

ESTIMATIVAS DA MORTALIDADE INFANTIL PARA O ESTADO DO RIO GRANDE DO NORTE - BRASIL (2000-2009): ADAPTAÇÃO DO ESTIMADOR BAYESIANO EMPÍRICO ÀS MICRORREGIÕES*

Antonino Melo dos Santos [†]

Josivan Ribeiro Justino [‡]

Renata Clarisse C. de Andrade [¶]

Maria Helena Constantino Spyrides[°]

RESUMO

A proposta do presente trabalho foi de elaborar correções atualizadas para os óbitos infantis para o Estado do Rio Grande do Norte, a partir das Microrregiões no período de 2000 a 2009. A metodologia utilizada foi o estimador Bayesiano empírico, onde foi possível proporcionar uma suavização nos dados de óbitos dessas Regiões do Estado. Utilizou-se também, o método de Regressão Linear Simples para encontrar um modelo para cada Mesorregião, de forma a fazer estimativas médias da taxa de mortalidade infantil, a partir dos óbitos corrigidos.

Para o manuseio dos dados utilizou-se planilhas em Excel. As informações foram extraídas do Sistema de Informações sobre Mortalidade e do Sistema de Informações sobre Nascidos, hospedado no DATASUS / MS.

A partir da análise das causas mal definidas foi possível verificar uma queda significativa desse grupo para o total de óbitos no estado do RN, o que indica uma forte melhoria da qualidade dessas informações. E com base nas correções desenvolvidas pelo estimador foi possível criar taxas de mortalidade corrigidas e notar que o grau de cobertura das Mesorregiões Central e Oeste apresentaram menores índices de cobertura, o que significa um maior sub-registro. Para cada ano (2000 a 2009) e Mesorregião o estimador possibilitou fazer correções, refletindo nos valores estimados. Contudo, o estimador mostrou-se satisfatório ao passo que permitiu fazer correções significantes nas mesorregiões estudadas.

Palavras-chaves: Correção de sub-registro; Estimador Bayesiano empírico; Microrregiões do RN; Óbitos Infantis.

* Trabalho apresentado no *V Congresso da Associação Latino Americana de População*, ALAP, realizado em Montevidéu, Uruguai, de 23 a 26 de Outubro de 2012.

[†] Estatístico e Mestrando do Programa de Pós-Graduação em Demografia – UFRN
(*nino_melo@hotmail.com*)

[‡] Estatístico e Mestrando do Programa de Pós-Graduação em Demografia – UFRN
(*josivanrjustino@hotmail.com*)

[¶] Estatística e Mestranda do Programa de Pós-Graduação em Demografia – UFRN
(*renataclarice@yahoo.com.br*)

[°] Professora do Departamento de Estatística/UFRN. (*spyrides@ccet.ufrn.br*)

ESTIMATIVAS DA MORTALIDADE INFANTIL PARA O ESTADO DO RIO GRANDE DO NORTE - BRASIL (2000-2009): ADAPTAÇÃO DO ESTIMADOR BAYESIANO EMPÍRICO ÀS MICRORREGIÕES*

Antonino Melo dos Santos [†]

Josivan Ribeiro Justino [†]

Renata Clarisse C. de Andrade [‡]

Maria Helena Constantino Spyrides[°]

1. INTRODUÇÃO

Muito se tem trabalhado em busca da melhora da qualidade das informações sobre mortalidade no Brasil, em especial, direcionada às Regiões Nordeste e Norte, estas que ainda sofrem pelo mal das informações de baixa qualidade. É fato já constatado, inclusive pelo próprio Sistema de Informações sobre Mortalidade (SIM/MS), a ocorrência de menor sub-registro nas áreas mais desenvolvidas do território nacional que, no caso brasileiro e no nível municipal de desagregação, correspondem às capitais, outros municípios da região metropolitana e cidades médias.

Direcionando o foco pra a mortalidade infantil, visto que, ainda se constitui em um dos mais importantes indicadores das condições de saúde das comunidades, sendo indispensável para direcionar políticas de saúde voltadas para a população infantil, apesar do declínio observado nos últimos anos, a mortalidade infantil ainda continua muito elevada nas regiões Norte e Nordeste do Brasil. Segundo o IBGE, a taxa nacional de 85 óbitos por mil nascidos vivos, no início dos anos 80, passou para cerca de 37,5 óbitos por mil nascidos vivos, em 1996, o que representa redução de 44,1%, entre 1980/1990 e de 21,5% entre 1990/1996 (Maranhão, 1998). Este declínio tem sido atribuído a uma concorrência de fatores como intervenções ambientais, avanços da medicina, expansão do acesso e dos cuidados de saúde, diminuição da taxa de

*Trabalho apresentado no *V Congresso da Associação Latino Americana de População*, ALAP, realizado em Montevideu, Uruguai, de 23 a 26 de Outubro de 2012.

[†] Estatístico e Mestrando do Programa de Pós-Graduação em Demografia – UFRN (nino_melo@hotmail.com)

[‡] Estatístico e Mestrando do Programa de Pós-Graduação em Demografia – UFRN (nino_melo@hotmail.com)

[‡] Estatística e Mestranda do Programa de Pós-Graduação em Demografia – UFRN (nino_melo@hotmail.com)

[°] Professora do Departamento de Estatística/UFRN. (spyrides@ccet.ufrn.br)

fecundidade, aumento do nível educacional da população, melhoria nutricional e das condições de vida, de uma maneira geral.

Tal indicador constitui de uma baixa confiabilidade dos dados envolvidos em seu cálculo (nascimentos vivos e óbitos ocorridos entre os menores de um ano), pois padecem de elevado sub-registro nas áreas menos desenvolvidas do país, como citado anteriormente.

Para se ter um pouco mais de conhecimento relacionado à qualidade das informações, destaca-se a classificação proposta por Chackiel (1987) onde apresentou um arranjo sobre a qualidade dos registros de mortalidade segundo percentual de óbitos no grupo de causas mal definidas (CMD): menos que 15% - informação muito boa; entre 15 e 25% - informação relativamente boa; entre 25 e 40% - informação pouco confiável; mais que 40% - informação deficiente.

No entanto, a classificação da qualidade não se restringe apenas a análise das causas mal definidas (CMD) de óbitos no grupo, pois um índice não satisfatório, não significa necessariamente, a existência de um elevado grau de sub-registro (JUSTINO, 2010).

Segundo IBGE (2007), a RIPSAs - Rede Interagencial de Informações para a Saúde que é coordenada pela OPAS - Organização Pan-Americana da Saúde e pelo MS - Ministério da Saúde, estabeleceu em 10% de sub-registro como limite para a utilização dos dados em indicadores demográficos e de saúde.

Para tanto, a utilização de métodos que possam contornar tais dificuldades faz-se necessário para melhorar a visualização da realidade para as áreas em déficit de qualidade.

Portanto, de uma forma mais especificada, têm-se como objetivos a serem atingidos, a elaboração de correções atualizadas de sub-registro a partir de estimativas da mortalidade infantil para as Microrregiões do Estado do Rio Grande do Norte, utilizando a suavização indicada pelo estimador Bayesiano empírico de James-Stein simplificada por Marshall (1991) e proposto por Cavalline e Leon (2007), para o período de 2000-2009. A partir dessas correções, também, analisar a qualidade dos dados de óbitos de crianças menores de um ano de idade, por microrregiões do Estado, a partir do comportamento das causas mal definidas (CMD) de óbito no período de 2000-2009; Estimar o sub-registro de óbitos de crianças menores de um ano de idade, por microrregiões do Estado, no período de 2000-2009; Avaliar o comportamento das

correções das informações com relação à cobertura das informações segundo as microrregiões do Estado; Criar uma modelagem capaz de estimar a taxa média de óbitos infantis por Mesorregião fazendo uso da Análise de Regressão linear.

