

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS NATURAIS E EXATAS
CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA E MODELAGEM
QUANTITATIVA**

**ESTUDO DA MORTALIDADE INFANTIL – UM
MODELO DE REGRESSÃO LOGÍSTICA MÚLTIPLA**

MONOGRAFIA DE ESPECIALIZAÇÃO

Meire Mezzomo

Santa Maria, RS, Brasil

2009

ESTUDO DA MORTALIDADE INFANTIL – UM MODELO DE REGRESSÃO LOGÍSTICA MÚLTIPLA

por

Meire Mezzomo

Monografia apresentada ao Curso de Especialização em Estatística e Modelagem Quantitativa, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Especialista em Estatística e Modelagem Quantitativa.**

Orientadora: Prof. Dra. Anaelena Bragança de Moraes
Co-Orientador: Prof. Dr. Luis Felipe Dias Lopes

Santa Maria, RS, Brasil

2009

**Universidade Federal de Santa Maria
Centro de Ciências Naturais e Exatas
Curso de Especialização em Estatística e Modelagem Quantitativa**

A Comissão Examinadora, abaixo assinada,
aprova a Monografia de Especialização

**ESTUDO DA MORTALIDADE INFANTIL – UM MODELO DE
REGRESSÃO LOGÍSTICA MÚLTIPLA**

elaborada por
Meire Mezzomo

como requisito parcial para obtenção do grau de
Especialista em Estatística e Modelagem Quantitativa

COMISSÃO EXAMINADORA:

Anaelena Bragança de Moraes, Dra.
(Presidente/Orientadora)

Roselaine Ruviaro Zanini, Dra. (UFSM)

Luciane Flores Jacobi, Dra. (UFSM)

Santa Maria, 31 de agosto de 2009.

AGRADECIMENTOS

Agradeço,

À Universidade Federal de Santa Maria pela singular oportunidade em fazer parte do curso de pós-graduação em Estatística e Modelagem Quantitativa.

Em especial, agradeço à minha Orientadora Anaelena Bragança de Moraes pelas excelentes aulas, pela amizade, paciência, disponibilidade, dedicação, atenção, confiança, carinho, acreditando em minha capacidade; valorizando e incentivando meu trabalho: serei eternamente grata por tudo que fizeste por mim. Tua contribuição foi fundamental, professora Anaelena!

Ao meu Co-orientador Luis Felipe Dias Lopes pela dedicação, pelo conhecimento e contribuição para o desenvolvimento do trabalho.

Agradeço, com muito carinho, à professora Roselaine Ruviaro Zanini pela disponibilidade, presteza, dedicação, incentivo e auxílio.

Agradeço também à especial professora Luciane Flores Jacobi pelos seus ensinamentos compartilhados, paciência, dedicação; estando sempre disponível a atender seus alunos com compreensão, carinho e disponibilidade. Obrigada professora Luciane!

Aos professores do curso, que demonstraram competência em transmitir seus conhecimentos.

Agradeço aos meus sinceros amigos, em especial à minha grande amiga Marília Amaral pela ajuda, presteza, disponibilidade, carinho e apoio nos momentos difíceis.

Finalmente, aos meus alunos, pelo carinho, confiança e paciência.

RESUMO

Monografia de Especialização
Curso de Especialização em Estatística e Modelagem Quantitativa
Universidade Federal de Santa Maria

ESTUDO DA MORTALIDADE INFANTIL – UM MODELO DE REGRESSÃO LOGÍSTICA MÚLTIPLA

AUTORA: MEIRE MEZZOMO

ORIENTADORA: ANAELENA BRAGANÇA DE MORAES

CO-ORIENTADOR: LUIS FELIPE DIAS LOPES

Data e Local da Defesa: Santa Maria, 31 de agosto de 2009.

O objetivo deste estudo foi analisar os dados do Sistema de Informações sobre Nascidos Vivos (SINASC) e do Sistema de Informações sobre Mortalidade (SIM) para identificar os fatores de risco para a mortalidade infantil de nascidos vivos de mães residentes nos municípios pertencentes à 4ª Coordenadoria Regional de Saúde, RS, no ano de 2006. Para a análise dos respectivos dados, foi realizado um estudo longitudinal, utilizando um modelo de regressão logística múltiplo. O coeficiente de mortalidade infantil (CMI), que expressa o risco de um nascido vivo morrer antes de completar um ano de vida, é considerado um dos mais eficientes indicadores da qualidade de assistência à saúde, bem como do nível sócio-econômico de uma população. Neste estudo a variável dependente (desfecho) corresponde à ocorrência ou não de óbito em bebês com menos de um ano de vida. Os resultados do ajuste do modelo indicam que os fatores de risco preditores significativos ($p \leq 0,05$) para mortalidade infantil incluem as variáveis: estado civil (mãe sem união estável: OR = 2,0), idade gestacional (premature: OR = 3,3), peso ao nascer (baixo peso: OR = 4,7), índice de Apgar no 1º minuto (≤ 8 : OR = 3,9), índice de Apgar no 5º minuto (≤ 8 : OR = 5,6) e anomalias congênitas (com anomalia: OR = 20,4). Estes resultados corroboram com a maioria daqueles encontrados na literatura sobre mortalidade infantil no Brasil. Este estudo possibilitou o uso das bases de dados de nascimentos e óbitos, sugerindo uma melhor vigilância e planejamento das políticas públicas de saúde, com ênfase aos cuidados e atenção às gestantes e aos recém-nascidos, com o intuito de identificar precocemente as possíveis doenças maternas e fetais, proporcionando uma melhor qualidade de vida do bebê.

Palavras-chaves: Sistema de Informações sobre Nascidos Vivos; Mortalidade Infantil; Fator de Risco; Modelo de Regressão Logística Múltipla.

ABSTRACT

Specialization Monograph
Specialization Course in Statistics and Quantitative Modeling
Federal University of Santa Maria

STUDY OF INFANT MORTALITY – A MULTIPLE LOGISTIC REGRESSION MODEL

AUTHOR: MEIRE MEZZOMO

ADVISOR: ANAELENA BRAGANÇA DE MORAES

CO-ADVISOR: LUIZ FELIPE DIAS LOPES

Date and place of defense: Santa Maria, August 31th, 2009.

The present study had the purpose of analyzing the data of the System of Information on Born Alive (SINASC) and of the System of Information on Mortality (SIM) in order to identify the risk factors for the infant mortality of born alive from mothers living in the municipal districts belonging to the 4th Local Coordination Department in 2006. A longitudinal study has been developed for the analysis of the respective data by means of a multiple logistic regression model. The Infant Mortality Coefficient (CMI) which expresses the risk of a born alive dying before to complete one year old is considered as one of the most efficient indicatives of the quality of health assistance, as well as of the socio-economical level of a population. In the present study, the dependent variable (outcome) corresponds to the occurrence or not of death in babies with less than one year old. The results of the model adjustment show that the significant risk predicting factors ($p \leq 0,05$) for infant mortality include the variables: marital status (mother without stable engagement OR= 2,0) gestational age (premature: OR= 3,3) birth weight (low weight: OR= 4,7), Apgar index in the first minute (≤ 8 OR= 3,9), Apgar index in the fifth minute (≤ 8 : OR = 5,6), and congenital anomalies (with anomaly: OR= 20,4). These results are in agreement with most of the results found in the literature about infant mortality in Brazil. This study allowed using birth and death database, suggesting a better surveillance and planning of the public health political, mainly regarding the attention and cares with pregnant women, aiming at the early diagnosis of the possible fetal and maternal diseases in order to provide a better quality of life for the baby.

Keywords: System of Information on Born Alive; Infant Mortality; Risk Factor; Multiple Logistic Regression Model

LISTA DE TABELAS

TABELA 1 – Associação entre fator de risco e doença numa população	17
TABELA 2 – Distribuição do peso ao nascer segundo a idade gestacional dos nascidos vivos, SINASC/RS, 2006	39
TABELA 3 – Análise descritiva dos fatores de risco para a mortalidade infantil, segundo características maternas, RS, 2006 (n = 6963)	41
TABELA 4 – Análise descritiva dos fatores de risco para a mortalidade infantil, segundo características do nascido vivo, RS, 2006 (n = 6963)	42
TABELA 5 – Análise de regressão logística univariada, segundo características maternas, tendo como desfecho o óbito, RS, 2006 (n = 6963)	43
TABELA 6 – Análise de regressão logística univariada, segundo características do nascido vivo, tendo como desfecho o óbito, RS, 2006 (n = 6963)	44
TABELA 7 – Análise de regressão logística múltipla, tendo como desfecho o óbito, RS, 2006 (n = 6963)	45
TABELA 8 – Análise de regressão logística múltipla - MODELO FINAL, tendo como desfecho o óbito, RS, 2006 (n = 6963)	47

LISTA DE ABREVIATURAS E SIGLAS

ASSTEPLAN – Assessoria Técnica e Planejamento
CMI – Coeficiente de Mortalidade Infantil
CRS – Coordenadoria Regional de Saúde
DATASUS – Departamento de Informática do Sistema Único de Saúde
DN – Declaração de Nascido Vivo
DO – Declaração de Óbito
Estatística D – *Deviance*
Estatística G – *Deviance* Parcial
EP – Erro Padrão
G. 1. – Graus de liberdade
IC – Intervalo de Confiança
IBGE – Instituto Brasileiro de Geografia e Estatística
LILACS – Literatura Latino-americana em Ciências de Saúde
MEDLINE – *Medical Literature Analysis and Retrieval System Online*
MMQ – Método dos Mínimos Quadrados
MS – Ministério da Saúde
NV – Nascido Vivo
OMS – Organização Mundial da Saúde
OPAS – Organização Pan-Americana da Saúde
OR – *Odds ratio* ou razão de chances
PIB – Produto Interno Bruto
PAHO – *Pan American Health Organization*
PSF – Programa de Saúde da Família
RS – Rio Grande do Sul
SEPLAG/RS – Secretaria de Planejamento e Gestão do Rio Grande do Sul
SES – Secretaria Estadual da Saúde
SVS – Secretaria de Vigilância em Saúde
SIM – Sistema de Informações sobre Mortalidade
SINASC – Sistema de Informações sobre Nascidos Vivos
SIS – Sistema de Informações em Saúde
SUS – Sistema Único de Saúde

LISTA DE ANEXO

ANEXO A - Declaração de Nascido Vivo	57
---	-----------

SUMÁRIO

1 INTRODUÇÃO	10
1.1 Justificativa	11
1.2 Objetivos	12
1.2.1 Objetivo geral	12
1.2.2 Objetivos específicos	12
2 REVISÃO DA LITERATURA	13
2.1 Coordenadorias regionais de saúde	13
2.2 Sistemas de informações em saúde	14
2.2.1 Sistema de informações sobre mortalidade	14
2.2.2 Sistema de informações sobre nascidos vivos	14
2.3 Mortalidade infantil	15
2.3.1 Coeficiente de mortalidade infantil	16
2.4 Medidas de avaliação em estudos epidemiológicos	17
2.4.1 Risco	17
2.4.2 <i>Odds</i> ou chance	17
2.4.3 <i>Odds ratio</i> ou razão de chances	18
2.5 Regressão logística binária	20
2.6 Função de verossimilhança	24
2.7 Teste da razão de verossimilhança	25
2.8 Teste de Wald	27
2.9 Estudos sobre a mortalidade infantil	28
3 MATERIAL E MÉTODOS	34
3.1 Caracterização do estudo	34
3.2 Obtenção dos dados	34
3.3 Ajustes do banco de dados	35
3.4 Variável dependente	35
3.5 Variáveis independentes ou covariáveis	35
3.6 Modelagem	37
4 RESULTADOS E DISCUSSÕES	39
5 CONCLUSÃO	50
6 REFERÊNCIAS BIBLIOGRÁFICAS	52

1 INTRODUÇÃO

Conforme a Organização Pan-Americana da Saúde (OPAS/OMS, 2002), os indicadores de saúde são medidas-síntese que contêm informações relevantes sobre determinados atributos e dimensões do estado de saúde, bem como do desempenho do sistema de saúde, e são utilizados em saúde pública para avaliar as condições de vida de uma população.

O coeficiente de mortalidade infantil (CMI) pode ser considerado como um dos mais eficientes indicadores da qualidade da assistência à saúde, e também do nível socioeconômico de uma população. Este coeficiente identifica a relação entre o número de óbitos de crianças menores de um ano de vida e o número de nascidos vivos em um determinado local e período, sendo calculado na base de mil nascidos vivos (VICTORA; GRASSI; SCHMIDT, 1994).

Os óbitos relacionados às crianças menores de um ano de vida são influenciados diretamente por condições como: número de consultas pré-natais, tipo de gravidez, história materna (conduta e doença materna), idade materna, condições e tipo de parto, diferenças raciais maternas e infantis, condições socioeconômicas da mãe (escolaridade), intervalos entre partos, prematuridade, baixo peso ao nascer, malformações congênitas, estado marital, sexo dos bebês, mortalidade perinatal, mães portadoras de doenças infecto-contagiosas, dentre outros (MACHADO e HILL, 2005).

Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE, 2006), a taxa de mortalidade infantil no Brasil continua em declínio, passando de 36,9 óbitos de menores de um ano de vida por mil nascidos vivos em 1996 para 25,1 por mil em 2006, como mostra a Síntese de Indicadores Sociais de 2006. Segundo observação dos técnicos do Instituto, no documento de divulgação da pesquisa "a melhoria das condições de habitação, particularmente, o aumento relativo do número de domicílios com saneamento básico adequado, vem contribuindo para reduzir o número de mortes infantis". Em 2006, o estado que registrou a menor taxa de mortalidade infantil (14,9 por mil) foi o Rio Grande do Sul, enquanto a mais elevada foi apurada no estado do Maranhão (40,7 por mil).

