da criança nas microrregiões do Nordeste brasileiro: uma investigação usando análise factorial. *Cad. Saúde Pública*, Rio de Janeiro, 26(2):311-322, fev, 2010.

PERPÉTUO, IGNES HELENA OLIVA; Contracepção e declínio da fecundidade na Região Nordeste, 1980-1996.

SILVA, RICARLLY; OLIVEIRA, CONCEIÇÃO; FERREIRA, DANIELA; BONFIM, CRISTINE VIEIRA; Avaliação da completude das variáveis do Sistema de Informações sobre Nascidos Vivos – SINASC – nos Estados da região Nordeste do Brasil, 2000 e 2009. *Epidemiologia e Serviços de Saúde – Revista do Sistema Único de Saúde*, Volume 22 – Número 2 – abril/junho de 2013.

SÃO PAULO, 2011. Prefeitura da São Paulo *Manual de Preenchimento da Declaração de Nascido Vivo*. São Paulo: Prefeitura de São Paulo, fevereiro, 2011.

CRESPO, C. D.; BASTOS, A. A.; CAVALCANTI, W. A. A Pesquisa do Registro Civil: condicionantes do sub-registro de nascimentos e perspectivas de melhorias da cobertura. IBGE, 2004.

HAIR, J. F. ET AL. ANÁLISE MULTIVARIADA DE DADOS. Bookman Editora, 5a Edição – São Paulo, 2005. Cap. 3 Análise Fatorial, p. 503-512.

MELLO JORGE MHP, GAWRYSZEWSKI VP, LATTORRE MR. I – Análise dos dados de mortalidade. Rev.Saúde Publica 1997;31:5-25.

MELLO-JORGE MHP, GOTLIEB SLD, SOBOLL MLMS, ALMEIDA MF, OLIVEIRA H. Avaliação do Sistema de Informações sobre Nascidos Vivos: primeira avaliação dos dados brasileiros. Rev Saúde Pública 1993; 27 Suppl:1-45.

MINGOTI, A. S. Análise dos dados através de métodos de estatística multivariada: uma abordagem aplicada. Editora UFMG – Belo Horizonte, 2005. Cap. 4 – Análise Fatorial, p. 99 – 138.

MINISTÉRIO DA SAÚDE – Sistema de Informação de Nascidos Vivos – SINASC (www.datasus.gov.br). Acessado em 26/Agosto/2010.

PAES, N.A. E SANTOS, C.S.A. A qualidade das estatísticas de nascimentos e as condições de favorecimento ao parto das microrregiões do Nordeste brasileiro em 2000, através do uso da Análise Fatorial. Cad. Saúde Pública, Rio de Janeiro, 26(2):311-322, 2010.

PAES, N.A. E SILVA, M.R.M. Cenários regionais prospectivos das estatísticas de nascimentos do Nordeste Brasileiro. Trabalho apresentado no XVII Encontro Nacional de Estudos Populacionais, ABEP, Caxambu, Minas Gerais, 2010.

PAES, NA. E SILVA, M.R.M. A qualidade das estatísticas de nascimento no Nordeste brasileiro. Trabalho apresentado no 19º SINAPE – São Pedro, 2010.

PEREIRA, J.C.R. Análise de Dados Qualitativos: Estratégias Metodológicas para as Ciências da Saúde, Humanas e Sociais. São Paulo: EDUSP, 2001.

ROMERO, D.E E CUNHA, C.B. Disponibilidade e qualidade da informação epidemiológica e demográfica no Sistema de Informações sobre Nascidos Vivos (SINASC) e no Sistema de Informação de Mortalidade (SIM) para menores de um ano. Trabalho apresentado no XV Encontro Nacional de Estudos Populacionais, ABEP, Caxambu, Minas Gerais, 2006.

SCOCHI, C. G. S. ET AL. Intervalo entre o nascimento e o Registro Civil: situação no município de Ribeirão Preto, São Paulo, Brasil. Rev. Bras. Saúde Matern. Infant., Recife 4 (2): 171-178 abr./jun. 2004.

SIEGEL, S. Estatística não-paramétrica para ciências do comportamento, McGraw-Hill, Brasil, 1975.

SILVA AAM, RIBEIRO VS, BORBA JÚNIOR AF, COIMBRA LC, SILVA RA. Avaliação da qualidade dos dados do Sistema de Informações sobre Nascidos Vivos em 1997-1998. Rev. Saúde Pública; 35:5169-181, 2001.

SOUZA, L. M. Avaliação do Sistema de Informação sobre Nascidos Vivos. Trabalho apresentado no XIV Encontro Nacional de Estudos Populacionais, ABEP, Caxambu, Minas Gerais, Setembro de 2004.

WINKLER, W.E. Methods for evaluating and creating data quality. Information Systems, 2004.

SOUZA LM. Avaliação do Sistema de Informação sobre Nascidos Vivos. In: Anais do XIV Encontro Nacional de Estudos Populacionais [CD-ROM]. Campinas: Associação Brasileira de Estudos Populacionais; 2004.

BARBOSA LM, MELO GHN. Avaliação da qualidade das informações sobre fecundidade provenientes do SINASC no Nordeste, 2000. Rev Bras Estud Popul 2005;

SCOCHI CGS, COSTA IAR, ROCHA SMM, LEITE AM, Nascimento LC. Intervalo entre o nascimento e o registro civil: situação no Município de Ribeirão Preto, São Paulo, Brasil. Rev Bras Saúde Matern Infant 2004;