2. FONTE DE DADOS E METODOLOGIA

Nesta seção serão apresentados os elementos utilizados no desenvolvimento do trabalho: base de dados, descrição da área estudada e técnicas estatísticas e demográficas de análise.

2.1. Fonte de dados

Foi utilizado como fonte dos dados deste estudo, o Sistema de Informações sobre Mortalidade (SIM) e o Sistema de Informações sobre nascidos vivos (SINASC) gerenciada pelo Datasus/MS. Foram coletados os dados de óbitos de crianças menores de um ano de idade, por causas de morte e os nascidos vivos, segundo as 19 Microrregiões do Estado, referente ao período 2000 - 2009, como proposto inicialmente. Em seguida foi realizada uma análise da qualidade das informações a partir das informações de causa de mortes mal definidas (CMD) para o período. As estimativas bayesianas foram operacionalizadas através de uma expressão matemática, de manuseio relativamente simples, com uso de planilhas do Microsoft Excel 2007. Para desenvolver um modelo que fosse capaz de representar a taxa média de mortalidade infantil por mesorregião foi empregada a análise de regressão linear simples. O manuseio das informações foram desenvolvidas com o auxílio do software R, STATISTICA 7 e planilhas do Excel.

2.2. Definição da área de estudo

A área definida para o estudo foi o Estado do Rio Grande do Norte e suas divisões por Microrregião e Mesorregião, de acordo com as divisões desenvolvidas pelo IBGE para fins estatísticos. Ao todo, as 4 Mesorregiões do Estado se subdividem em 19 Microrregiões. As classificações estão dispostas da seguinte forma (Wikipédia, 2011):

- **Mesorregião Leste:** é formada pela união de vinte e cinco municípios agrupados em quatro Microrregiões (Litoral Nordeste, Litoral Sul, Macaíba e Natal). Pelo fato de nela estar localizada a capital do Estado, é a mais populosa, reunindo mais de 40% da população Potiguar. Municípios importantes dessa Mesorregião são Natal, Parnamirim, São Gonçalo do Amarante, Macaíba, Ceará-Mirim, Touros, São Miguel do Gostoso, Canguaretama e Tibau do Sul Miguel do Gostoso, Canguaretama e Tibau do Sul;
- **Mesorregião Agreste:** é a terceira mais populosa do estado, formada pela união de quarenta e três Municípios agrupados em três Microrregiões (Agreste Potiguar, Baixa Verde e Borborema Potiguar), além de ser a única do estado onde nenhum de seus municípios tem litoral. Municípios importantes dessa Mesorregião são Santa Cruz, João Câmara e São Paulo do Potengi;
- **Mesorregião Central:** é a menos populosa do Estado, formada pela união de cinco Microrregiões (Angicos, Serra de Santana, Seridó Ocidental, Seridó Oriental e Macau) que compartilham trinta e sete municípios. É a única Mesorregião limítrofe com todas as demais Mesorregiões do Rio Grande do Norte. Municípios importantes da Mesorregião são Caicó, Currais Novos, Macau, Guamaré, Angicos, Acari, Pedro Avelino e Galinhos, Macau, Guamaré, Angicos, Acari, Pedro Avelino e Galinhos;
- **Mesorregião Oeste:** é a segunda mais populosa do estado, dividida em sete Microrregiões (Pau dos Ferros, Serra de São Miguel, Umarizal, Vale do Açu, Mossoró, Médio Oeste e Chapada do Apodi) que, juntos, abrigam sessenta e dois Municípios, sendo a Mesorregião com o maior número de Municípios potiguares. Municípios importantes dessa mesorregião são Mossoró, Assu, Apodi, Pau dos Ferros, Areia Branca, São Miguel, São Rafael, Caraúbas, Patu, Tibau e Alexandria.

2.3. Estimador Bayesiano

Para a estimação dos óbitos gerais por microrregião, empregou-se a técnica de estimadores bayesianos de James-Stein, proposta por Cavalini e Leon (2007). Este

modelo é indicado quando existe o envolvimento de áreas de diferentes magnitudes, por exemplo, apresentando grandes flutuações de tamanhos populacionais. Neste trabalho, as microrregiões do Estado (totalizando dezenove), foram tomadas como unidade de análise, as quais, por sua vez foram aglutinadas formando as chamadas mesorregiões (totalizando quatro mesorregiões: Oeste, Central, Agreste e Leste).

Os contingentes populacionais envolvidos nos cálculos das taxas (nascidos vivos das microrregiões e óbitos de menores de um ano, por exemplo), estão sujeitos às chamadas oscilações dos pequenos números, pelo que se fez opção pela aplicação de uma metodologia capaz de contornar tais oscilações. Dessa forma, os dados foram submetidos ao processo de suavização propiciada pelo emprego do modelo bayesiano empírico de James-Stein (Efron & Morris, 1975), largamente recomendado por diversos autores (Assunção et al 1998; Freire e Assunção, 1998, 2002; Formiga, 2003; Cavalini, 2005 e Cavalini e Leon, 2007), como forma de reduzir as referidas flutuações. Como resultado da aplicação do método, obtém-se uma média ponderada (taxa de mortalidade ajustada – TMA) entre a taxa bruta da localidade (área menor, no caso, microrregião) e a taxa da área maior (região de vizinhança, no caso, mesorregião) tomada como referência.

As estimativas são operacionalizadas através da seguinte expressão:

$$\theta_i = T_m + C_i * (t_i - T_m),$$

Para,

$$C_i = [s^2 - (T_m / P_m)] / \{[s^2 - (T_m / P_m)] + (T_m / P_i)\},$$

Onde:

θ_i - *é a taxa suavizada;*

T_m - *é a taxa média global ou a taxa média dos vizinhos;*

t_i - *é a taxa da área i;*

s^2 - *é a variância da taxa a ser medida;*

P_m - *é a população média global ou a população média dos vizinhos;*

P_i - *é a população da área i.*

Pode-se observar, na expressão acima, que o multiplicador C_i será próximo de 1 se a população P_i for grande, o que fará a taxa suavizada (θ_i) tender para o mesmo valor da taxa (t_i) calculada inicialmente. Por outro lado, se a área em tela possuir uma

população muito pequena, então, C_i , será também muito pequeno (próximo de zero), fazendo com que a taxa suavizada da pequena área aproxime-se da taxa média de referência. Essa taxa de referência deve ser selecionada em função dos dados observados, com o critério de vizinhança podendo ser o de uma região maior ou da média de áreas vizinhas.