Com o objetivo de identificar os fatores de risco associados à mortalidade infantil, pode ser utilizada a análise de regressão logística múltipla, que é uma técnica estatística que tem, também, como objetivo produzir um modelo que permita a predição de valores de uma variável categórica, frequentemente binária, neste caso a ocorrência ou não do óbito, a partir de uma série de variáveis explicativas contínuas e/ou categóricas (PAGANO, 2004).

Essa técnica, assim como a regressão linear múltipla, estuda a relação entre uma variável resposta e uma ou mais variáveis independentes. A diferença entre elas se deve ao fato de que, na regressão logística, a variável dependente é categórica, enquanto, na regressão linear, essa variável é contínua. Outra diferença expressiva é que, na regressão logística, a resposta é expressa por meio de uma probabilidade de ocorrência, enquanto na regressão linear, obtém-se um valor numérico, não necessariamente entre zero e um (HOSMER e LEMESHOW, 1989).

Neste estudo utilizou-se a regressão logística múltipla com o objetivo de identificar os fatores de risco da mortalidade infantil dos nascidos vivos de mães residentes nos municípios pertencentes à 4ª Coordenadoria Regional de Saúde, no ano de 2006. A escolha dessa região do RS deve-se ao fato de pertencer à região central do estado, cuja sede encontra-se em Santa Maria, região de referência em saúde no estado do Rio Grande do Sul.

1.2 Justificativa

O estudo de técnicas como a regressão logística múltipla é de grande interesse em diversas áreas do conhecimento, principalmente para a análise de dados oriundos da área de saúde.

Atualmente, essas técnicas têm sido empregadas em um número significativo de trabalhos científicos que levam em consideração dados reais, quando indivíduos estão expostos a fatores que levam a riscos de um evento.

A técnica de regressão logística permite realizar uma modelagem preditiva, utilizando-se um modelo de dados binários, que permite calcular a probabilidade do sucesso de um determinado evento, que, no presente estudo, representa a ocorrência de óbitos de nascidos vivos com menos de um ano de vida.

As informações sobre mortalidade infantil são provenientes de pesquisas estatísticas, demográficas e de estudos epidemiológicos. Essas pesquisas oferecem inúmeras vantagens para um melhor entendimento dos fatores determinantes da mortalidade infantil de uma determinada região, contribuindo, assim, para uma melhor fidedignidade dos dados obtidos e, conseqüentemente, possibilitando uma maior qualidade das informações sobre as causas relacionadas à mortalidade infantil e seus diversos fatores determinantes.

Este estudo pretende ampliar a compreensão dos elementos da estrutura relacionada à mortalidade infantil de bebês com menos de um ano de vida de mães residentes nos municípios pertencentes à 4ª Coordenadoria Regional de Saúde, a fim de contribuir com os setores específicos e responsáveis da saúde pública dessa região para um planejamento efetivo e preventivo, justificando-se, assim, a importância e relevância deste trabalho.

1.2 Objetivos

1.2.1 Objetivo geral

Identificar os fatores de risco para a mortalidade infantil de nascidos vivos de mães residentes nos municípios pertencentes à 4ª Coordenadoria Regional de Saúde, RS, no ano de 2006, utilizando um modelo de regressão logística múltipla e os Sistemas de Informações de Saúde - SINASC e SIM.

1.2.2 Objetivos específicos

- Realizar uma análise descritiva das variáveis do respectivo banco de dados;
- Encontrar um modelo de regressão logística múltipla binária para a mortalidade infantil, obtendo-se os fatores de risco a ela associados;
- Construir um texto em português sobre a análise de regressão logística múltipla, para servir como referência para futuros estudos dessa técnica de análise estatística.

2 REVISÃO DA LITERATURA

2.1 Coordenadorias Regionais de Saúde

Conforme a Secretaria Estadual de Saúde (SES/RS, 2006), o Estado do Rio Grande do Sul é constituído por dezenove Coordenadorias Regionais de Saúde (CRS) que possuem as seguintes atribuições:

- Coordenar o provimento de recursos financeiros para o desenvolvimento e manutenção das atividades administrativas das CRSs;
- Dar suporte técnico e administrativo à descentralização das ações e serviços das Coordenadorias Regionais de Saúde;
- Promover, em conjunto com Departamentos, Assessorias e Escola de Saúde Pública, o desenvolvimento de recursos humanos, visando a adequar as ações e serviços de saúde aos princípios do SUS;
- Acompanhar o desempenho dos serviços de assistência hospitalar e ambulatorial da rede pública e privada, em nível regional;
- Promover a articulação das Coordenadorias Regionais de Saúde no processo de Regionalização;
- Avaliar a descentralização dos serviços em conjunto com a Assessoria Técnica e Planejamento (ASSTEPLAN);
- Executar outras atividades correlatas ou que lhe venham a ser atribuídas;
- Atribuições do apoio administrativo;
- Planejar, coordenar e executar as atividades relacionadas à recepção, protocolo, digitação, arquivo e cópia de documentos, bem como demais rotinas administrativas (SES/RS, 2006).

A 4ª Coordenadoria Regional de Saúde do Rio Grande do Sul abrange 32 municípios do estado: Agudo, Cacequi, Capão do Cipó, Dilermando de Aguiar, Dona Francisca, Faxinal do Soturno, Formigueiro, Itaara, Ivorá, Jaguari, Jarí, Júlio de Castilhos, Mata, Nova Esperança do Sul, Nova Palma, Paraíso do Sul, Pinhal Grande, Quevedos, Restinga Sêca, Santa Maria, Santiago, São Francisco de Assis, São João do Polêsine, São Martinho da Serra, São Pedro do Sul, São Sepé, São Vicente do Sul, Silveira Martins, Toropi, Tupanciretã, Unistalda e Vila Nova do Sul, sendo Santa Maria o município-sede (SES/RS, 2006).

2.2 Sistemas de Informações em Saúde

A implantação do Sistema de Informações em Saúde (SIS), por meio de instrumentos padronizados de coleta de dados no país, é relativamente recente.

Segundo o Ministério da Saúde (MS/SV Sa, 2002), o SIS abrange um conjunto de informações de natureza administrativa, epidemiológica, demográfica e gerencial. A implantação desse sistema representa uma melhor estruturação dos serviços de vigilância em saúde, bem como uma maior organização dos serviços de atenção e da qualidade no atendimento aos seus usuários.

2.2.1 Sistema de Informações sobre Mortalidade

O Sistema de Informações sobre Mortalidade (SIM) foi criado, pelo Ministério da Saúde em 1975, para a obtenção regular de dados sobre mortalidade no Brasil.

A implantação desse sistema proporciona uma melhor captação de dados sobre mortalidade a fim de subsidiar as diversas esferas de gestão na saúde pública e, conseqüentemente, realizar análises de situação, planejamento e avaliação das ações e panoramas na área (MS/SV Sb, 2002).

O documento básico é a Declaração de Óbito (DO), padronizada nacionalmente e distribuída pelo Ministério da Saúde em três vias. As DOs são coletadas pelas secretarias estaduais ou municipais e são codificadas e transcritas para um sistema informatizado, sendo posteriormente disponibilizadas para o Departamento de Informática do SUS (DATASUS).

2.2.2 Sistema de Informações sobre Nascidos Vivos

O Sistema de Informações sobre Nascidos Vivos (SINASC) foi implantado pelo Ministério da Saúde em 1990 com o objetivo de reunir informações epidemiológicas referentes aos nascimentos informados em todo o território nacional. Por intermédio desses registros, é possível subsidiar as intervenções relacionadas à saúde da mulher e da criança, bem como as ações de atenção à gestante e ao recém-nascido (MS/SV Sc, 2002).

O documento básico é a Declaração de Nascido Vivo (DN), padronizada nacionalmente e distribuída pelo Ministério da Saúde em três vias, que são repassadas às

secretarias estaduais e municipais e posteriormente são repassados aos estabelecimentos de saúde e cartórios (MS/SVSc, 2002).

2.3 Mortalidade infantil

Conhecer o perfil da mortalidade infantil é fundamental para a formulação de estratégias que permitam o seu controle e prevenção. Altas taxas de mortalidade infantil refletem os baixos níveis de saúde e o desenvolvimento socioeconômico de uma população (KOZU et al., 2005).

Segundo a Secretaria do Planejamento e Gestão do RS (SEPLAG/RS, 2009), no Brasil, as ações desenvolvidas para a promoção da saúde da criança estabeleceram, de maneira geral, um expressivo declínio da mortalidade infantil nas diversas regiões do país. A taxa passou de 48,0 óbitos por mil nascidos vivos em 1990, para 34,6 em 1999 e 27,8 em 2002. Essa redução pode estar associada à ampliação da oferta dos serviços do Sistema Único de Saúde, à ampliação dos serviços de saneamento básico, ao aumento do aleitamento materno, ao declínio da fecundidade e ao aumento das coberturas vacinais, sendo que esses fatores resultam em melhores condições e qualidade de vida para uma população.

Embora a mortalidade infantil no Rio Grande do Sul tenha apresentado um sensível decréscimo, de 48,4 óbitos por mil em 1990, para 13,1 óbitos por mil em 2006, esse valor é superior ao considerado aceitável pela Organização Mundial de Saúde (OMS), que é de 10 mortes por mil nascimentos (SEPLAG/RS, 2009).

A distribuição espacial dos dados da mortalidade infantil expressa que, mesmo tendo aumentado o número de municípios com taxas mais baixas, ainda persistem um número significativo de municípios com taxas acima da média do estado do RS. No período de 2001 a 2003, 57,5% dos municípios gaúchos apresentaram taxas de mortalidade infantil inferiores à média do estado, que foi de 15,7 óbitos por mil nascidos vivos, e 33,1% dos municípios apresentaram mortalidade infantil inferior a 10 óbitos por mil nascidos vivos. No ano de 2006, o percentual de municípios com taxas de mortalidade infantil inferiores à média do estado aumentou: 64,12% dos municípios possuíam taxas menores que 13,1 por mil, enquanto que 55,24% apresentavam taxas inferiores a 10 óbitos por mil nascidos vivos (SEPLAG/RS, 2009).

Deve-se ressaltar também que a grande maioria dos municípios que apresentam taxas de mortalidade infantil mais elevadas são municípios com população e número de nascimentos baixos, onde os casos de morte repercutem com mais intensidade nas taxas de mortalidade infantil (SEPLAG/RS, 2009).

2.3.1 Coeficiente de Mortalidade Infantil

Para medir as condições de saúde, faz-se o uso de indicadores, que são medidas que possibilitam a verificação da ocorrência de um fato específico. Assim, um dos indicadores fundamentais para se avaliar a qualidade de vida de uma população é o Coeficiente de Mortalidade Infantil (CMI), que mensura quantas crianças menores de um ano, dentre as que nasceram vivas, morreram em um determinado período (VICTORA; GRASSI; SCHMIDT, 1994).

Conforme Simões (1999), o Coeficiente de Mortalidade Infantil é utilizado para analisar variações geográficas e temporais da mortalidade infantil na identificação de tendências e situações de desigualdade, permitindo-se, dessa forma, contribuir na avaliação dos níveis de saúde e desenvolvimento socioeconômico da população, bem como comparar a qualidade de vida entre regiões.

O CMI pode ser calculado pela utilização da equação (1) e expressa o risco de um nascido vivo morrer antes de completar um ano de vida, sendo obtido dividindo-se o número de óbitos em crianças menores de um ano pelo número total de crianças nascidas vivas (no mesmo período e local) e multiplicando por 1000 (SIMÕES, 1999).

$$\text{CMI} = \frac{\text{número total de óbitos}}{\text{número total de crianças nascidas vivas}} \times 1000 \quad (1)$$

O resultado obtido para o CMI, conforme equação (1), permitirá conhecer quantas crianças menores de um ano morreram em cada mil nascidas vivas.

No próximo item, serão apresentadas algumas medidas de avaliação epidemiológica de interesse para o desenvolvimento deste estudo.

2.4 Medidas de avaliação em estudos epidemiológicos

2.4.1 Risco

Segundo Jekel; Elmore; Katz (1999), o risco designa a proporção de incidência de pessoas que não estão afetadas no início do período de estudo, mas que passam pelo evento de risco (morte, doença, acidente, etc.) durante o período em estudo.

O risco de um evento de quem está exposto ao fator de risco pode ser descrito pela equação 2

$$\text{Fator de risco na exposição} = \frac{\text{número de doentes expostos}}{\text{número total de expostos}} = \frac{a}{a + b} \quad (2)$$

sendo os valores de a e b definidos de acordo com a Tabela 1.

Tabela 1 – Associação entre fator de risco e doença numa população

	Doente (D)	Não-doente (\bar{D})	Total
Exposto (E)	a	b	a + b
Não-exposto (\bar{E})	c	d	c + d
Total	a + c	b + d	n = a + b + c + d

Conforme apresentado na Tabela 1, a doença representa o desfecho, ou seja, doente corresponde à ocorrência, e não-doente, à ausência do desfecho. O exposto corresponde à presença do fator de risco, e não-exposto, à ausência do fator (JEKEL; ELMORE; KATZ, 1999).

2.4.2 Odds ou chance

Para Medronho (2003), a chance (*odds*) de um doente no grupo exposto pode ser descrito pela equação (3)

$$Odds = \frac{\text{número de doentes expostos}}{\text{número de não - doentes expostos}} = \frac{a}{b} \quad (3)$$

e a chance de um doente no grupo não-exposto é expresso por

$$Odds = \frac{\text{número de doentes não - expostos}}{\text{número de não - doentes não - expostos}} = \frac{c}{d} \quad (4)$$

sendo os valores de a , b , c e d definidos de acordo com a Tabela 1.

Quando o risco do evento é muito pequeno, esse valor aproxima-se da chance (*odds*) do evento.

2.4.3 *Odds ratio* ou razão de chances

A razão de chances (OR) pode ser calculada pela divisão da chance da doença na exposição pela chance da doença na não-exposição (ver equação 5).