Do exposto, fica claro que a taxa de mortalidade ajustada poderá ser maior ou menor do que a taxa da grande área, tomada como região de referência, podendo ser menor do que a taxa de mortalidade observada para a microrregião, o que implicaria numa supressão de óbitos informados pelo SIM. Nesses casos, foi adotado o mesmo ajuste utilizado por Cavalini e Leon (2007) e Cavalini (2005), de desprezar essa taxa ajustada e adotar a taxa original da microrregião. Conforme observado pelos autores, como resultado final deste ajuste, o número total de óbitos do Estado tende a aumentar, pois somente serão permitidos acréscimos, e nunca reduções, ao número de óbitos em cada microrregião. De posse dos óbitos estimados ou finais, será possível estimar o sub-registro de óbitos para cada microrregião, tomado como a diferença entre a unidade e a chamada cobertura dos óbitos. Esta, por sua vez, é obtida da razão entre os óbitos observados e estimados (ou finais). Dada pela seguinte expressão:

$$\textit{Sub-registro} = 1 - (\textit{Óbitos observados} / \textit{Óbitos estimados})$$

Pode-se encontrar, também, o fator de correção aplicado a cada microrregião. Como mostra a expressão a seguir:

$$\textit{Fator de Correção} = 1 / \textit{Grau de cobertura}$$

Vale salientar que, vão existir microrregiões onde a cobertura terá representatividade de 100%, ou seja, o sub-registro será nulo. Isso ocorrerá quando o número de óbitos estimados for inferior ao número de óbitos observados, dado que, quando ocorre isto na estimativa, é adotado o valor de óbitos observado. Fazendo com que não seja possível fazer reduções nos números, e sim, acréscimos.

2.4. Análise de Regressão Linear Simples

Para estimar a taxa média de mortalidade infantil por mesorregião, utilizou-se da análise de regressão linear simples. A técnica é empregada para estimar uma relação que possa existir na população.

A regressão linear simples constitui uma tentativa de estabelecer uma equação matemática linear (linha reta) que descreva o relacionamento entre duas variáveis. Da mesma forma como é usada a média para resumir uma variável aleatória, a reta de regressão é usada para resumir a estimativa linear entre duas variáveis aleatórias.

Há diversas formas de utilização de equações de regressão, no caso estudado, a forma trabalhada foi de:

Estimar valores de uma variável, com base em valores conhecidos da outra. Em situações em que as duas variáveis medem aproximadamente a mesma coisa, mas uma delas é relativamente dispendiosa, ou difícil de lidar, enquanto que a outra não.

A Equação Linear

Principais características:

- Coeficiente angular da reta, dado pela tangente denominado de “b”.
- Coeficiente linear, denominado “a”, que é o valor de Y quando X=0.

Modelo: $\hat{Y} = a + bX$

Verifica-se que:

- Para um valor X_i podem existir um ou mais valores de Y_i amostrados.
- Para esse mesmo valor X_i se terá apenas um valor projetado \hat{Y}_i .
- Para cada valor de X_i existirá um desvio d_i (ou erro e_i) dos valores de \hat{Y}_i .
- Sempre teremos observações que não são pontos da reta.

Determinação da Equação Matemática

Na regressão, os valores y são preditos com base em valores dados ou conhecidos de x. A variável y é chamada variável dependente, e a variável x, variável independente.

Critério para obter os valores dos coeficientes a e b:

- Ajustar uma reta horizontal de valor igual à média dos valores de y, isto é, \bar{Y} , pois a média é uma reta de regressão com $b = 0$.

- Ajustar uma reta que divida os pontos observados de forma que a soma dos desvios seja nula. No entanto, a simples soma dos desvios leva à compensação dos desvios positivos e negativos, como já se viu no cálculo da variância.

O método dos mínimos quadrados

O critério é encontrar os coeficientes a e b da reta de regressão que minimizem a soma dos quadrados dos desvios.

Características importantes:

- A soma dos desvios verticais dos pontos em relação à reta é zero;
- A soma dos quadrados desses desvios é mínima (isto é, nenhuma outra reta daria menor soma de quadrados de tais desvios).

Simbolicamente, o valor que é minimizado é:

$$\sum d_i^2 = \sum (y_i - y_c)^2$$

Onde:

y_i = valor observado de y

y_c = o valor calculado de y utilizando-se a equação de mínimos quadrados com o valor de x correspondente a y_i .

Os coeficientes são calculados pelas fórmulas abaixo.

Tendo presente que $Cov(x,y) = r_{xy} \sigma_x \sigma_y$, o coeficiente b será igual a estas quatro fórmulas possíveis:

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} = \frac{Cov(x,y)}{Var(x)} = \frac{r_{xy}\sigma_x\sigma_y}{\sigma_x^2} = r_{xy} \frac{\sigma_y}{\sigma_x}$$

$$a = \frac{\sum y - b\sum x}{n} = \bar{Y} - b\bar{X}$$

Minimizar a soma dos quadrados dos desvios não garante que se tenha obtido a melhor reta ajustada, é apenas uma propriedade desejada de ajuste de reta.

O método de ajuste dos mínimos quadrados foi preferível por que:

- Obtém as melhores estimações, isto é, as estimativas não terão tendenciosidade.
- Oneram os desvios maiores, fato desejável que evita grandes desvios.
- Permite realizar testes de significância na equação de regressão.
- A reta de regressão passa pelo ponto formado pelos valores das médias das duas séries de observações.

3. RESULTADOS E DISCUSSÕES

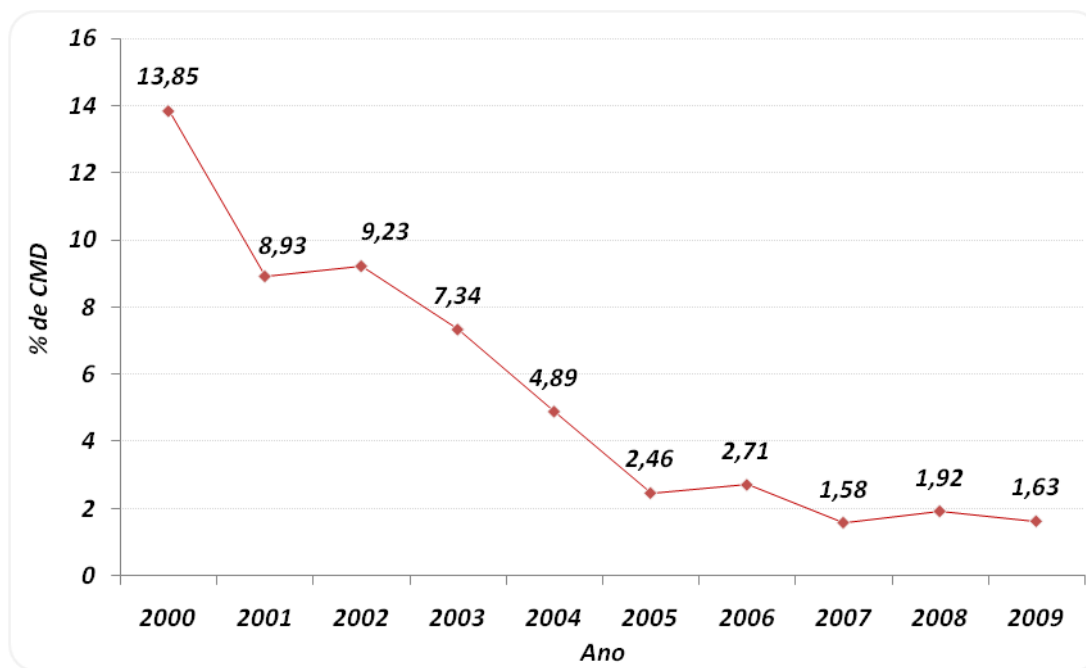
Nesta seção serão expostos os resultados encontrados, condizentes com os objetivos propostos, seguindo a seguinte sequência: Causas Mal Definidas, correções e observações do grau de cobertura, análise das correções, avaliando o método a partir da análise de variância, modelagem utilizando a técnica de regressão linear.

3.1. Análise das causas mal definidas

A omissão do registro do óbito em cartório, seja pela dificuldade de acesso, ou pela falta de orientação, existência de cemitérios irregulares ou pela falta de conhecimento da população sobre a importância da declaração de óbito, compromete o real dimensionamento do problema e a identificação das ações adequadas de saúde para a diminuição das taxas de mortalidade. A qualidade das informações nas declarações de óbito, representada pelo grande contingente de causas mal definidas de óbito – imprecisões na declaração da "causa da morte" – e campos não preenchidos, prejudica a análise dos fatores que influenciam a mortalidade e conseqüentemente, dificulta as ações de intervenção. Para conhecer a distribuição e qualidade dos óbitos infantis no Rio Grande do Norte, o Gráfico 01 apresenta a distribuição proporcional das Causas de Morte Mal Definidas (CDM).

Para o conhecimento inicial da distribuição e qualidade dos óbitos infantis, foi desenvolvido um gráfico (Gráfico 01) onde apresenta a distribuição proporcional das Causas de Morte Mal Definidas (CDM), dado que, é um representante de peso nas análises de qualidade das informações de óbitos em todos os níveis federais.

Gráfico 01: Evolução por ano da proporção de Causa Mal Definidas (CMD) ou XVIII - Sintomas, sinais e achados anormais de exames clínicos e de laboratório, não classificados em outra parte, para os óbitos infantis (menores de 1 ano) do Estado do Rio Grande do Norte.



Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

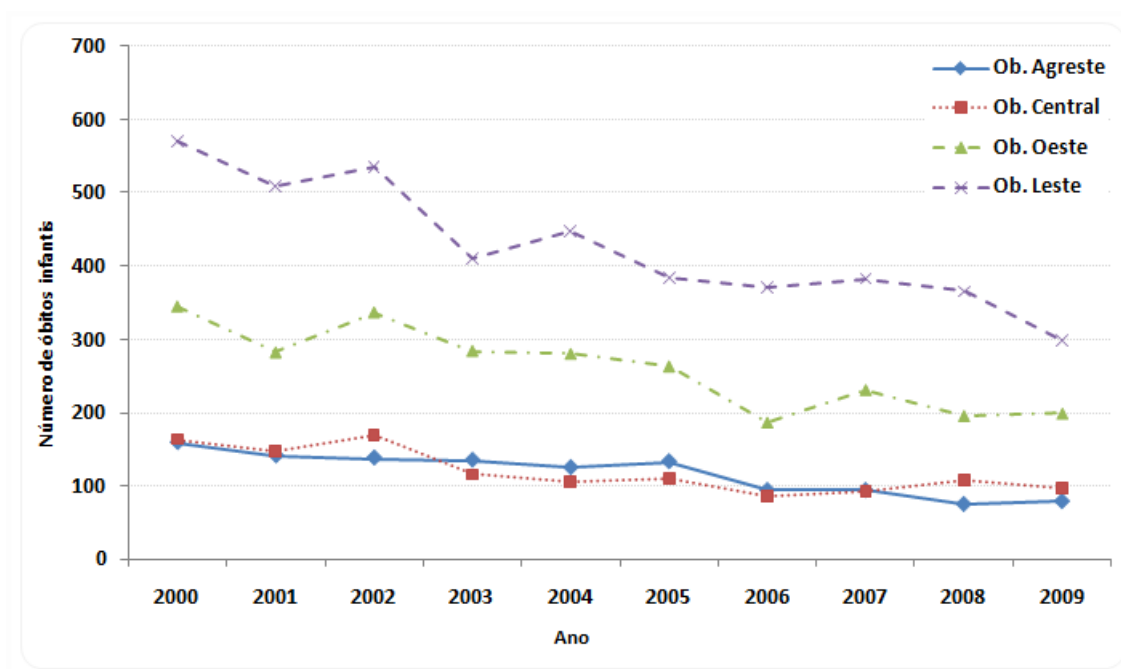
Fato notável foi a queda significativa da proporção nos registros de mortalidade no grupo de Causas Mal Definidas (CMD). Onde, no ano de 2000 representava 13,85% no Estado do Rio Grande do Norte, para os óbitos infantis (menores de 1 ano). E entre 2007 e 2009, chegou aos seus patamares mais baixos até agora registrados, 1,58% e 1,63%, respectivamente. Queda de cerca de 12,05%. Segundo a classificação apresentada por Chackiel (1987), coloca as proporções do Estado do Rio Grande do Norte, para os óbitos infantis, classificado como “informações de muito boa qualidade” (Menos de 15% de óbitos no grupo CMD).

2.5. Correções e apresentação do grau de cobertura segundo o estimador Bayesiano

O grau de cobertura de óbitos nas Mesorregiões foi calculado a partir das estimativa feitas para as Microrregiões. A unidade de análise final, onde efetivamente são estimados os óbitos esperados para calcular o grau de cobertura é a Mesorregião. Portanto, nas Mesorregiões calcula-se o grau de cobertura através do quociente entre a soma dos óbitos observados e a soma dos óbitos estimados das Microrregiões que pertencem a esta Mesorregião.

O Gráfico 2 apresenta o comportamento dos óbitos estimados para as Mesorregiões, com destaque para uma queda substancial dos óbitos nas Mesorregiões Leste e Oeste.

Gráfico 02: Evolução por ano dos óbitos estimados a partir do estimador bayesiano calculado com base nas Microrregiões para as Mesorregiões do Estado do Rio Grande do Norte.

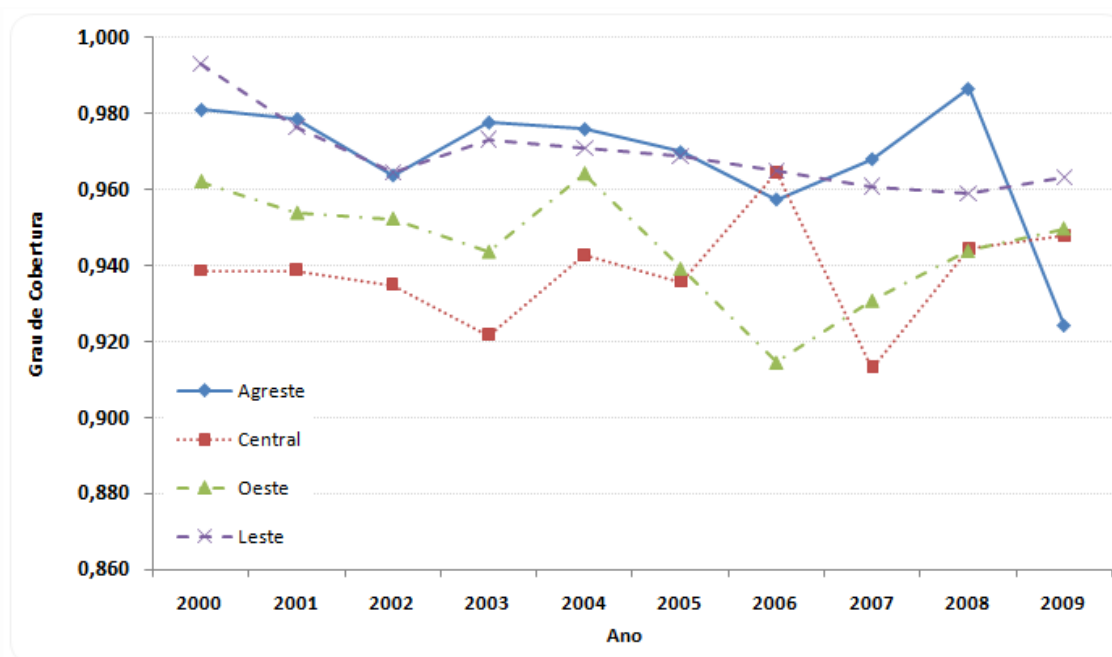


Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

O Gráfico 3 apresenta a evolução por ano, do comportamento do grau de cobertura nas Mesorregiões do Rio Grande do Norte. A Mesorregião Leste apresenta uma estabilidade na cobertura dos óbitos a partir do ano 2002, mantendo-se em um nível bom de cobertura. A Mesorregião Agreste mantinha um padrão semelhante ao da Mesorregião Leste, com exceção do ano 2009, onde foi registrado um aumento das mortes para os valores estimados, o que ocasionou uma redução drástica do grau de

cobertura. As Mesorregiões: Oeste e Central apresentam um grau de cobertura baixo, o que indica um número elevado de sub-registro, oscilando ano após ano, em valores muito próximos, o que demonstra uma similaridade das informações nestas Mesorregiões.