Segundo Kahn e Sempos (1989), a OR é utilizada como uma medida de associação entre a exposição e a doença (evento), assumindo valores de zero a infinito (nunca menores que zero). Para o cálculo da OR, pode-se utilizar as equações (5) ou (6)

$$OR = \frac{\text{"odds" na exposição}}{\text{"odds" na não - exposição}} = \frac{a/b}{c/d} = \frac{a}{b} \times \frac{d}{c} \quad (5)$$

$$OR = \frac{\text{"odds" na doença}}{\text{"odds" na não - doença}} = \frac{a/c}{b/d} = \frac{a}{c} \times \frac{d}{b} \quad (6)$$

Observa-se que a razão de *odds* na exposição é sempre igual à razão de *odds* na doença, podendo-se também chamar a OR de razão de produtos cruzados, conforme demonstrado acima.

A interpretação da OR depende da sua magnitude e significância:

- Se $OR = 1$, o fator não é considerado de risco; ou seja, não há associação entre a exposição e a doença;

- Se $OR > 1$, o fator é considerado de risco, ou seja, indica que a *odds* da doença é maior quando exposto ao fator de risco (associação positiva);
- Se $OR < 1$, o fator é considerado de proteção, ou seja, indica a redução da *odds* para a doença na exposição ao fator de risco (associação negativa)

Para Kahn e Sempos (1989), o Método de Woolf é utilizado para estimar o valor da *odds ratio* populacional (Ω) pelo intervalo de confiança (IC), mostrado na equação (10)

$$P[\exp[\ln OR \pm Z_{\alpha/2} EP(\ln OR)]] = 1 - \alpha \quad (7)$$

com o erro padrão do $\ln OR$ igual a

$$EP(\ln OR) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}} \quad (8)$$

então

$$P[\exp(\ln OR) \times \exp(\pm Z_{\alpha/2} \times \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}})] = 1 - \alpha \quad (9)$$

e

$$P[OR \times \exp(-Z_{\alpha/2} \times \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}) < \Omega < OR \times \exp(+Z_{\alpha/2} \times \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}})] = 1 - \alpha \quad (10)$$

Para estimar a magnitude de associação (OR) entre o fator de risco e um evento de interesse, pode ser utilizada a análise de regressão logística, que consiste num poderoso instrumento que considera o ajuste simultâneo para os efeitos de variáveis de confundimento (MEDRONHO, 2003), que será apresentada no próximo capítulo.

2.5 Regressão logística binária

Segundo Berkson (apud WARREN, 1990), o modelo de regressão logística começou a ser difundido nos últimos anos pela alta capacidade de processamento dos computadores e desenvolvimento dos pacotes estatísticos.

De maneira geral, existem três procedimentos distintos para a manipulação dos dados, no que diz respeito à regressão logística: a regressão logística binária, ordinal e nominal. A escolha do método depende do número de categorias e das características da variável resposta. Uma variável binária é aquela que aceita apenas dois níveis de respostas, como sim ou não, alto ou baixo, enquanto que uma variável ordinal segue uma classificação ou ordenação natural, como pequeno, médio e grande. Já a nominal pode ter mais de três níveis e não considera nenhuma ordenação como a previsão do tempo: ensolarado, nublado ou chuvoso (PENHA, 2002).

A regressão logística é uma técnica semelhante à regressão linear, utilizada quando a variável dependente é categórica e, em geral, dicotômica. O objetivo desta técnica é identificar quais as variáveis independentes que influenciam no resultado (variável dependente) e usá-las numa equação para prever a probabilidade de as variáveis independentes explicarem o desfecho (HOSMER e LEMESHOW, 1989). Estas variáveis independentes podem ser categóricas e/ou contínuas, sendo que a variável dependente (y) é normalmente binária.

Para Pagano (2004), ao estudar uma regressão linear com p variáveis independentes, tenta-se estimar uma equação de regressão da população

$$\mu_y | x_1, x_2, \dots, x_p = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (11)$$

No caso de uma análise univariada, pode-se ajustar um modelo da forma

$$p = \beta_0 + \beta_1 x \quad (12)$$

no qual x representa uma variável independente. Esse é um modelo de regressão linear padrão, no qual y é o resultado de uma variável aleatória contínua, normalmente distribuída, que foi substituída por p, sendo que β_0 é o intercepto da linha, e β_1 é a sua inclinação. Sob

inspeção, no entanto, esse modelo não é apropriado. Uma vez que p é uma probabilidade, ela está restrita a assumir valores entre 0 e 1, entretanto, a equação (12) pode produzir um valor que se encontra fora desse intervalo (PAGANO, 2004).

Pode-se, então, tentar resolver esse problema ajustando-se o modelo na forma

$$p = e^{\beta_0 + \beta_1 x} \quad (13)$$

A equação (13) garante que a estimativa de p seja positiva. Embora esse modelo não produza uma estimativa negativa de p , em contrapartida, pode resultar em um valor maior do que 1.

Para satisfazer essa restrição final, ajusta-se o modelo da forma

$$p = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} = P(y/x) \quad (14)$$

sendo essa expressão denominada de função logística, que é a forma específica do modelo da regressão logística. Essa equação (14) ajustada não produz um valor negativo ou maior do que 1, ou seja, restringe o valor estimado de p para o intervalo exigido (0 a 1).

Sabe-se que, se um evento ocorre com probabilidade p , a chance a seu favor (sucesso)

é de $\frac{p}{(1-p)}$

$$\frac{p}{(1-p)} = \frac{e^{\beta_0 + \beta_1 x}}{(1 + e^{\beta_0 + \beta_1 x})} = e^{\beta_0 + \beta_1 x} \quad (15)$$

Aplicando-se o logaritmo neperiano em ambos os lados da equação (15) obtém-se

$$\ln \left[\frac{p}{1-p} \right] = \ln [e^{\beta_0 + \beta_1 x}] = \beta_0 + \beta_1 x = \text{logit} (p) \quad (16)$$

Essa transformação de p é denominada função *logit* e é linear nos seus parâmetros.

Segundo Cox (1969 apud HOSMER e LEMESHOW, 1989), muitas funções de distribuições têm sido propostas, mas a função ideal para a variável resposta dicotômica é a função *logit*, por ser extremamente flexível e fácil de ser usada e interpretada.

Com isso, modelar a probabilidade p como uma função logística é equivalente a ajustar um modelo de regressão linear, no qual a resposta contínua (y) foi substituída pelo logaritmo da chance de sucesso de uma variável aleatória dicotômica, sendo que o logaritmo é utilizado por questões matemáticas.

No entanto, em vez de assumir que a relação entre p e x seja linear, assume-se que a relação entre $\ln \left[\frac{p}{(1-p)} \right]$ e x é linear. Essa técnica de ajuste do modelo denomina-se regressão logística (PAGANO, 2004).

Conforme Kahn e Sempos (1989), a razão de chances ou *odds ratio* (OR) é a razão da chance do desfecho (y) na exposição ao fator de risco, em relação à chance do desfecho (y) na não-exposição ao fator de risco. Esta razão permite conhecer qual a chance de um evento acontecer, em relação a ele não acontecer sob as mesmas condições, ou seja,

$$odds = \frac{p}{1-p} \quad (17)$$

Na equação (16), se x for uma variável categórica binária, dicotomizada, onde x pode assumir o valor 0 ou 1, substituindo-se

$$\ln (odds)_{x=0} = \beta_0 + \beta_1 \times 0 = \beta_0 \quad (18)$$

$$\ln (odds)_{x=1} = \beta_0 + \beta_1 \times 1 = \beta_0 + \beta_1 \quad (19)$$

então, substituindo β_0 da equação (18) na equação (19), tem-se

$$\beta_1 = \ln (odds)_{x=1} - \beta_0 = \ln (odds)_{x=1} - \ln (odds)_{x=0} \quad (20)$$

usando-se a propriedade do logaritmos neperianos, obtém-se

$$\beta_1 = \ln \left(\frac{(odds)_{x=1}}{(odds)_{x=0}} \right) = \ln (OR) \quad (21)$$

Destaca-se que a OR corresponde ao aumento de uma unidade na variável independente (x), e tem-se

$$OR = e^{\beta_1} \quad (22)$$

onde $e = 2,718282$, que é o número neperiano, sendo a base do logaritmo mais utilizada no modelo de regressão logística múltipla.

O coeficiente β_0 é o $\ln (OR)$ quando se compara qualquer valor de x com o valor (x – 1), ou seja, com x subtraído em uma unidade.

Se x for uma variável contínua, o coeficiente de regressão β_1 , representa o aumento no $\ln odds$ de y para cada aumento unitário em x.

Generalizando para p variáveis independentes, tem-se

$$\ln odds = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (23)$$

substituindo o *odds* por $\left[\frac{p}{1-p} \right]$, resulta em:

$$\ln \left[\frac{p}{1-p} \right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (24)$$

e de acordo com a equação (14), tem-se

$$P (y/x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}} = \frac{1}{1 + e^{-(\ln odds)}} \quad (25)$$

Na regressão linear, o método mais usado para a estimação de parâmetros β é o método dos mínimos quadrados (MMQ), no qual são estimados os valores dos coeficientes que minimizam a soma dos quadrados dos desvios dos valores observados em relação aos valores preditos, baseados no modelo. Quando o MMQ é usado em modelos com resultado

dicotômico, os estimadores não apresentam propriedades estatísticas desejáveis para o ajuste do modelo. Nesse caso, utiliza-se o método de máxima verossimilhança, que produz estimativas para os parâmetros desconhecidos que maximizam a probabilidade de obtenção do conjunto de dados observados (HOSMER e LEMESHOW, 1989).

2.6 Função de verossimilhança

Para Hosmer e Lemeshow (1989), a função de verossimilhança expressa a probabilidade dos dados observados como uma função de parâmetros desconhecidos, sendo as estimativas resultantes aquelas que mais se aproximam do resultado. Essa função é obtida com o produto dos termos da expressão (26)

$$\zeta(x_i) = p(x_i)^{y_i} \times [1 - p(x_i)]^{1-y_i} \quad (26)$$

Assim, considerando os pares de valores (x_i, y_i) , onde $y_i = 1$, a contribuição para a função de verossimilhança é $p(x_i)$ e, para $y_i = 0$, a contribuição para a função de verossimilhança é $1 - p(x_i)$.

Quando as observações são consideradas independentes, a função de verossimilhança (equação 27) é

$$l(\beta) = \prod_{i=1}^n \zeta(x_i) \quad (27)$$

O princípio da máxima verossimilhança afirma que se deve utilizar como estimativa o valor de β que maximiza a expressão (27). Por questões matemáticas, é mais conveniente utilizar o logaritmo na expressão (27)

$$L(\beta) = [\ln l(\beta)] = \sum_{i=1}^n y_i \ln[p(x_i)] + (1 - y_i) \ln[1 - p(x_i)] \quad (28)$$

Para determinar os valores de β que maximizam $L(\beta)$ deve-se fazer a diferenciação em relação a β_0 e β_1 e igualar a zero. Essas equações são descritas na forma

$$\sum_{i=1}^n [y_i - p(x_i)] = 0 \quad (29)$$

e

$$\sum_{i=1}^n x_i [y_i - p(x_i)] = 0 \quad (30)$$

e são denominadas de equações de verossimilhança.

Na regressão linear, as equações de verossimilhança (29) e (30), obtidas por diferenciação, são lineares nos parâmetros, sendo facilmente resolvidas. Na regressão logística, essas equações de verossimilhança não são lineares nos parâmetros (β_0 e β_1), sendo necessária a utilização de métodos especiais para a obtenção da solução. Esses métodos são iterativos e encontrados em programas computacionais de análise de regressão logística. A solução é obtida usando o procedimento de mínimos quadrados ponderados generalizados.

2.7 Teste da razão de verossimilhança

Segundo Hosmer e Lemeshow (1989), esse teste tem por objetivo comparar os valores observados para a variável resposta (y) com os valores preditos obtidos no modelo com e sem a presença da variável em questão.

Dito de outra maneira, esse teste tem o objetivo de verificar se um subconjunto de variáveis x pode ser retirado do modelo de regressão logística múltiplo, ou seja, se os coeficientes de regressão β_i 's são iguais a zero. Neste caso, pode-se utilizar o teste da razão de verossimilhança, que é baseado na estatística de *deviance* (D) do modelo.

Esse teste é útil para verificar a significância do modelo geral:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{algum } \beta_i \neq 0.$$

A hipótese nula indica que nenhuma das variáveis independentes incluídas no modelo são significativas, do ponto de vista estatístico, para prever a probabilidade de $y = 1$.

Na regressão logística, essa comparação é baseada no logaritmo da função de verossimilhança, equação (28), e utiliza-se a seguinte expressão

$$D = -2 \ln \left(\frac{\text{verossimilhança do modelo atual}}{\text{verossimilhança do modelo saturado}} \right) \quad (31)$$

sendo que, o modelo saturado é aquele que se ajusta completamente aos dados, isto é, para cada observação tem-se um parâmetro.

A equação (31) é denominada de razão de verossimilhança (D), e a explicação sobre o uso de $-2 \ln$ é matemática e necessária para se obter a quantidade cuja distribuição é conhecida como qui-quadrado (χ^2). Sendo assim, D pode ser utilizada para teste de hipóteses. Esse teste é chamado de teste da razão de verossimilhança.

$$D = -2 \sum_{i=1}^n \left[y_i \ln \left(\frac{\hat{p}_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{p}_i}{1 - y_i} \right) \right] \quad (32)$$

onde, $\hat{p}_i = \hat{p}(x_i)$.

A estatística D é chamada de *deviance* por alguns autores. A *deviance* para a regressão logística tem o mesmo significado da soma dos quadrados residual na regressão linear.

Para Hosmer e Lemeshow (1989), a *deviance* (desvio) de um modelo compara o logaritmo da verossimilhança desse modelo com o logaritmo da verossimilhança do modelo saturado.

Se a *deviance* é pequena, a explicação do modelo ajustado é praticamente igual ao do modelo saturado, ou seja, pode-se utilizar o modelo ajustado, pois geralmente possui menos parâmetros, sendo mais simples (princípio da parcimônia).