Gráfico 03: Evolução por ano do Grau de Cobertura de óbitos infantis, segundo estimador bayesiano calculado com base nas Microrregiões para as Mesorregiões do Estado do Rio Grande do Norte.

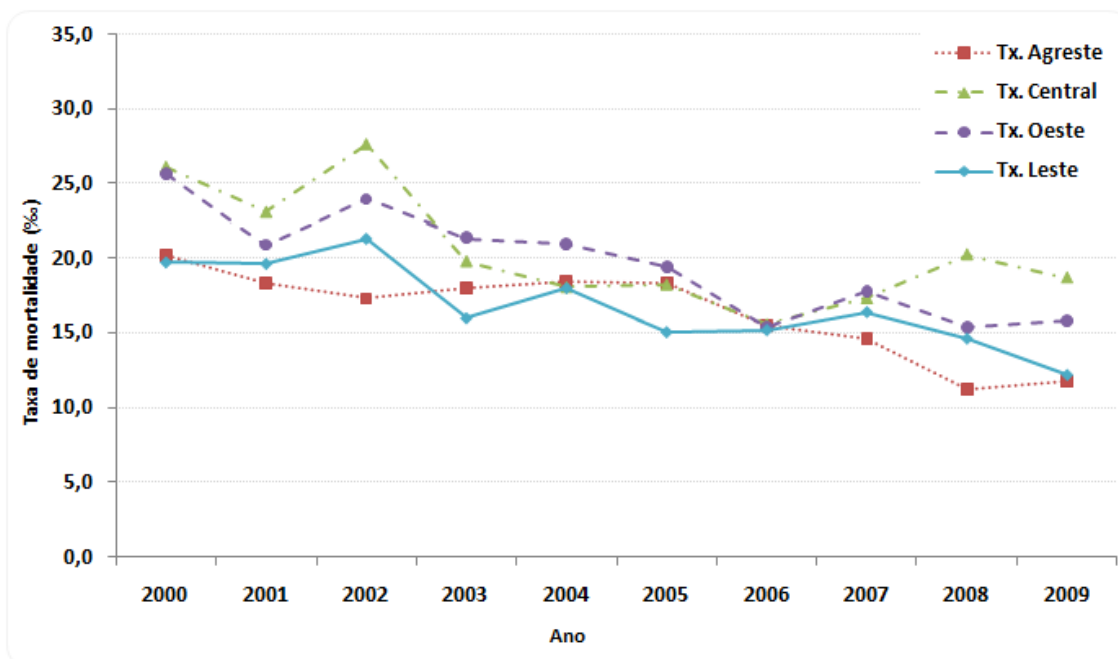


Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

2.6. Análise das correções através das taxas de mortalidade

As taxas de mortalidade infantil foram calculadas a partir dos óbitos estimados, sendo estas avaliadas por Mesorregião do Estado. O gráfico 4 abaixo, apresenta as taxas para cada ano em estudo, de 2000 a 2009 por Mesorregião, onde percebe-se que há uma tendência de queda em todas as Mesorregiões. Porém, as Mesorregiões Oeste e Central ainda apresentam taxas superiores às demais Mesorregiões. Fato interessante a se destacar, foi o ano de 2006, onde todas as Mesorregiões se aproximaram de uma taxa única.

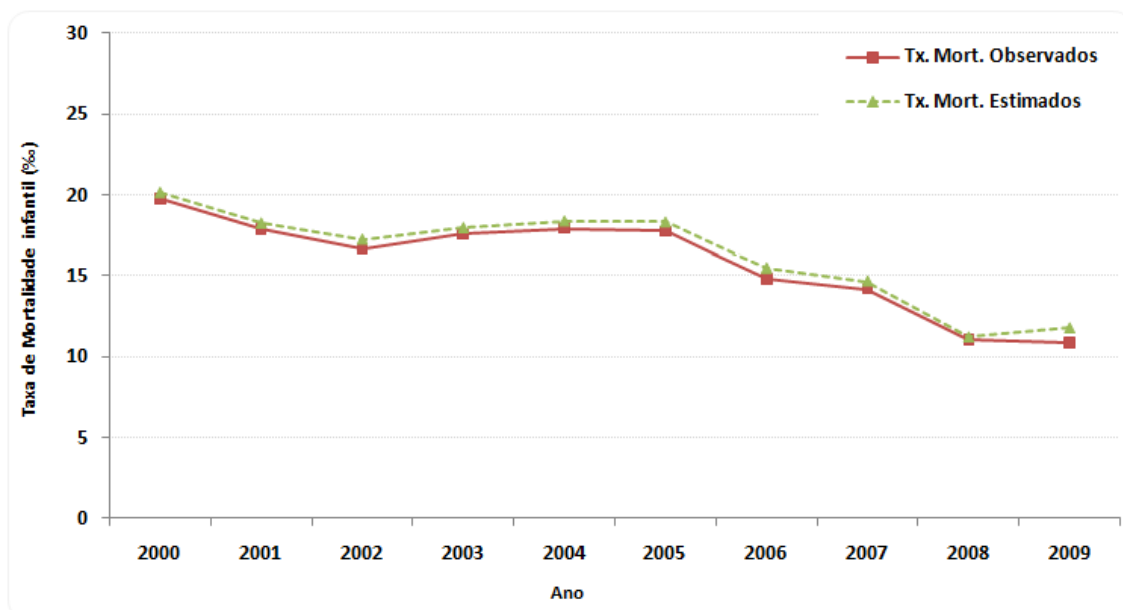
Gráfico 04: Evolução por ano das Taxas de Mortalidade infantil a partir dos óbitos estimados com o estimador bayesiano por Microrregiões para as Mesorregiões do Estado do Rio Grande do Norte.



Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

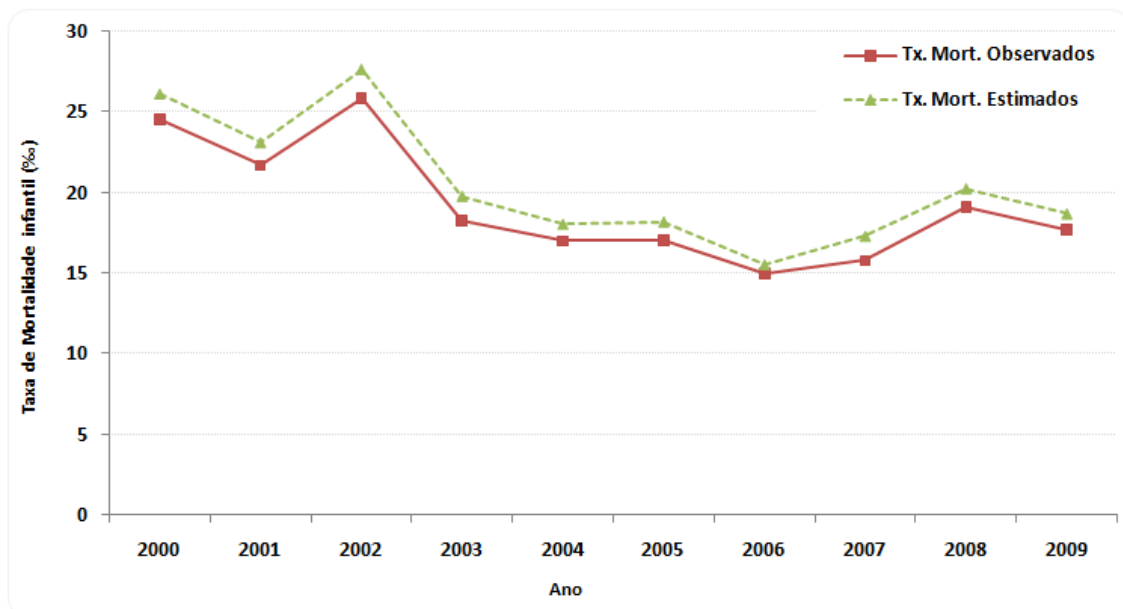
Os gráficos 5, 6, 7 e 8 apresentam a evolução por ano das taxas de mortalidade observadas e as estimadas para cada Mesorregião, no qual verifica-se que houve uma correção significativa em todas Meso. Com destaque para as Mesorregiões Oeste e Central, onde o estimador se mostrou mais ativo na correção.

Gráfico 05: Evolução comparativa por ano das Taxas Estimadas com as Taxas Observadas para a Mesorregião Agreste do Estado do Rio Grande do Norte.



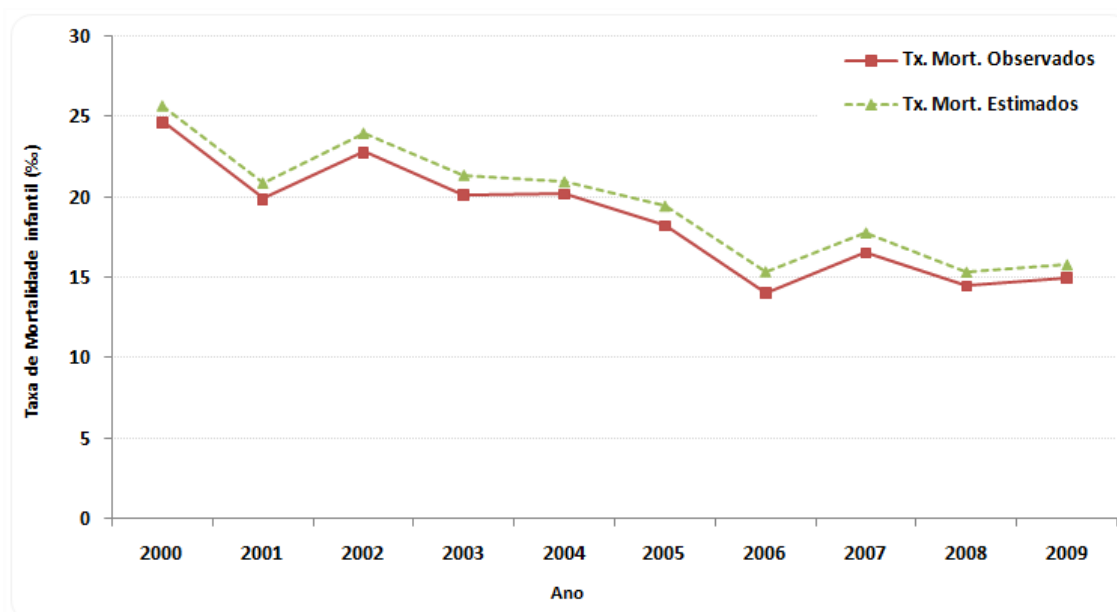
Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

Gráfico 06: Evolução comparativa por ano das Taxas Estimadas com as Taxas Observadas para a Mesorregião Central do Estado do Rio Grande do Norte.



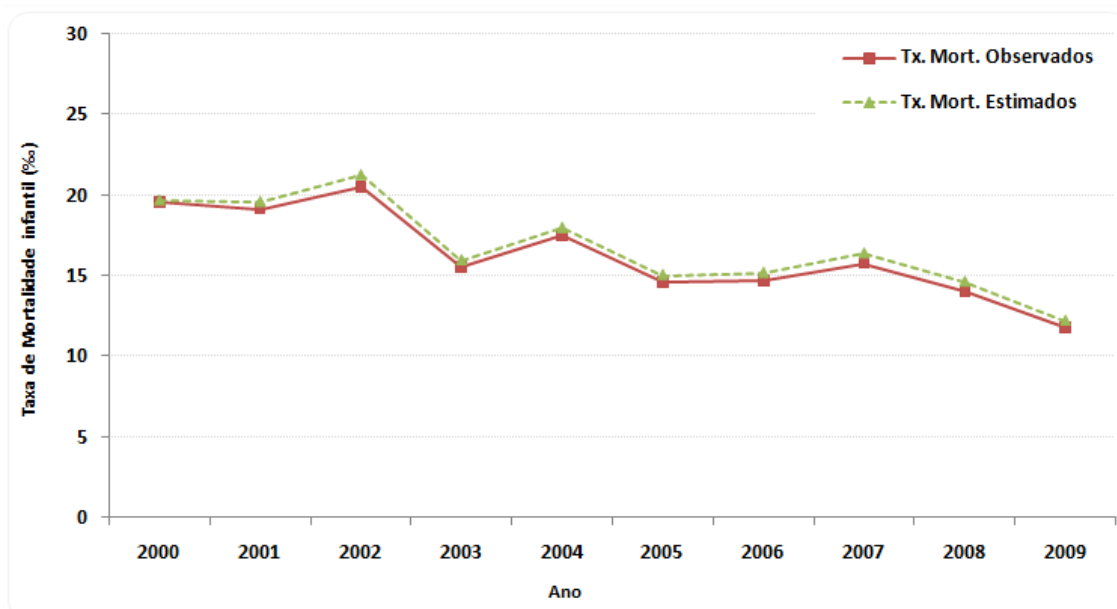
Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

Gráfico 07: Evolução comparativa por ano das Taxas Estimadas com as Taxas Observadas para a Mesorregião Oeste do Estado do Rio Grande do Norte.



Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

Gráfico 08: Evolução comparativa por ano das Taxas Estimadas com as Taxas Observadas para a Mesorregião Leste do Estado do Rio Grande do Norte.