Se a *deviance* é grande, a explicação do modelo ajustado é considerada pobre, ou seja, não se pode usar o modelo ajustado.

Com o objetivo de verificar a significância de uma variável independente incluída no modelo, compara-se o valor de D com e sem essa variável na equação (32). A alteração em D, devido à inclusão da variável independente no modelo, é obtida como

$$G = D \{ \text{para o modelo sem a variável} \} - D \{ \text{para o modelo com a variável} \} \quad (33)$$

A diferença entre as *deviances* de dois modelos é denominada de *deviance* parcial (G), e, por esta medida, é possível testar se determinada(s) variável(eis) explanatória(s) pode(m) ser retirada(s) do modelo.

A *deviance* parcial G pode ser expressa conforme a equação (34)

$$G = - 2 \ln \left(\frac{\text{verossimilhança sem a variável}}{\text{verossimilhança com a variável}} \right) \quad (34)$$

Se a *deviance* do modelo reduzido não é muito maior do que a *deviance* do modelo saturado, conclui-se que as variáveis extras podem ser retiradas do modelo de regressão logístico múltiplo. Por outro lado, se houver uma grande diferença entre as *deviances*, significa que as variáveis preditoras (x_i) devem ser mantidas no modelo, pois elas melhoram muito o ajuste do modelo, ou seja, melhoram sua explicação.

A *deviance* parcial, estatística G, segue aproximadamente uma distribuição do qui-quadrado com $p - q$ graus de liberdade, para um n razoavelmente grande. Os graus de liberdade correspondem à diferença nos graus de liberdade para os dois modelos ajustados: $(n - p) - (n - q) = p - q$. Se o valor da *deviance* parcial for maior que $\chi^2 (1 - \alpha; p - q)$, rejeita-se $H_0 (\beta_{r_s} = 0)$ portanto, as variáveis explanatórias devem permanecer no modelo (HOSMER e LEMESHOW, 1989).

2.8 Teste de Wald

É um teste estatisticamente equivalente ao teste de razão de verossimilhança (os pressupostos são os mesmos para ambos os testes).

Para Hosmer e Lemeshow (1989), o teste de Wald é utilizado para testar a significância dos coeficientes no modelo de regressão logística. Esse teste é obtido pela comparação da estimativa de máxima verossimilhança do parâmetro de inclinação β_i em relação à estimativa do seu erro padrão. A razão resultante, sob a hipótese de que $\beta_i = 0$, segue uma distribuição normal padrão. A estatística do teste é dada pela equação (35)

$$W_i = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (35)$$

onde

$\hat{\beta}_i$ = estimativa do coeficiente de regressão;

$SE(\hat{\beta}_i)$ = erro padrão do coeficiente de regressão β_i .

Sendo as hipóteses: $H_0: \beta_i = 0$ e $H_1: \beta_i \neq 0$, e comparando-se a estatística W_i com um valor tabelado $Z_{\alpha/2}$ (bilateral) para o nível de significância (α) considerado, pode-se concluir pela aceitação ou rejeição de H_0 .

Tanto o teste da razão de verossimilhança como o teste de Wald necessitam da estimativa de máxima verossimilhança dos parâmetros β_i 's.

Para ambos os testes, considerando um conjunto de dados com muitas variáveis independentes (x), a computação iterativa necessária para obter as estimativas de máxima verossimilhança são consideráveis (HOSMER e LEMESHOW, 1989).

Em suma, o método para testar a significância dos coeficientes para uma variável numa regressão logística é similar ao usado na regressão linear, mas utiliza uma função de verossimilhança para uma variável de desfecho dicotômica (HOSMER e LEMESHOW, 1989).

Uma observação importante a ser feita é que o teste de Wald frequentemente rejeita coeficientes que seriam significativos (HAUK e DONNER, 1977).

2.9 Estudos sobre a mortalidade infantil

Segundo informações do Ministério da Saúde (MS/SVS, 2009), no Brasil a mortalidade infantil, apesar de apresentar uma tendência decrescente, ainda constitui um importante problema para a Saúde Pública. Dentre os principais fatores de risco associados à mortalidade infantil, pode-se ressaltar: características do recém-nascido, condições socioeconômicas, características maternas e do sistema de saúde.

Sabe-se que inúmeros estudos sobre a mortalidade infantil no Brasil foram descritos por diferentes autores. Alguns estudos descritivos e analíticos serão apresentados neste item.

Victoria; Grassi; Schmidt (1994) analisaram a evolução temporal no período de 1980 a 1992 e a distribuição espacial em 1990 de alguns indicadores de saúde infantil no estado do Rio Grande do Sul. Os fatores analisados foram: mortalidade infantil, mortalidade proporcional em menores de um ano de vida, baixo peso ao nascer e cobertura vacinal. Os

autores observaram uma expressiva redução da mortalidade infantil nesse período (39,0% para 19,3%). A prevalência de baixo peso ao nascer apresentou-se estável, entre 8 e 10%, com um discreto aumento até 1991. A mortalidade e o baixo peso ao nascer mostraram-se desigualmente distribuídos dentro do estado, sendo que a região Norte apresentou as melhores condições de saúde.

O estudo conduzido por Victora et al. (1996) descreveu a tendência temporal da mortalidade infantil, utilizando dois estudos de coorte realizados em Pelotas, RS, nos anos de 1982 e 1993. Ambas as coortes, incluíram todos os nascimentos hospitalares e óbitos registrados em hospitais, cartórios e cemitérios. O levantamento dos dados ocorreu por entrevistas, revisão de prontuários e necropsias. Segundo os autores, as principais causas para a mortalidade infantil em 1993 foram: as perinatais, as malformações congênitas, as diarreias e as infecções respiratórias. Os nascidos vivos com baixo peso ao nascer apresentaram taxa de mortalidade doze vezes maior que os com peso ideal, e os bebês pré-termo, duas vezes maior que os bebês com retardo de crescimento intra-uterino. Em relação à renda familiar, os bebês de família de baixa renda (1 salário mínimo) registraram mortalidade sete vezes maior em relação aos de alta renda (10 salários mínimos). O coeficiente de mortalidade infantil apresentou um decréscimo de 36,4 por mil em 1982 para 21,1 por mil em 1993, uma redução de 42%, principalmente à custa do componente de óbitos pós-natal, cuja queda foi de 57%. Ocorreram 111 óbitos infantis em 1993 contra 215 em 1983, sendo que destes 111 óbitos, 18,9% ocorreram em domicílios, 3,6% em berçários, 2,7% em enfermarias pediátricas, e os demais, 74,8%, em UTIs. Quanto às causas de mortalidade infantil, as reduções de 1982 para 1993 foram: 28% para causas perinatais, 60% para diarreia, 69% para infecções respiratórias e 93% para outras infecções, sendo que as malformações congênitas e demais causas não sofreram alterações nesse período.

Araújo; Bozzetti; Tanaka (2000) analisaram os fatores associados à mortalidade neonatal precoce (7 dias ou menos) no município de Caxias do Sul, RS, no ano de 1995. A população em estudo foi constituída de todos os nascimentos hospitalares até o 7º dia de vida, sendo excluídos os partos domiciliares (0,6%). O levantamento dos óbitos foi obtido pela revisão dos prontuários, dos livros de registros hospitalares e na Delegacia Regional de Saúde. Para as análises estatísticas, foram utilizadas a regressão logística univariada e múltipla e o risco relativo (RR) como medida de risco. A probabilidade de morte neonatal precoce observada foi de 7,44 por mil nascidos vivos, sendo que 42 morreram na primeira semana de vida e destes, 20 óbitos (47,6%) foram no 1º dia de vida. Os fatores de risco associados à mortalidade neonatal incluíram as variáveis: história de natimortalidade (RR = 18,25); índice

de Apgar 1º minuto inferior a 7 (RR = 10,11); peso ao nascer inferior a 2500g (RR = 5,93); idade materna igual ou superior a 35 anos (RR = 4,52) e o sexo masculino (RR = 4,16). A principal causa dos óbitos neonatais registrados neste estudo foi a doença da membrana hialina, seguida das cardiopatias congênitas, partos prematuros e o deslocamento externo da placenta.

No estudo de Sarinho et al. (2001), foram analisados os fatores de risco para os óbitos neonatais na cidade do Recife, em Pernambuco, por meio de um estudo caso-controle cuja validação dos dados deu-se utilizando o método de *linkage* do banco do SINASC e SIM. No estudo foi utilizada a análise de regressão logística múltipla para avaliar os fatores de risco e o *odds ratio* como medida de risco. Foram considerados na análise todos os óbitos dos recém-nascidos com peso superior a 500g. Dos óbitos neonatais, 84,9% ocorreram no período precoce, concentrando-se 46,6% destes óbitos nas primeiras 24 horas, sendo 79,7% de baixo peso ao nascer (inferiores a 2500g). Os autores identificaram os seguintes determinantes associados à mortalidade neonatal: peso ao nascer menor que 1500g (OR = 49,6); índice de Apgar 5º minuto menor que 7 (OR = 44,1); peso ao nascer entre 1500g e 2500g (OR = 8,2) e idade gestacional inferior a 37 semanas (OR = 4,3).

Almeida et al. (2002) analisaram a influência do peso ao nascer, os fatores sociodemográficos e assistenciais relacionados a óbitos neonatais no município de São Paulo, SP, em 1998, por meio de um estudo caso-controle, vinculando-se as bases de dados do SINASC e SIM. As variáveis utilizadas no estudo foram divididas em quatro categorias: características maternas, características do recém-nascido, condições socioeconômicas e características dos serviços de saúde. Foi utilizada uma análise hierárquica, separada por faixas de peso, utilizando a regressão logística não-condicional e a *odds ratio* como medida de risco. Nesse estudo os autores analisaram os óbitos totais e as faixas de peso, excluindo os nascimentos sem informações sobre o tipo de parto, peso ao nascer não informado, com menos de 500g e óbitos relacionados a anomalias congênitas. Os fatores de risco encontrados para a mortalidade neonatal incluem: peso ao nascer inferior a 1500g (OR = 125,43), peso ao nascer entre 1500 e 2499g (OR = 6,96), idade gestacional inferior a 37 semanas (OR = 3,75) e idade materna (OR = 1,51). Para o grupo dos recém-nascidos com muito baixo peso ao nascer (inferior a 1500g), mostraram significância as variáveis relacionadas aos serviços de saúde: hospital do SUS (OR = 3,26), maternidade (OR = 2,73) e parto cesáreo (OR = 0,34). Para o grupo de baixo peso ao nascer (1500 a 2499g), os fatores associados foram: idade gestacional inferior a 37 semanas (OR = 3,87), ter 4 ou mais filhos (OR = 2,43) e sexo masculino (OR = 1,52). Para o grupo com peso ao nascer igual ou superior a 2500g incluem as variáveis: idade

gestacional inferior a 37 semanas (OR = 6,58), nascimento dentro do hospital do SUS (OR = 2,02), idade materna inferior a 20 anos (OR = 1,70), parto cesáreo (OR = 1,63) e precariedade residencial (OR = 1,60).

Os autores Martins e Velasquez-Mélendez (2004) analisaram os fatores determinantes relacionados à mortalidade neonatal a partir de uma coorte de nascidos vivos em Montes Claros, MG nos anos de 1997 a 1999, utilizando o método de *linkage* entre os bancos do SINASC e SIM. Para o estudo desses determinantes, utilizaram a técnica de regressão logística e a *odds ratio* como medida de risco. Foram identificados os seguintes determinantes associados à mortalidade neonatal: idade gestacional inferior a 37 semanas (OR = 5,68), peso ao nascer inferior a 2500g (OR = 4,94), índice de Apgar 1º minuto inferior a 7 (OR = 0,75), índice de Apgar 5º minuto inferior a 7 (OR = 0,76). As variáveis: idade materna, tipo de gestação, tipo de parto, consultas pré-natais e sexo do recém-nascido não se mostraram significativas para a mortalidade neonatal nesse estudo. O coeficiente de mortalidade infantil encontrado para essa região foi de 13,4 por mil nascidos vivos.

Andrade et al. (2006) verificaram se as condições de vida e infra-estrutura dos municípios do estado do Paraná no período de 1997 a 2001 se relacionavam com a taxa de mortalidade infantil e seus componentes, utilizando a técnica de componentes principais multivariada. Foi realizado um estudo de agregados com 399 municípios agrupados em cinco clusters de condições de vida. O levantamento dos dados foi obtido a partir do Censo Demográfico de 2000. Os autores descreveram que a taxa média de mortalidade infantil no estado foi de 19,3 por mil, a neonatal de 12,0 por mil e a pós-neonatal de 7,3 por mil nascidos vivos. A taxa de mortalidade infantil pós-neonatal oscilou de 8,7 (clusters de piores condições) para 6,3 por mil nascidos vivos (clusters de melhores condições). Em relação à componente neonatal, os clusters de melhor e pior situação refletiram as menores taxas, sendo que a taxa do cluster de melhor situação foi de 11,8 por mil e do de pior, foi de 11,4 por mil nascidos vivos.

Silva et al. (2006) analisaram os fatores de risco para a mortalidade infantil de crianças menores de um ano de vida para o município de Maracanaú no estado do Ceará, nos anos de 2000 a 2002, utilizando a *linkage* dos dados do SINASC e SIM. O levantamento dos dados foi elaborado pelas equipes do Programa de Saúde da Família (PSF). Os fatores de risco foram estimados por meio da análise de regressão logística bivariada e multivariada, e a *odds ratio* (OR) foi utilizada como medida de risco. Os autores observaram que 63,0% dos óbitos ocorreram no período neonatal (menos de 28 dias de vida), sendo 56% no período precoce (menos de 7 dias), 44% no período tardio (7 a 28 dias) e 37,0% no período pós neonatal (28

dias a 1 ano). A taxa encontrada de um nascido vivo morrer antes de completar um ano de vida foi de 10,7 por mil nascidos vivos, sendo 6,7 por mil no período neonatal. Os fatores de risco associados à mortalidade infantil foram: peso ao nascer inferior a 2500g (OR = 3,16); idade gestacional inferior a 37 semanas (OR = 2,70); índice de Apgar 1º minuto inferior a 7 (OR = 4,40); índice de Apgar 5º minuto inferior a 7 (OR = 5,56) e consultas pré-natais inferiores a 7 (OR = 2,05). As variáveis: tipo de gravidez, tipo de parto, sexo do bebê, número de filhos, idade materna e grau de instrução e local do nascimento não se mostraram significativamente associadas aos fatores de risco para a mortalidade infantil nesse estudo.

Duarte (2007) realizou uma revisão sistemática de artigos científicos publicados entre 1998 e 2006 em três relevantes bases de dados: MEDLINE (*Medical Literature Analysis and Retrieval System Online*), LILACS (Literatura Latino-americana em Ciências de Saúde) e PAHO (*Pan American Health Organization*), com o objetivo de analisar o comportamento da mortalidade infantil na última década e sua possível associação com as mudanças na organização e financiamentos do SUS. Para a autora, o CMI é considerado um eficiente indicador das condições de vida de uma população, além de ser facilmente calculado. Os CMIs são classificados conforme a proximidade ou distância de valores já obtidos em sociedades mais desenvolvidas; consideram-se altas taxas quando os valores encontrados são superiores a 50 por mil; médias taxas, entre 20 e 49 por mil e baixas taxas, quando inferiores a 20 por mil nascidos vivos. Foram analisados 59 artigos divididos de acordo com os temas: mortalidade infantil (28 artigos) ou seus componentes por idade e causa (31 artigos). As variáveis explicativas utilizadas foram: taxas de fecundidade total e natalidade, taxa de desemprego, índice de analfabetismo, variação percentual do PIB *per capita*, o índice de Gini, a variação da proporção de domicílios ligados à rede de abastecimento de água de esgotamento sanitário ou que possuíam fossa séptica, o número de estabelecimentos de saúde por mil habitantes e a cobertura vacinal anti-sarampo nos menores de um ano. Segundo a autora, a mortalidade infantil apresentou tendência decrescente (-59,3%) e forte correlação com a maioria dos indicadores analisados. Para ela, os estudos produzidos até o momento não possibilitam avaliar com clareza o provável impacto das alterações produzidas na organização e financiamento do SUS, sendo evidente a necessidade de monitoramento do indicador e a importância dos estudos locais, principalmente nos municípios com informações incipientes.

Nascimento et al. (2008) realizaram um estudo para identificar os possíveis fatores de risco para óbitos de menores de um ano de vida em Salvador (2000), por meio de um estudo caso-controle, utilizando a técnica de *linkage* dos bancos SINASC e SIM. Para a descrição dos fatores de risco de óbitos foi utilizada a análise de regressão logística bivariada e

multivariada e como medida de risco, a *odds ratio* (OR). Dos bebês que nasceram no ano de 2000, 2,75% evoluíram a óbito nos primeiros onze meses de vida. O coeficiente de mortalidade infantil encontrado foi de 24,4 óbitos por mil nascidos vivos. Segundo os autores, aproximadamente 66% dos óbitos ocorreram no período neonatal, 34,5% nas primeiras 24 horas de vida. Cerca de 59% dos bebês apresentaram peso inferior a 2500g e 40% peso entre 1000 e 2499g. As variáveis significativas associadas ao modelo incluem: idade gestacional inferior a 37 semanas (OR = 24,30), ocupação materna: para empregadas domésticas (OR = 3,65) e para donas de casa ou estudantes (OR = 2,67), partos em unidade do setor público (OR = 2,80), número não adequado de consultas pré-natais (OR = 2,50) e baixo peso ao nascer (OR = 1,89).

Em recente artigo publicado por Zanini et al. (2009) sobre a tendência da mortalidade infantil no RS, no período de 1994 a 2004, foi realizado um estudo ecológico longitudinal em que, além da análise de regressão logística clássica, utilizou-se a modelagem multinível para identificar os fatores de risco no nível do nascido vivo e também no nível da microrregião de moradia da mãe. Entre os resultados obtidos, destaca-se a constatação do declínio (-2,2%) dos CMI's de 1994 (19,25 por mil) para 2004 (15,16 por mil) no RS. Este declínio também ocorreu nas respectivas microrregiões, mas com diferenças significativas entre elas.

No próximo capítulo, serão apresentados os métodos utilizados para o desenvolvimento deste trabalho.

3 MATERIAL E MÉTODOS

As técnicas estatísticas empregadas no presente estudo foram a análise descritiva e as análises de regressão logística simples e múltipla.

3.1 Caracterização do estudo

Este estudo longitudinal compreende os registros de nascimentos e óbitos de crianças de mães residentes nos municípios pertencentes à 4ª Coordenadoria Regional de Saúde do Rio Grande do Sul, sendo que as informações utilizadas sobre os nascidos vivos (NV) foram provenientes das Declarações de Nascido Vivo (DN), cujas informações constam no Sistema de Informações sobre Nascidos Vivos (SINASC), e os óbitos de bebês menores de um ano de vida, provenientes das Declarações de Óbitos (DO), no ano de 2006, que constam no Sistema de Mortalidade (SIM).

A população considerada neste estudo foi de 6983 NV. A identificação dos óbitos foi realizada vinculando-se o número da DN comum às bases de nascimentos e óbitos (*linkage*), permitindo que as probabilidades de morte pudessem ser calculadas em função da presença ou ausência de alguns fatores de risco presentes nas DNs (ZANINI, 2007).

3.2 Obtenção dos dados

Os dados foram obtidos no site do DATASUS, utilizando-se o programa TAB para *Windows* – Tabwin (<http://www.datasus.gov.br/tabwin/tabwin.htm>). Inicialmente foi obtido um arquivo de dados com extensão DBC, que corresponde a arquivos em padrão DBase, extensão DBF compactados. Utilizou-se, então, a função “arquivo/comprime/expande.DBF” para obter os dados em DBF. Posteriormente o arquivo expandido foi aberto no programa SPSS 13.0, como arquivo do tipo DBase (DBF).

3.3 Ajustes do banco de dados

Não foi possível a vinculação de oito Declarações de Óbito, devido a não vinculação do número da DN nos dois sistemas (SINASC e SIM).

Uma das limitações deste estudo diz respeito ao fato de considerar apenas os bebês que nasceram e morreram no ano de 2006. Além disso, conforme Souza (2004), para a identificação dos erros de registro dos dados, foi considerada a classificação do NV com base no peso e idade gestacional. Essa classificação considera que: NV com idade gestacional inferior a 22 semanas; peso inferior a 500g; idade gestacional de 22 a 27 semanas e peso superior a 1500g; idade gestacional de 28 a 31 semanas e peso superior a 2500g; idade gestacional superior a 37 semanas e peso inferior a 1500g são informações preenchidas de forma errônea, sendo que essas informações podem estar relacionadas a um nascido morto. Segundo Margotto (2001), não há viabilidade biológica para um recém-nascido com peso inferior a 500g ou com idade gestacional inferior a 22 semanas, sendo que recém-nascidos com essas características não deveriam ser incluídos no SINASC. Conforme esses critérios foram excluídos 20 registros que se enquadraram em uma dessas situações, restando 6.963 NV no banco de dados para a análise.

3.4 Variável dependente

Neste estudo a variável dependente (desfecho dicotômico) é a ocorrência ou não de óbito em crianças menores de um ano de vida, ou seja, óbito infantil.

3.5 Variáveis independentes ou covariáveis

Foram consideradas como covariáveis aquelas disponíveis no SINASC com potencial de serem fatores de risco associados à mortalidade infantil.

Assim, foram analisadas as covariáveis associadas às características maternas: idade (menos de 20 anos (mães adolescentes), 20 a 34 anos e 35 anos ou mais); paridade (2 filhos

ou menos, 3 filhos ou mais, vivos e mortos de gestações anteriores); escolaridade (0 a 3 anos, 4 a 11 anos, 12 anos concluídos ou mais de estudo); estado civil (casada ou união consensual, outro); número de consultas pré-natais (nenhuma, de 1 a 6, 7 ou mais); tipo de gravidez (única, múltipla); tipo de parto (vaginal, cesáreo); local do nascimento (hospital, outro) e idade gestacional (pré-termo (< 37 semanas), a termo (\geq 37 semanas)).

As covariáveis relativas às características do nascido vivo foram: peso ao nascer (< 2500g (baixo peso), \geq 2500g); índice de Apgar 1º minuto (< 7, 7 ou 8, 9 ou 10); índice de Apgar 5º minuto (< 7, 7 ou 8, 9 ou 10); sexo (masculino, feminino); raça/cor (branca, outra); malformação congênita e/ou anomalia cromossômica (sem anomalia, com anomalia).

O índice de Apgar consiste na avaliação de 5 itens do exame físico (cor da pele, frequência cardíaca, esforço respiratório, irritabilidade reflexa e tônus muscular) do recém-nascido com um e com cinco minutos de vida (APGAR, 1958).

Algumas covariáveis foram recategorizadas em faixas, conforme apresentado anteriormente, com base na revisão bibliográfica, com o objetivo de possibilitar uma melhor interpretação e a discussão dos resultados.

As covariáveis cujas frequências nas categorias eram pequenas, foram reagrupadas em novas categorias, sendo que o código atribuído à faixa de referência foi zero, e os códigos para as faixas de maior risco foram considerados números inteiros de 1 a n.

Para as covariáveis contínuas: peso, índice de Apgar 1º minuto e índice de Apgar 5º minuto, foi construído o histograma para a avaliação do seu comportamento e também foi feito o cálculo de algumas medidas descritivas que auxiliaram nesta avaliação. Posteriormente essas covariáveis foram agrupadas em faixas com o objetivo de facilitar a interpretação dos coeficientes no modelo a ser ajustado aos dados.

Foram realizados os cruzamentos entre as covariáveis categóricas e a variável dependente dicotômica. As tabelas resultantes destes cruzamentos possibilitaram a avaliação das possíveis categorias de risco para o desfecho (y). Dos cruzamentos foram obtidos os percentuais de óbitos, por categoria, para cada covariável, sendo também calculados os coeficientes de mortalidade infantil para as mesmas categorias.

As associações entre as covariáveis categóricas e o desfecho podem ser verificadas pelo teste de independência do qui-quadrado. Se o resultado fosse significativo, considerando um nível de significância de 25% (α), a covariável era considerada como um potencial fator de risco para entrar no modelo múltiplo.

Neste estudo, optou-se pela análise de regressão logística univariada, que fornece a mesma informação que o teste de independência do qui-quadrado para o critério de seleção das covariáveis para o modelo múltiplo.

3.6 Modelagem

Para definir o modelo mais adequado para os dados, foram estabelecidos os seguintes passos:

1º) Seleção das variáveis candidatas ao modelo de regressão múltipla: foi realizada a análise de regressão logística univariada de cada covariável com o desfecho. Para a seleção das variáveis que entrarão no modelo, pode-se utilizar: o método *enter* ou um método automático *stepwise*, sendo que as variáveis são selecionadas para serem incluídas ou excluídas do modelo, com base somente no critério estatístico. A covariável que apresentou o *p-value* menor ou igual a 25% no teste univariado foi incluída no modelo múltiplo, ou seja, se $p \leq 0,25$, a covariável foi significativa e houve evidência de associação desta covariável com o desfecho. Se $p > 0,25$, a covariável foi desconsiderada para ser testada juntamente com as outras no modelo múltiplo. A regressão logística univariada possibilitou a estimação das razões de chance (OR bruto) e os respectivos intervalos de confiança de 95% (IC 95%);

2º) Para a análise de regressão logística múltipla, foi utilizada a estratégia de análise não-condicional que incluiu, simultaneamente no modelo, todas as covariáveis independentes significativas da análise de regressão logística univariada, método *enter* e nível de significância de 0,25, verificados pelo teste de Wald;

3º) Foram eliminadas do modelo múltiplo, uma a uma, as covariáveis não significativas ($p > 0,05$), verificando-se o efeito da saída de cada covariável nos coeficientes (β_i) que permaneciam no modelo. A cada retirada de uma covariável (com maior valor de p), o modelo era processado para que os coeficientes fossem reajustados para aquele novo conjunto de covariáveis. Para verificar o efeito da retirada de uma covariável do modelo, fez-se o teste da razão de verossimilhança. Neste estudo optou-se pelo teste de Wald para decidir pela permanência das covariáveis no modelo;

4º) O modelo final foi o que apresentou todas as covariáveis significativas em nível de 5% de significância e cujo IC 95% não incluiu o valor de OR = 1;

5º) Com o modelo encontrado, foram estimadas algumas probabilidades da ocorrência do desfecho para os valores das covariáveis (riscos).

As análises estatísticas foram realizadas com o auxílio do aplicativo computacional SPSS 13.0 (SPSS Inc., Chicago, EUA).

No próximo capítulo serão apresentados os resultados da análise dos dados, bem como a discussão dos mesmos.

4 RESULTADOS E DISCUSSÕES

No banco de dados considerado para a análise, 20 NV (0,3%) apresentaram registros incorretos e foram excluídos da análise, restando então 6963 NV (99,7%).

Do total de nascidos vivos de mães residentes nos municípios pertencentes a 4ª CRS, no ano de 2006, registrados no SINASC, aproximadamente 3,06% dos bebês evoluíram a óbito no período neonatal precoce (0 a 6 dias), 1,02% no período neonatal tardio (7 a 27 dias) e 95,92% no período pós-neonatal (28 dias a uma ano).