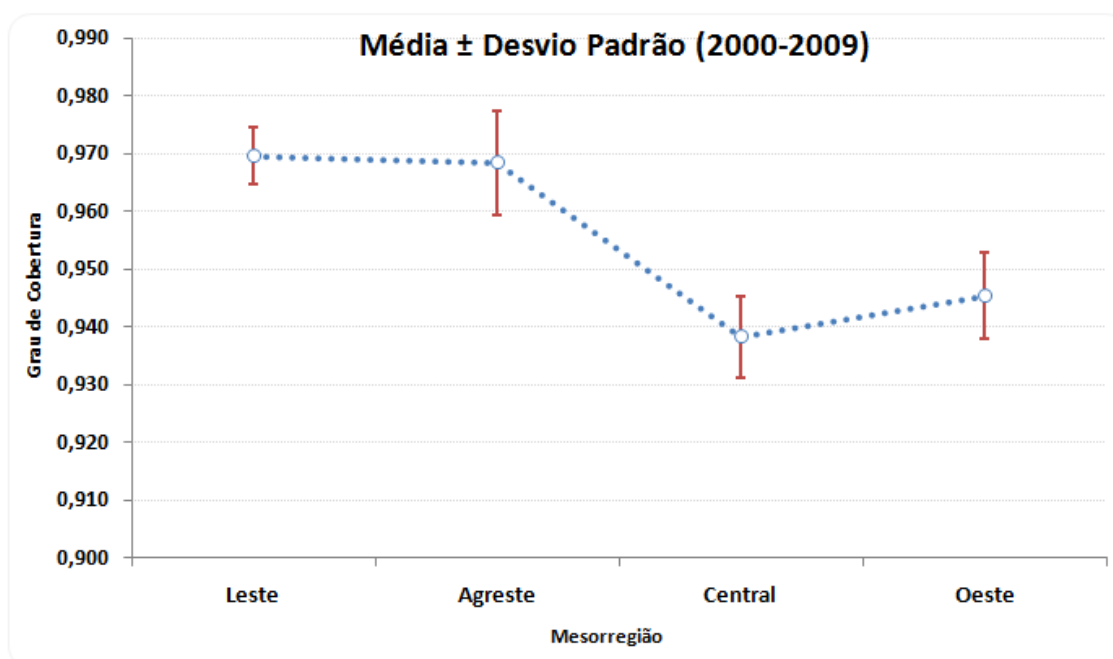
Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

2.7. Avaliação do método de suavização do estimador

Foi desenvolvido um método de avaliação das estimativas, pois se pretende verificar a diferença entre os graus de cobertura das Mesorregiões levando em consideração todo período de 2000 a 2009.

De forma inicial, pode-se verificar (Gráfico 09) as diferenças entre as Mesorregiões levando em consideração as médias e o desvio padrão. As Mesorregiões onde apresentaram maiores índices de Grau de cobertura foram as Mesorregiões Leste e Agreste. Regiões onde apresentam um índice significativo de desenvolvimento, tendo como principais encabeçadoras da lista, a capital do Rio Grande do Norte (Natal) e as cidades pertencentes à Grande Natal.

Gráfico 09: Estatísticas básicas, segundo o grau de cobertura de óbitos infantis para as Mesorregiões do Estado do Rio Grande do Norte com base na junção dos anos de 2000 a 2009.



Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

O gráfico das estatísticas só vem a comprovar realmente a diferença existente entre os graus de cobertura dessas Mesorregiões após a estimação ou suavização desenvolvidas nas Microrregiões para as Mesorregiões.

Contudo, foi elaborada uma análise estatística para dar a uma certificação da conclusão anterior. Com a análise de variância (ANOVA) foi possível verificar que ao nível de significância e 5%, as Mesorregiões apresentam níveis de grau de cobertura diferenciados ($p < 0,0001$) com relação as médias, no período (2000 a 2009).

Tabela 01: Análise de variância para verificar a diferença entre as estimativas das Mesorregiões, a partir do grau de cobertura, levando em consideração o período 2000 a 2009.

	Soma de Quadrados	Gl	Média de Quadrados	F	P valor
Intercepto	36,51100	1	36,51100	174776,7	0,000000
Mesorregião	0,00759	3	0,00253	12,1	0,000013
Erro	0,00752	36	0,00021		

Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

Mas apenas com a ANOVA, não é possível verificar quais Mesorregiões se assemelham mais. Para contornar tal problema, foi empregado o Teste de Tukey. Trata-se de um teste que subdivide os g grupos de observações em sub-grupos dentro dos quais pode-se considerar que as médias não apresentam diferenças significativas.

Tabela 02: Teste de Tukey para verificar a diferença significativa entre grupos da análise de variância.

Mesorregião	Grau de Cobertura	1	2
Central	0,93828	****	
Oeste	0,94544	****	
Agreste	0,96836		****
Leste	0,96949		****

Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

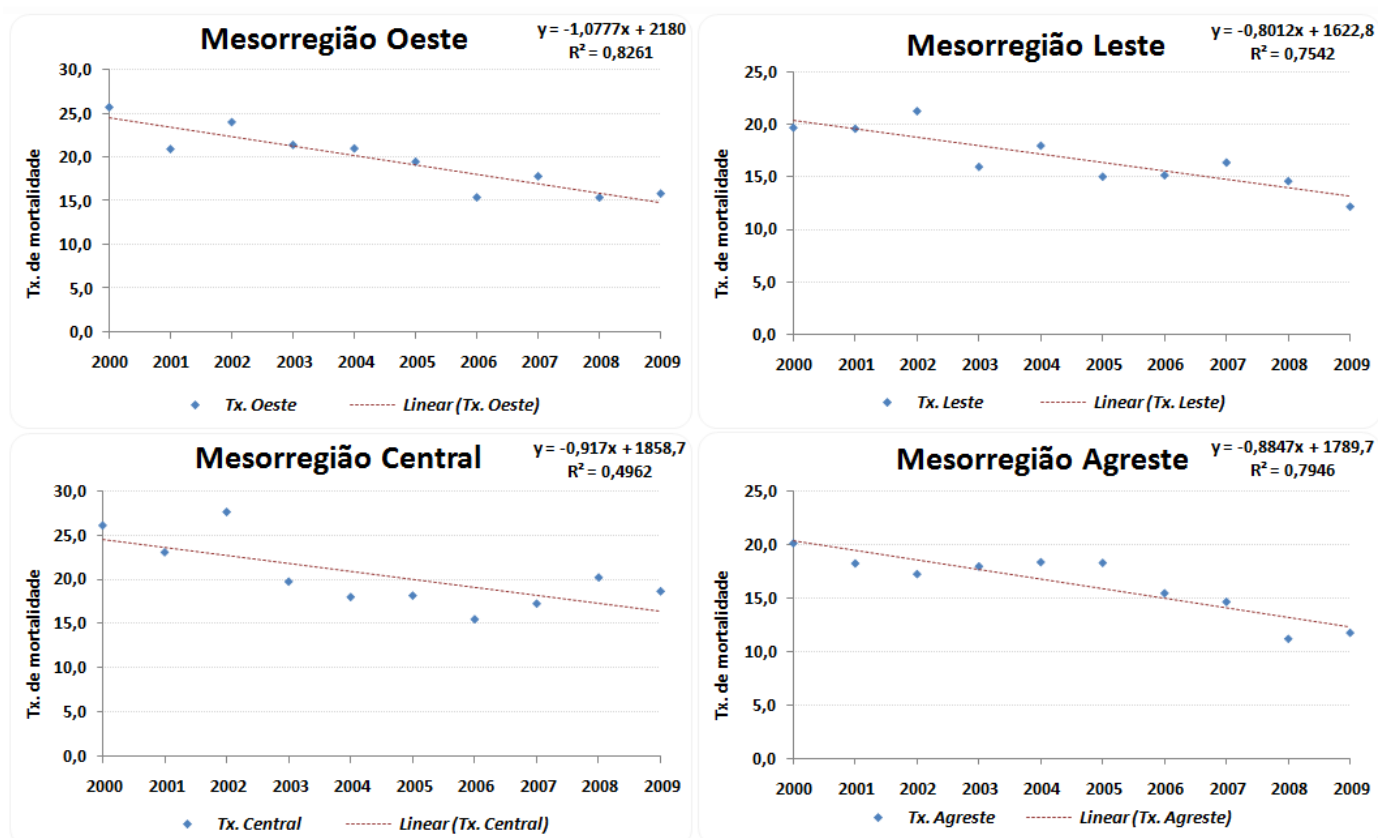
Entre os tratamentos 1 e 2, não existe diferença significativa entre as médias. Assim como, nos tratamento 3 e 4. Porém, 1 e 2 diferem do 3 e 4. A partir das médias acima verifica-se a diferença entre os tratamentos com base no Grau de Cobertura.