Na Tabela 2, é apresentada a distribuição do peso ao nascer, em gramas e por faixas, segundo a idade gestacional em semanas. Verifica-se que 9,0% dos NV apresentaram baixo peso ao nascer (< 2500g), sendo este considerado um dos principais fatores de risco associados à mortalidade infantil. Pode-se observar, também, que, dos 680 nascidos vivos prematuramente, 389 (57,2%) nasceram com baixo peso. Dos nascidos a termo, 239 (3,8%) dos bebês apresentaram baixo peso, evidenciando a importância do controle da idade gestacional na ocorrência do baixo peso.

Tabela 2 – Distribuição do peso ao nascer segundo a idade gestacional dos nascidos vivos, SINASC/RS, 2006

	Idade gestacional (semanas)						Total (%)
	Pré-termo			Pré-termo total	A termo	Pós-termo	
Peso (g)	22 a 27	28 a 31	32 a 36	< 37	37 a 41	≥ 42	
500 a 1499	43	38	22	103	-	-	103 (1,48)
1500 a 2499	-	24	262	286	239	-	525 (7,56)
2500 a 2999	-	-	203	203	1391	10	1604 (23,09)
3000 a 3999	-	-	87	87	4237	36	4360 (62,76)
> 4000	-	-	1	1	346	8	355 (5,11)
Total	43	62	575	680	6213	54	6947 (100,00)

Nota: existem 15 dados faltantes para a idade gestacional e 1 para o peso ao nascer.

Comparando-se os resultados obtidos sobre a incidência de baixo peso ao nascer com outros estudos que utilizaram a mesma abordagem temática, observa-se que o baixo peso ao nascer é considerado um forte determinante de risco relacionado a óbitos antes do primeiro ano de vida (SARINHO et al., 2001; MARTINS e VELÁSQUEZ-MELÉNDEZ, 2004). Esse determinante é mencionado como um indicador do risco social relacionado às precárias

condições socioeconômicas e nutricionais de uma população, mostrando, assim, a complexidade dos fatores associados e a necessidade de um aprofundamento nos estudos das causas desse determinante (MORAIS NETO e BARROS, 2000).

Em estudo realizado por Moraes (2007) sobre a ocorrência do baixo peso ao nascer no Rio Grande do Sul, no ano de 2003, a autora constatou que: 8,4% dos NV apresentaram baixo peso ao nascer, sendo que 60,9% eram prematuros e 4,02% nascidos a termo apresentaram baixo peso. Os resultados obtidos no presente estudo para a 4ª CRS corroboram com os resultados encontrados pela autora para o estado do RS.

Na Tabela 3, são apresentados os prováveis fatores de risco para a mortalidade infantil, segundo as características da mãe do NV.

Verificou-se que 1296 (18,6%) das mães eram adolescentes, ou seja, apresentavam idade inferior a 20 anos. Observou-se, também que 593 (8,6%) mães apresentavam um baixo nível de instrução escolar, sendo estas consideradas analfabetas ou semi-analfabetas. Em relação ao estado civil, verificou-se que 3975 (57,3%) mães não possuíam união estável durante o período gestacional. Em relação ao número de consultas pré-natais 2206 (32%) das futuras mães realizaram menos de seis consultas, sendo que o preconizado pelo MS é o mínimo de 7 consultas pré-natais. Dos nascidos vivos, 131 (1,9%) nasceram de gestações múltiplas, ou seja, de uma gestação única com dois ou mais bebês. O percentual de partos cesáreos foi de 55,7%, muito acima dos 15% recomendados pela Organização Mundial de Saúde.

O coeficiente de mortalidade infantil para os nascidos vivos pertencentes a 4ª CRS foi de 11,49‰, conforme Tabela 3, considerado baixo segundo a classificação descrita por Duarte (2007). Pode-se observar, também, nas Tabelas 3 e 4, que os valores dos CMIs são maiores para as faixas das covariáveis que representam potencial risco para a mortalidade infantil. Alguns dos valores mais elevados para os CMIs são descritos para: mães adolescentes (18,52‰), mães com alta paridade (17,34‰), baixa escolaridade (18,55‰), nenhuma consulta pré-natal (81,08‰), gravidez múltipla (76,33‰), nascimento fora do hospital (161,29‰) e nascimento prematuro (75,00‰).

Partindo-se para uma análise em termos nacionais, Duarte (2007) contextualiza o Brasil como um país de taxas elevadas, pois afirma que, apesar do decréscimo da mortalidade infantil, o país, assim como o estado do RS, apresenta diferenças importantes dentro de seu território. Fundamentando-se nos dados de Victora (2001) e na pesquisa de Szwarcwald et al. (2002), as maiores taxas de mortalidade infantil registradas são no Nordeste e Norte brasileiro, enquanto as menores são no Sudeste e região Sul do país.

Tabela 3 – Análise descritiva dos fatores de risco para a mortalidade infantil, segundo características maternas, RS, 2006 (n = 6963)

Variáveis independentes	Nascidos Vivos (%)	Sobrevida (%)	Óbitos (%)	CMI (‰)
Idade materna (anos) *12-52; 26,27±6,87				
20 a 34	4669 (67,1)	4626 (99,1)	43 (0,9)	9,21
≥ 35	998 (14,3)	985 (98,7)	13 (1,3)	13,03
< 20	1296 (18,6)	1272 (98,1)	24 (1,9)	18,52
Paridade				
≤ 2	5891 (85,0)	5830 (99,0)	61 (1,0)	10,35
≥ 3	1038 (15,0)	1020 (98,3)	18 (1,7)	17,34
Escolaridade (anos)				
12 ou mais	1140 (16,5)	1128 (98,9)	12 (1,1)	10,53
4 a 11	5181 (74,9)	5124 (98,9)	57 (1,1)	11,00
0 a 3	593 (8,6)	582 (98,1)	11 (1,9)	18,55
Estado civil				
Casada ou união consensual	2960 (42,7)	2937 (99,8)	23 (0,8)	7,77
Outro	3975 (57,3)	3919 (98,6)	56 (1,4)	14,09
Consultas pré-natais				
7 ou mais	4693 (68,0)	4660 (99,3)	33 (0,7)	7,03
1 a 6	2095 (30,4)	2059 (98,3)	36 (1,7)	17,18
Nenhuma	111 (1,60)	102 (91,9)	9 (8,1)	81,08
Tipo de gravidez				
Única	6832 (98,1)	6762 (99,0)	70 (1,0)	10,24
Múltipla	131 (1,9)	121 (92,4)	10 (7,6)	76,33
Tipo de parto				
Vaginal	3088 (44,3)	3054 (98,9)	34 (1,1)	11,01
Cesáreo	3875 (55,7)	3829 (98,8)	46 (1,2)	11,87
Local do nascimento				
Hospital	6932 (99,6)	6857 (98,9)	75 (1,1)	10,82
Outro	31 (0,4)	26 (83,9)	5 (16,1)	161,29
Idade gestacional (semanas)				
A termo (≥ 37)	6267 (90,2)	6238 (99,5)	29 (0,5)	4,63
Pré-termo (< 37)	680 (9,8)	629 (92,5)	51 (7,5)	75,00
Total	6963(100,0)	6883 (98,9)	80 (1,1)	11,49

*mínimo – máximo; média ± desvio padrão; excluídos da análise os casos sem informação: paridade (34), escolaridade (49), estado civil (525), consultas pré-natais (64), idade gestacional (16); CMI (‰): coeficiente de mortalidade infantil por mil nascidos vivos.

Na Tabela 4, tem-se os potenciais fatores de risco para a mortalidade infantil, segundo características do nascido vivo.

Verificou-se que o peso médio dos bebês ao nascer foi de 3168,5 gramas (desvio padrão igual a 61,1). O valor do Apgar para os nascidos vivos no 1º minuto apresentou a média de 8,6 (desvio padrão igual a 1,2) e para o Apgar 5º minuto foi de 9,7 (desvio padrão

igual a 0,7). O percentual de nascidos vivos em relação à presença de malformação congênita foi de 1,0%.

Outros valores elevados do CMI foram para: NV de baixo peso (80,32‰), NV com índice de Apgar no 1º minuto inferior a 7 (123,56‰), índice de Apgar no 5º minuto inferior a 7 (420,00‰) e nascidos vivos com anomalias congênitas (279,41‰).

Tabela 4 – Análise descritiva dos fatores de risco para a mortalidade infantil, segundo características do nascido vivo, RS, 2006 (n = 6963)

Variáveis independentes	Nascidos Vivos (%)	Sobrevida (%)	Óbitos (%)	CMI (‰)
Peso ao nascer (g)* 500-5380; 3168,5±61,1				
≥ 2.500	6327 (90,9)	6298 (99,5)	29 (0,5)	4,58
< 2.500	635 (9,1)	584 (92,0)	51 (8,0)	80,32
Apgar 1º minuto* 0-10; 8,6±1,2				
9 ou 10	4775 (68,8)	4763 (99,7)	12 (0,3)	2,51
7 ou 8	1820 (26,2)	1799 (98,8)	21 (1,2)	11,54
< 7	348 (5,0)	305 (87,6)	43 (12,4)	123,56
Apgar 5º minuto* 0-10; 9,7±0,7				
9 ou 10	6641 (95,6)	6604 (99,4)	37 (0,6)	5,57
7 ou 8	254 (3,7)	235 (92,5)	19 (7,5)	74,80
< 7	50 (0,7)	29 (58,0)	21 (42,0)	420,00
Sexo				
Masculino	3564 (51,0)	3516 (98,7)	48 (1,3)	13,47
Feminino	3397 (49,0)	3366 (99,1)	31 (0,9)	9,13
Raça/cor				
Branca	6281 (90,2)	6210 (98,9)	71 (1,1)	11,30
Outra	679 (9,8)	671 (98,8)	8 (1,2)	11,78
Anomalias				
Sem	6886 (99,0)	6827 (99,1)	59 (0,9)	8,57
Com	68 (1,0)	49 (72,1)	19 (27,9)	279,41
Total	6963(100,0)	6883 (98,9)	80 (1,1)	11,49

*mínimo – máximo; média ± desvio padrão; CMI (‰): coeficiente de mortalidade infantil por mil nascidos vivos.

As covariáveis testadas na análise de regressão logística univariada foram: idade materna, paridade, escolaridade, estado civil, número de consultas pré-natais, tipo de gravidez, tipo de parto, local do nascimento, idade gestacional, peso ao nascer, índice de Apgar 1º minuto, índice de Apgar 5º minuto, sexo, raça/cor e anomalias congênitas, todas com $p \leq 0,25$.

Na Tabela 5, são apresentados os resultados da análise de regressão logística univariada para os nascidos vivos, considerando-se as covariáveis correspondentes às características da mãe. Observa-se que as variáveis significativas ($p \leq 0,25$) foram: idade materna, paridade, estado civil, número de consultas pré-natais, tipo de gravidez, tipo de parto, local do nascimento e idade gestacional, selecionadas para serem testadas no modelo de regressão logística múltiplo.

Tabela 5 – Análise de regressão logística univariada, segundo características maternas, tendo como desfecho o óbito, RS, 2006 (n = 6963)

Variáveis independentes	OR Bruto	IC 95%	p
Idade materna (anos)			0,021*
20 a 34	1		
≥ 35	1,420	0,761 - 2,650	
< 20	2,030	1,227 - 3,357	
Paridade			0,053*
≤ 2	1		
≥ 3	1,687	0,993 - 2,865	
Escolaridade (anos)			0,257
12 ou mais	1		
4 a 11	1,046	0,559 - 1,955	
0 a 3	1,777	0,779 - 4,051	
Estado civil			0,016*
Casada ou união consensual	1		
Outro	1,825	1,120 - 2,972	
Consultas pré-natais			< 0,001*
7 ou mais	1		
1 a 6	2,469	1,535 - 3,971	
Nenhuma	12,460	5,811 - 26,715	
Tipo de gravidez			< 0,001*
Única	1		
Múltipla	7,983	4,018 - 15,862	
Tipo de parto			0,738
Vaginal	1		
Cesáreo	1,079	0,691 - 1,685	
Local do nascimento			< 0,001*
Hospital	1		
Outro	17,582	6,574 - 47,024	
Idade gestacional (semanas)			< 0,001*
A termo (≥ 37)	1		
Pré-termo (< 37)	17,441	10,975 - 27,715	

OR: Odds ratio; OR Bruto = 1: Categoria de referência; IC 95%: Intervalo de Confiança de 95%; * $p \leq 0,25$.

O valor da significância da covariável foi verificado pelo valor do p global e não pelos valores de p por faixas.

Na Tabela 6, são apresentados os resultados da análise de regressão logística univariada considerando-se as covariáveis correspondentes às características do nascido vivo. Observa-se que as variáveis significativas ($p \leq 0,25$) foram: peso ao nascer, índice de Apgar 1º minuto, índice de Apgar 5º minuto, sexo, raça/cor e anomalias, selecionadas para serem testadas no modelo de regressão logística múltiplo.

Tabela 6 - Análise de regressão logística univariada, segundo características do nascido vivo, tendo como desfecho o óbito, RS, 2006 (n = 6963)

Variáveis independentes	OR Bruto	IC 95%	p
Peso ao nascer (g)			< 0,001*
≥ 2.500	1		
< 2.500	18,965	11,929 - 30,153	
Apgar 1º minuto			< 0,001*
9 ou 10	1		
7 ou 8	4,633	2,275 - 9,437	
< 7	55,959	29,205 - 107,221	
Apgar 5º minuto			< 0,001*
9 ou 10	1		
7 ou 8	14,431	8,175 - 25,474	
< 7	129,249	67,615 - 247,066	
Sexo			0,089*
Masculino	1		
Feminino	1,482	0,941 - 2,334	
Raça/cor			0,911
Branca	1		
Outra	1,043	0,500 - 2,175	
Anomalias			< 0,001*
Sem	1		
Com	44,868	24,911 - 80,814	

OR: Odds ratio; OR Bruto = 1: Categoria de referência IC 95%: Intervalo de Confiança de 95%; * $p \leq 0,25$.

Inicialmente foram testadas, no modelo múltiplo, as 12 covariáveis selecionadas (excluídas a escolaridade da mãe, o tipo de parto e a raça/cor do NV). Para o modelo encontrado, foi obtido o valor de $-2 \ln$ da verossimilhança de 435,969 com 16 graus de liberdade (g.l.).

Na etapa posterior, foram excluídas do modelo, uma a uma, as covariáveis não significativas ($p > 0,05$): sexo do NV, idade materna, paridade, número de consultas pré-natais e tipo de gravidez.

Na Tabela 7, são apresentados os valores do OR ajustado do modelo resultante, considerando-se as covariáveis correspondentes às características maternas e do NV significativas no modelo. Observa-se que todas as covariáveis apresentaram pelo menos uma categoria significativa ($p \leq 0,05$), sendo estas: estado civil, local do nascimento, idade gestacional, peso ao nascer, índice de Apgar 1º minuto, índice de Apgar 5º minuto e anomalias, representando os fatores de risco para a mortalidade infantil encontrados pelo ajuste deste modelo de regressão logística múltipla.

Tabela 7 – Análise de regressão logística múltipla, tendo como desfecho o óbito, RS, 2006 (n = 6963)

Variáveis independentes	OR Ajustado*	IC 95%	p
Estado civil			
Casada ou união consensual	1		
Outro	2,099	1,141 - 3,862	0,017
Local do nascimento			
Hospital	1		
Outro	86,322	9,493 - 784,969	< 0,001
Idade gestacional (semanas)			
A termo (≥ 37)	1		
Pré-termo (< 37)	3,595	1,732 - 7,462	0,001
Peso ao nascer (g)			
≥ 2.500	1		
< 2.500	4,130	1,991 - 8,565	< 0,001
Apgar 1º minuto			
9 ou 10	1		
7 ou 8	3,326	1,558 - 7,099	0,002
< 7	8,978	3,605 - 22,357	< 0,001
Apgar 5º minuto			
9 ou 10	1		
7 ou 8	1,860	0,848 - 4,077	0,121
< 7	13,544	5,018 - 36,557	< 0,001
Anomalias			
Sem	1		
Com	18,486	8,531 - 40,061	< 0,001

* *Odds ratio* ajustado para as outras covariáveis da tabela por meio da regressão logística múltipla; OR ajustado =1: categoria de referência; IC 95%: Intervalo de Confiança de 95%.

Para o modelo que inclui as 7 covariáveis apresentadas na Tabela 7, onde todas são significativas, o valor obtido para a quantidade $-2 \ln$ da verossimilhança é igual a 471,213 com 9 g.l. Pode-se observar na Tabela 7 que, embora significativo ($p < 0,001$), o valor da OR da faixa de risco da covariável local do nascimento é bastante elevado, apresentando um intervalo de confiança muito amplo, indicando um valor expressivo para a estimativa do erro padrão deste coeficiente, o que, segundo Hosmer e Lemeshow (1989), é um forte indicativo de problemas numéricos na regressão logística (overfitting).

Avaliando o valor da *deviance* deste modelo com a covariável local do nascimento (471,213; 9 g.l.) em relação ao modelo sem ela (477,898; 8 g.l.), tem-se $G = 477,898 - 471,213 = 6,685$ com 1 g.l. Comparando-se o valor de G com o valor $\chi^2_{5\%,1} = 3,84$, verifica-se um resultado significativo, pois $G > \chi^2$, indicando que o coeficiente da covariável local do nascimento é significativo, devendo ser mantido no modelo. Contrariando este resultado, a autora optou em retirá-la do modelo pelo motivo explicado no parágrafo anterior.

Portanto, o modelo final foi aquele cujos coeficientes estão apresentados na Tabela 8. Observa-se que as faixas das covariáveis índice de Apgar no 1º minuto e no 5º minuto foram reagrupadas em função da significância das mesmas neste modelo.

Com o modelo obtido, pode-se concluir que o risco de um nascido vivo morrer antes de completar um ano de vida é aumentado se a mãe não possui união estável, se for prematuro, tiver baixo peso, se apresentar um índice de Apgar no 1º minuto e no 5º minuto inferior a 9 e apresentar anomalia congênita. Victora et al. (1996), em estudos de coortes realizados em Pelotas, RS, nos anos de 1982 e 1993, também encontraram como significativas, entre outras causas para a mortalidade infantil, o baixo peso ao nascer, a prematuridade e a malformação congênita.

Considerando o peso ao nascer, a OR = 4,67 indica que a ocorrência de mortalidade nos bebês que nasceram com baixo peso ($< 2500g$) é quatro vezes a ocorrência de mortalidade em bebês que não nasceram com baixo peso ($\geq 2500g$), ou ainda, o baixo peso aumenta em 300% o risco de óbito.

Silva et al. (2006) encontraram os seguintes fatores de risco para a mortalidade infantil para o município de Maracanaú no estado do Ceará nos anos de 2000 a 2002: peso ao nascer inferior a 2500g (OR = 3,16); idade gestacional inferior a 37 semanas (OR = 2,70); índice de Apgar 1º minuto inferior a 7 (OR = 4,40); índice de Apgar 5º minuto inferior a 7 (OR = 5,56) e número de consultas pré-natais inferiores a 6 (OR = 2,05), sendo esses resultados inferiores aos encontrados no presente estudo.

Tabela 8 – Análise de regressão logística múltipla – MODELO FINAL, tendo como desfecho o óbito, RS, 2006 (n = 6963)

Variáveis independentes	OR Ajustado*	IC 95%	p
Estado civil			
Casada ou união consensual	1		
Outro	1,984	1,117 - 3,524	0,019
Idade gestacional (semanas)			
A termo (≥ 37)	1		
Pré-termo (< 37)	3,256	1,624 - 6,530	$< 0,01$
Peso ao nascer (g)			
≥ 2.500	1		
< 2.500	4,667	2,325 – 9,369	$< 0,001$
Apgar 1º minuto			
9 ou 10	1		
≤ 8	3,926	1,915 – 8,051	$< 0,001$
Apgar 5º minuto			
9 ou 10	1		
≤ 8	5,613	3,112 – 10,123	$< 0,001$
Anomalias			
Sem	1		
Com	20,402	9,773 – 42,595	$< 0,01$

* *Odds ratio* ajustado para as outras covariáveis da tabela por meio da regressão logística múltipla; OR ajustado = 1: categoria de referência; IC 95%: Intervalo de Confiança de 95%; $\beta_0 = -7,152$.

O modelo final, obtido com a regressão logística múltipla, é representado pela aplicação da equação (23) cujas covariáveis representam os fatores de risco para a mortalidade infantil dos nascidos vivos na região da 4ª CRS, no ano de 2006, sendo eles: estado civil, idade gestacional, peso ao nascer, índice de Apgar 1º minuto, índice de Apgar 5º minuto e anomalias congênitas.

Em relação aos determinantes para a mortalidade infantil obtidos por meio do modelo ajustado, salienta-se a ocorrência significativa de quatro covariáveis: o baixo peso ao nascer, a prematuridade e os índices de Apgar 1º e 5º minuto, fatores citados por Victora et al., 1996; Sarinho et al., 2001; Martins e Velasquez-Méendez, 2004 e reforçados como sendo potenciais fatores de risco associados à mortalidade infantil.

Os coeficientes (β_i 's) da equação (23) correspondem aos valores do ln (OR ajustado) da Tabela 8, para cada categoria das covariáveis.

$$\ln odds = \beta_0 + \beta_1 X_{\text{estcivil}} + \beta_2 X_{\text{gestação}} + \beta_3 X_{\text{peso ao nascer}} + \beta_4 X_{\text{Apgar1}} + \beta_5 X_{\text{Apgar5}} + \beta_6 X_{\text{anomalia}} \quad (23)$$

$$\ln odds = -7,152 + 0,685 X_{\text{estcivil}} + 1,181 X_{\text{gestação}} + 1,540 X_{\text{peso ao nascer}} + 1,368 X_{\text{Apgar1}} + 1,725 X_{\text{Apgar5}} + 3,016 X_{\text{anomalia}}$$

A seguir, são exemplificadas três situações hipotéticas que ilustram a utilização do modelo múltiplo encontrado (Tabela 8) para estimar a probabilidade da ocorrência da mortalidade infantil.

Situação 1: Qual a probabilidade de óbito de um nascido vivo, com menos de um ano de vida, que, ao nascer, apresentar: mãe sem união estável, prematuro, baixo peso (peso < 2500g), índice de Apgar 1º minuto menor que 8, índice de Apgar 5º minuto menor que 8 e anomalia congênita?

$$\ln odds = -7,152 + 0,685 \times (1) + 1,181 \times (1) + 1,540 \times (1) + 1,368 \times (1) + 1,725 \times (1) + 3,016 \times (1) = 2,363$$

$$P(Y/X) = \frac{1}{1 + e^{-(\ln odds)}} = \frac{1}{1 + e^{-2,363}} = 0,91396 \dots = 91,4\%$$

Situação 2: Qual a probabilidade de óbito de um nascido vivo, com menos de um ano de vida, que, ao nascer, apresentar: mãe sem união estável, prematuro, baixo peso (peso < 2500g), índice de Apgar 1º minuto menor do que 8, índice de Apgar 5º minuto maior do que 8, sem anomalia congênita?

$$\ln odds = -7,152 + 0,685 \times (1) + 1,181 \times (1) + 1,540 \times (1) + 1,368 \times (1) + 1,725 \times (0) + 3,016 \times (0) = -2,378$$

$$P(Y/X) = \frac{1}{1 + e^{-(\ln odds)}} = \frac{1}{1 + e^{+2,378}} = 0,08487 \dots = 8,5\%$$

Situação 3: Qual a probabilidade de óbito de um nascido vivo, com menos de um ano de vida, que, ao nascer, apresentar: mãe com união estável, não prematuro, peso normal (peso \geq 2500g), índice de Apgar 1º minuto 9 ou 10, índice de Apgar 5º minuto 9 ou 10 e sem anomalia congênita?

$$\ln odds = - 7,152 + 0,685 \times (0) + 1,181 \times (0) + 1,540 \times (1) + 1,368 \times (0) + 1,725 \times (0) + 3,016 \times (0) = -7,152$$

$$P(Y/X) = \frac{1}{1 + e^{-(\ln odds)}} = \frac{1}{1 + e^{+7,152}} = 0,00078 \dots = 0,08\%$$

Com base nessas três situações hipotéticas, observa-se o ajuste do modelo encontrado, pois colocadas duas situações extremas: na situação 1, o nascido vivo apresenta todas as covariáveis com valores nas faixas de risco, e o modelo estima uma elevada probabilidade de óbito, 91,4%; na situação 3, o nascido vivo apresenta todas as covariáveis fora da faixa de risco, e o modelo estima uma pequena probabilidade de óbito, ou seja, 0,08%; na situação 2, o nascido vivo apresenta algumas covariáveis na faixa de risco, sendo a probabilidade encontrada pelo modelo de 8,5% de óbitos.

5 CONCLUSÃO

Ao término deste estudo, conclui-se que foram atingidos os objetivos propostos, pois o desenvolvimento desta pesquisa contribui de forma relevante para a compreensão dos procedimentos de cálculos envolvidos na técnica de análise de regressão logística múltipla binária, assim como para a identificação dos fatores de risco relevantes associados à mortalidade infantil nos municípios pertencentes a 4ª CRS.

Considerando-se os resultados obtidos neste estudo, conclui-se que a existência de um banco de dados oficial, com acesso livre e gratuito, é de suma importância, proporcionando a realização de estudos científicos que tenham por objetivo identificar, neste caso, fatores de risco para eventos de grande relevância para a Saúde Pública, como a mortalidade infantil.

No que diz respeito ao desenvolvimento da pesquisa, o modelo ajustado para os dados foi ao encontro do principal objetivo desta pesquisa, ou seja, a identificação dos fatores de risco para a mortalidade infantil de nascidos vivos para a região central do estado do Rio Grande do Sul, cujas variáveis identificadas foram: o estado civil da mãe (mãe sem união estável), idade gestacional (inferior a 37 semanas), peso ao nascer (inferior a 2500 gramas), índice de Apgar no 1º minuto (inferior a 9), índice de Apgar no 5º minuto (inferior a 9) anomalias congênitas (presente). Os achados deste estudo confirmam a maioria dos resultados encontrados por outros autores para diversas regiões do Estado brasileiro, sendo que os modelos podem diferir quanto à inclusão de outras covariáveis, tais como: escolaridade da mãe, o número de consultas pré-natais, entre outros, devido às diferenças de condições econômicas, sociais, ambientais e principalmente de saúde nas regiões do estado do RS.

Com isso, acredita-se que estes resultados podem contribuir para orientação e planejamento das políticas públicas de saúde com ênfase nos cuidados às gestantes, no período de gestação e no momento do parto, orientando, desta forma, para uma especial atenção da mãe em relação ao seu recém-nascido em prol da sobrevivência dos bebês com uma maior qualidade vida.

Analisando-se os resultados mais relevantes deste estudo, sugere-se mais vigilância por parte dos serviços de saúde aos recém-nascidos em relação a intervenções que possibilitem uma redução nas taxas de baixo peso ao nascer, proporcionando maior atenção à gestante e aos cuidados no momento do parto, alcançando-se uma elevação nos índices do NV, refletidos nos valores dos índices de Apgar 1º minuto e no 5º minuto. Certamente esses

cuidados proporcionarão a redução das taxas de prematuridade, um dos fatores de risco mais importantes a serem considerados.

Estas conclusões evidenciam a necessidade de uma assistência em saúde mais adequada para cada um dos níveis de complexidade do sistema de saúde, bem como a garantia de assistência pré-natal adequada às gestantes e ao recém-nascido, com intuito de identificar precocemente eventuais doenças maternas e fetais, possibilitando, desta maneira, a sobrevivência dos recém-nascidos.

Assim, este trabalho poderá servir de referência para futuras pesquisas realizadas no campo das ciências da saúde, da mesma forma que poderá auxiliar e fundamentar outros trabalhos que utilizem a análise de regressão logística múltipla.