Portanto, as Mesorregiões mais semelhantes são Central e Oeste com os menores índices de Grau de cobertura e, Agreste e Leste com maiores graus de cobertura.

2.8. Análise de regressão e modelagem a partir dos óbitos estimados

De posse dos óbitos infantis estimados, após a aplicação Bayesiana, foi desenvolvida uma modelagem utilizando a Análise de Regressão Linear Simples, para verificar a tendência de queda das taxas de mortalidade infantil por Mesorregião e estimar a taxa média de mortalidade para anos posteriores a 2009.

Gráfico 10: Análise de Regressão linear simples por Mesorregião para estimativa média da taxa de mortalidade infantil já com óbitos estimados.



Fonte: Elaboração própria com base nos dados do SIM e SINASC/MS.

Com base nas estimativas acima, pode-se destacar os seguintes modelos para cada Mesorregião:

Mesorregião Agreste: $y = -0,8847 * x + 1789$ ($R^2 = 0,7946$)

Mesorregião Oeste: $y = -1,0777 * x + 2180$ ($R^2 = 0,8261$)

Mesorregião Central: $y = -0,917 * x + 1858$ ($R^2 = 0,4962$)

Mesorregião Leste: $y = -0,8012 * x + 16228$ ($R^2 = 0,7542$)

Contudo, avaliando o índice R^2 , onde mede a proporção da variação total da resposta explicada pela regressão, percebe-se que foram representativamente altos (acima de 0,7), com exceção da Mesorregião Central (0,4962). Portanto, pode-se afirmar que os modelos desenvolvidos foram bem ajustados.

4. CONSIDERAÇÕES FINAIS

A partir das informações referentes às proporções de causas mal definidas (CDM), percebeu-se que houve uma melhoria significativa na qualidade, pois houve um declínio satisfatório com relação aos óbitos infantis mal declarados no Estado.

A correção, ou suavização, proporcionada pelo estimador bayesiano, apresentou resultados satisfatórios, ao passo que foi capaz revelar observações, as quais, consideradas sub-registro, mediante o conhecimento prévio desta deficiência não só nas Microrregiões do Estado, mas também em outras UF's do Brasil. Fato interessante a se destacar é comportamento das Mesorregiões Oeste e Central, onde apresentaram altos índices de sub-registro, se comparado com as demais Mesorregiões (Leste e Agreste).

Com relação às taxas de mortalidade, todas as mesorregiões apresentaram quedas dentro do período estudado, apesar da correção realizada, o que realmente condiz com a realidade brasileira de queda na mortalidade infantil, graças ao esforço incalculável do Ministério da Saúde e demais órgãos competentes.

Utilizando a Análise de Regressão Linear foi possível elaborar uma modelagem capaz de estimar a taxa média de mortalidade infantil por Mesorregião do Estado (fazendo uso dos dados estimados), dado visto, que o Sistema de Informações sobre Mortalidade (SIM/MS) ainda não publicou definitivamente as informações sobre óbitos infantis para

o ano de 2010. Considerando que, no trabalho, por se tratar de óbitos infantis (menores de 1 anos), seria de grande valia, trabalhar a correção dos óbitos infantis por divisões etárias para estes menores (Crianças com idades contidas no período Perinatal, Neonatal e Pós-Neonatal), uma vez que estes fatores estão ligados a fatores genéticos, socioambientais, socioeconômicos, e condições de vida da criança e da mãe. Assim como Justino (2010), fez uso das divisões etárias, também seria possível realizar essa correção para as Microrregiões utilizando a metodologia adotada neste trabalho, deixando assim, uma análise elaborada mais a fundo com relação ao sub-registro de óbitos infantis.

5. RE FERÊNCIAS

Assunção, R. M.; Barreto, S.A.; Guerra, H. L.; Sakurai, E. 1998. Figuras de taxas epidemiológicas: uma abordagem Bayesiana. *Cadernos de Saúde Pública*, 14: 713-723,

Cavalini, L.T. 2005. Morbi-mortalidade nos municípios brasileiros: associação com indicadores sócio-econômicos e de assistência à saúde. Um estudo multinível (Tese Doutorado). Instituto de Medicina Social – IMS/UERJ, Rio de Janeiro.

Cavalini, L.T.; Leon, A. C. P. 2007. Correção de sub-registros de óbitos e proporção de internações por causas mal definidas. *Rev. De Saúde Pública. FSP*, (1): 85-93.

Chakiel, Juan. 1987. La investigacion sobre causas de morte en la América Latina. In *Notas de Poblacion*, Santiago, Chile, n.44,p.9-30, ago.

Efron, B.; Morris, C. 1975. Data analysis using Stein's estimation rule and its competitors: an empirical Bayes approach. *J Am Stat Assoc.*;70:311-9.

Freire, F. H. M. A; Assunção, R. M. 2002. Estimadores bayesianos para níveis de mortalidade e fecundidade em pequenas áreas. In: *SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA*, 15, 2002, Águas de Lindóia, SP, Resumos... São Paulo, SP, ABE, v. 1.

Formiga, M. C. C. ; Ramos, P. C. F. ; Barboza, I. M. ; Silva, S. C. A. S. 2003. Mortalidade Infantil no estado do Rio Grande do Norte: Uma análise da influência de fatores socioeconômicos. In: *XIV Congresso de Iniciação Científica da UFRN*, 2003, Natal - RN. Resumos. Natal - RN : UFRN.

IBGE. Estatísticas do registro civil 2007. Rio de Janeiro, 2007. Disponível em:
<<http://www.ibge.gov.br/home/estatistica/populacao/registrocivil/2007/default.shtm>>.
Acesso em: 30 jun. 2011.

Jusino, J. R. 2010. Estimação de sub-registros de óbitos em pequenas áreas com os métodos: Bayesiano Empírico e algoritmo EM. 2010. 52 f. Monografia (Bacharel) - Curso de Estatística, Departamento de Estatística, UFRN, Natal.

Maranhão, A. G. K.; Joaquim, M. M. C.; Siu, C. Mortalidade Neonatal e perinatal no Brasil. Brasília: Ministério da Saúde; 1998.

Marshall R. 1991. Mapping disease and mortality rates using empirical Bayes estimators. *Appl Stat.*40:283-94.

Trussel, T.J. 1975. A re-estimation of the multiplying factors for the Brass Technique for determining childhood survivorship rates. *Population Studies*, v. 29(1): 97-107, London.

Wikipédia - Enciclopédia Livre. Rio Grande do Norte. Disponível em:
<http://pt.wikipedia.org/wiki/Rio_Grande_do_Norte#Subdivis.C3.B5es>. Acesso em:
25 out. 2011.