6 REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, M. F. et al. Mortalidade neonatal no município de São Paulo: influência do peso ao nascer e de fatores sócio-demográficos e assistenciais. **Revista Brasileira de Epidemiologia**, São Paulo, v. 5, n. 1, p. 93-107, 2002.

ANDRADE, S. M. et al. Condições de vida e mortalidade infantil no Estado do Paraná, Brasil. (1997 - 2001). **Cadernos de Saúde Pública**, Rio de Janeiro, v. 22, n. 1, jan. 2006.

APGAR. V. **Os Primeiros dez minutos da vida**: a escala de Apgar, 1958. Medicina Intensiva. Disponível em: <<http://www.medicinaintensiva.com.br/apgar.htm>>. Acesso em 28 jun.2009.

ARAÚJO, B. F.; BOZZETTI, M. C.; TANAKA, A. C. A. Mortalidade neonatal precoce no município de Caxias do Sul: um estudo de coorte. **Jornal de Pediatria**, Rio de Janeiro, v. 76, n. 3, 2000.

COX, D. R.; HINKLEY D. V. **Monographs on Statistics and Applied Probability: Analysis of Binary Data** – Chapman and Hall, London, 1969.

DUARTE, C. M. R. Reflexos das políticas de saúde sobre as tendências da mortalidade infantil no Brasil: revisão da literatura sobre a última década. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 23, n. 7, p. 1511-1528, jul., 2007.

HAUK, W. W.; DONNER, Wald's test as applied to hypothesis in logistic analysis. **Journal of the American Statistical Association**. v. 72, p.851-853, 1977.

HOSMER, Jr. D.W.; LEMESHOW, S. **Applied Logistic Regression**. New York: John Wiley, 1989.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Ministério do Planejamento, Orçamento e Gestão. Censo Demográfico – 2000. Disponível em: <http://www.ibge.gov.br/home/estatistica/populacao/censo2000/fecundidade_mortalidade/conceitos.shtm>. Acesso em: 10 jun. 2009.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Mortalidade infantil cai para 25,1 por mil em 2006. Ministério do Planejamento, Orçamento e Gestão. Disponível em: <http://ultimosegundo.ig.com.br/brasil/2007/09/28/ibge_mortalidade_infantil_cai_para_251_por_mil_em_2006_1024251.html>. Acesso em: 23 jun. 2009.

JEKEL, J. F.; ELMORE, J. G.; KATZ, D. L. **Epidemiologia, Bioestatística e Medicina Preventiva**. Tradução de Ricardo Savaris. Porto Alegre: Artmed, 1999.

KAHN, H. A.; SEMPOS, C. T. **Statistical Methods in Epidemiology**. Edited by Brian MacMahon. New York: Oxford University Press, 1989.

KOZU, K. T. et al. **Mortalidade Infantil**: causas e fatores de risco: um estudo bibliográfico, 2005. Disponível em:
<<http://www.medstudents.com.br/original/original/mortinf/mortinf.htm>>. Acesso em: 12 jul. 2009.

MACHADO, C. J.; HILL, K. Maternal, Neonatal and Community Factors Influencing Neonatal Mortality in Brazil. **Journal of Biosocial Science**. v. 37, n. 2, p.193-208, 2005.

MARGOTTO P. R. **Recém-nascido pré-termo extremo**: limite de viabilidade. 2001. Disponível em:
<<http://www.paulomargotto.com.br/especialidade/neumatologia/LIMVIA.doc>> Acesso em: 02 jul. 2009.

MARTINS, E, F.; VELÁSQUEZ-MELÉNDEZ, G. Determinantes da mortalidade neonatal a partir de uma coorte de nascidos vivos, Montes Claros, Minas Gerais, 1997-1999. **Revista Brasileira de Saúde Materno Infantil**, Recife, v. 4, n. 4, p. 405-412, out./dez., 2004.

MEDRONHO, R. A. et al. **Epidemiologia**. São Paulo: Atheneu, 2003.

MINISTÉRIO DA SAÚDE. Secretaria de Vigilância em Saúde (MS/SVSA, 2002). Disponível em: < http://portal.saude.gov.br/portal/saude/Gestor/area.cfm?id_area=1133>. Acesso em: 25 jun. 2009.

MINISTÉRIO DA SAÚDE (MS/SVSB, 2002). Sistema de Informações sobre Mortalidade – SIM. Disponível em:
<http://portal.saude.gov.br/portal/saude/Gestor/visualizar_texto.cfm?idtxt=21377>. Acesso em: 03 jun. 2009.

MINISTÉRIO DA SAÚDE (MS/SVSc, 2002). Sistema de Informações sobre Nascidos Vivos – SINASC. Disponível em:
<http://portal.saude.gov.br/saude/visualizar_texto.cfm?idtxt=21379>. Acesso em: 03 jun. 2009.

MINISTÉRIO DA SAÚDE (MS/SVSD, 2006). Sistema de Informações sobre Nascidos Vivos – SINASC. Disponível em: <<http://tabnet.datasus.gov.br/cgi/sinasc/nvdescr.htm#origem>>. Acesso em: 03 jun. 2009.

MINISTÉRIO DA SAÚDE (MS/SVSe, 2006). Mortalidade Infantil. Informações em saúde. Brasil, 2006. Disponível em: <<http://tabnet.datasus.gov.br/cgi/mortinf/mibr.htm>>. Acesso em: 06 jun. 2009.

MORAES, A. B. **Baixo peso de nascidos vivos no Rio Grande do Sul, Brasil**: uma análise estatística multinível. 2003. p.168. Tese (Doutorado em Epidemiologia) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

MORAIS NETO, O. L.; BARROS, M. B. A. Fatores de risco para a mortalidade neonatal e pós-neonatal na região centro-oeste do Brasil: linkage entre banco de dados de nascidos vivos e óbitos infantis. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 16, n. 2, 2000.

NASCIMENTO, E. M. R. et al. Estudo de fatores de risco para óbitos de menores de um ano mediante compartilhamento de bancos de dados. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 24, n. 11, p. 2593-2602, nov., 2008.

ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE (OPAS/OMS, 2002). Indicadores básicos de saúde no Brasil: Conceitos e aplicações. Rede Interagencial de Informações de Saúde – Ripsa – Brasília, 2002. Disponível em: <<http://www.opas.org.br/sistema/arquivos/matriz.pdf>>. Acesso em: 04 de jul. 2009

PAGANO, M. **Princípios de Bioestatística**. Tradução de Luiz Sérgio de Castro Paiva. São Paulo: Pioneira Thomson Learning, 2004.

PENHA, R. N. **Um estudo sobre regressão logística binária**. 2002. Monografia (Graduação em Engenharia de Produção). Departamento de Produção - Universidade Federal de Itajubá. Itajubá, MG, Brasil, 2002. Disponível em: <<http://www.epr.unifei.edu.br/PFG/producao2002/trabalhos/Renata.PDF>> Acesso em: 10 jun. 2009.

SARINHO, S. W. et al. Fatores de risco para óbitos neonatais no Recife: um estudo caso-controle. **Jornal de Pediatria**, Rio de Janeiro, v. 77, n. 4, 2001.

SECRETARIA ESTADUAL DE SAÚDE - RS (SES/RS, 2003). Mortalidade Infantil. Tabela detalhada de alguns indicadores de saúde por município e CRS 2003. Disponível em: <<http://www.saude.rs.gov.br>>. Acesso em: 04 jun. 2009.

SECRETARIA ESTADUAL DE SAÚDE - RS (SES/RS, 2006). Disponível em:
< <http://www.saude.rs.gov.br/wsa/portal/index.jsp?menu=organograma&cod=1340> >.
Acesso em: 04 jul. 2009

SECRETARIA DO PLANEJAMENTO E GESTÃO DO RS (SEPLAG/RS, 2009).
Mortalidade Infantil, Atlas Socioeconômico do Rio Grande do Sul. Disponível em:
<<http://www.seplag.rs.gov.br/atlas/atlas.asp?menu=312>>. Acesso em 03 jul. 2009

SILVA, C. F. et al. Fatores de risco para mortalidade infantil em município do Nordeste do Brasil: *linkage* entre bancos de dados de nascidos vivos e óbitos infantis, 2000 a 2002. **Revista Brasileira de Epidemiologia**, São Paulo, v. 9, n. 1, p. 69-80, 2006.

SIMÕES, C. C. S. **Estimativas da Mortalidade Infantil por microrregiões e municípios**. Brasília: Ministério da Saúde, 1999.

SOUZA L. M. **Avaliação do Sistema de Informação sobre Nascidos Vivos – SINASC: Minas Gerais e mesorregiões 2000**. Dissertação (Mestrado em Demografia do Centro de Desenvolvimento e Planejamento Regional da FACE) – Universidade Federal de Minas Gerais, 2004.

SZWARCWALD, C. L. et al. Estimação da mortalidade infantil no Brasil: o que dizem as informações sobre óbitos e nascimentos do Ministério da Saúde? **Cadernos de Saúde Pública**, Rio de Janeiro, v. 18, n. 6, dez. 2002.

VICTORA, C. G.; GRASSI, P. R.; SCHMIDT, A. M. Situação de saúde da criança em área da região sul do Brasil, 1980-1992: tendências temporais e distribuição espacial. **Revista de Saúde Pública**, São Paulo, v. 28, n. 6, dez. 1994.

VICTORA, C. G. Intervenções para reduzir a mortalidade infantil pré-escolar e materna no Brasil. **Revista Brasileira de Epidemiologia**, São Paulo, v. 4, n. 1, abr. 2001.

VICTORA, C. G et al. Mortalidade infantil em duas coortes de base populacional no sul do Brasil, 1982 e 1993: tendências e diferenciais. **Cadernos de Saúde Pública**, Rio de Janeiro, n. 12, p.79-86, 1996.

ZANINI, R. R. **Modelos multiníveis aplicados ao estudo da mortalidade infantil no Rio Grande do Sul, Brasil, de 1994 a 2004**. 2007. 200f. Tese (Doutorado em Epidemiologia) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

ZANINI, R. R. et al. Infant mortality trends in the State of Rio Grande do Sul, Brazil, 1994-2004: a multilevel analysis of individual and community risk factors. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 25, n. 5, may. 2009.

WARREN, R. E. Predictive modeling of archaeological site location: a primer. In: ALLEN, K. M. S., GREEN, S. W., ZUBROW, E. B. W., (eds). **Interpreting Space: GIS and Archaeology**. London: Taylor e Francis, 1990.

ANEXOS

ANEXO A - Declaração de Nascido Vivo

Estrutura de Dados SINASC

ESTRUTURA DO SINASC PARA O CD-ROM

Os arquivos são DBF e estão compactados na forma de DBC. Para descompactá-los, ou expandi-los utilize o TABWIN, opção Arquivos Comprime/Expande arquivos DBF.

Os campos dos arquivos são os seguintes:

CAMPO	NOME	TIPO/TAM	DESCRIÇÃO
01	NúmeroDN	C(08)	Número da DN, seqüencial por UF informante e por ano
02	LOCNASC	C(01)	Local de ocorrência do nascimento, conforme a tabela: 0: Ignorado 1: Hospital 2: Outro Estab Saude 3: Domicílio 4: Outro
03	CODESTAB	C(09)	Código de estabelecimento
04	CODBAINASC	C(03)	Código Bairro nascimento
05	CODMUNNASC	C(07)	Município de ocorrência, em codificação idêntica a de CODMUNRES, conforme tabela TABMUN.
06	IDADEMAE	C(02)	Idade da mãe em anos.
07	ESTCIVMAE	C(01)	Estado civil, conforme a tabela: 1: Solteiro 2: Casado 3: Viuvo 4: Separado judicialmente 9: Ignorado
08	ESCMAE	C(01)	Escolaridade, anos de estudo concluídos: 1: Nenhum 2: 1 a 3 anos 3: 4 a 7 anos 4: 8 a 11 anos 5: 12 e mais 9: Ignorado
09	CODOCUPMAE	C(05)	Ocupação, conforme a Classificação Brasileira de Ocupações (CBO)
10	QTDFILVIVO	C(02)	Número de filhos vivos.
11	QTDFILMORT	C(02)	Número de filhos mortos, ignorados, não incluindo o próprio.
12	CODMUNRES	C(07)	Município de residência, em codificação idêntica a de CODMUNOCOR, conforme tabela TABMUN

CAMPO	NOME	TIPO/TAM	DESCRIÇÃO
13	GESTACAO	C(01)	Semanas de gestação, conforme as tabelas: 9: Ignorado 1: Menos de 22 semanas 2: 22 a 27 semanas 3: 28 a 31 semanas 4: 32 a 36 semanas 5: 37 a 41 semanas 6: 42 semanas e mais
14	GRAVIDEZ	C(01)	Tipo de gravidez, conforme a tabela: 9: Ignorado 1: Única 2: Dupla 3: Tripla e mais
15	PARTO	C(01)	Tipo de parto, conforme a tabela: 9: Ignorado 1: Vaginal 2: Cesáreo
16	CONSULTAS	C(01)	Número de consultas de pré-natal: 1: Nenhuma 2: de 1 a 3 3: de 4 a 6 4: 7 e mais 9: Ignorado
17	DTNASC	C(08)	Data do nascimento, no formato ddmmaaa
18	SEXO	C(01)	Sexo, conforme a tabela: 0: Ignorado, não informado 1: Masculino 2: Feminino
19	APGAR 1	C(02)	Apgar no primeiro minuto 00 a 10
20	APGAR 5	C(02)	Apgar no quinto minuto 00 a 10
21	RACACOR	C(01)	Raça/Cor: 1: Branca 2: Preta 3: Amarela 4: Parda 5: Indígena
22	PESO	C(04)	Peso ao nascer, em gramas.
23	IDANOMAL	C(01)	1: Com anomalia 2: Sem Anomalia
24	CODANOMAL	(C04)	Código de malformação congênita ou anomalia cromossômica, de acordo com a CID-10