

Bernardo José Moura Fonseca dos Santos

**Modelagem do baixo peso ao nascer para bebês
nascidos no Estado do Rio de Janeiro, em 2019**

Niterói - RJ, Brasil

29 de abril de 2021



Universidade Federal Fluminense

Bernardo José Moura Fonseca dos Santos

Modelagem do baixo peso ao nascer para bebês nascidos no Estado do Rio de Janeiro, em 2019

Trabalho de Conclusão de Curso

Projeto Final II apresentado para obtenção do grau de Bacharel em Estatística pela Universidade Federal Fluminense

Orientador: Prof. José Rodrigo de Moraes

Niterói - RJ, Brasil

29 de abril de 2021



Universidade Federal Fluminense

Bernardo José Moura Fonseca dos Santos

Modelagem do baixo peso ao nascer para bebês nascidos no Estado do Rio de Janeiro,
em 2019

Projeto Final II de graduação sob o título “*Modelagem do baixo peso ao nascer para bebês nascidos no Estado do Rio de Janeiro, em 2019*”, defendido por Bernardo José Moura Fonseca dos Santos em 29 de abril de 2021, na cidade de Niterói, no Estado do Rio de Janeiro, pela banca examinadora constituída pelos professores:

Prof. Dr. José Rodrigo de Moraes

Orientador

Departamento de Estatística – UFF

Prof. Me. Mauricio Franca Lila

Instituto Brasileiro de Geografia e Estatística – IBGE

Profa. Me. Rafele Febrone Meloni

Instituto de Nutrição – UERJ

Niterói, 29 de abril de 2021

Ficha catalográfica automática - SDC/BIME
Gerada com informações fornecidas pelo autor

S237m Santos, Bernardo José Moura Fonseca dos
Modelagem do baixo peso ao nascer para bebês nascidos no
Estado do Rio de Janeiro, em 2019 / Bernardo José Moura
Fonseca dos Santos ; José Rodrigo de Moraes, orientador.
Niterói, 2021.
51 f.

Trabalho de Conclusão de Curso (Graduação em
Estatística)-Universidade Federal Fluminense, Instituto de
Matemática e Estatística, Niterói, 2021.

1. Modelo Log-linear de Poisson. 2. Baixo Peso ao Nascer. 3.
Sistemas de Informação em Saúde. 4. Estudos Transversais.
5. Produção intelectual. I. Moraes, José Rodrigo de,
orientador. II. Universidade Federal Fluminense. Instituto de
Matemática e Estatística. III. Título.

CDD -

Resumo

O baixo peso ao nascer (BPN) está associado à ocorrência de morbimortalidade neonatal, assim como ao risco de mortalidade durante o primeiro ano de vida e a problemas de desenvolvimento na infância. É considerado um importante indicador de saúde pública relacionado a fatores como saúde materna, nutrição, assistência à saúde e pobreza. Usando os dados do Sistema de Informações sobre Nascidos Vivos (SINASC) de 2019, este trabalho teve como objetivo avaliar a associação entre as características sociodemográficas maternas, da gestação e do parto e a prevalência de baixo peso ao nascer, no Estado do Rio de Janeiro, em 2019. Utilizou-se o modelo de regressão log-linear de Poisson (com estimador de variância robusta), pelo método de máxima verossimilhança (MV). A prevalência de BPN foi cerca de 9,5%; e observou-se maior prevalência de BPN entre os bebês de mães com 35 anos ou mais (RP=1,104, p-valor<0,001), que vivem sem companheiro (RP=1,045, p-valor=0,003), menos escolarizadas e de cor não branca (RP=1,079, p-valor<0,001). Presença de nuliparidade, prematuridade (RP=11,812, p-valor<0,001) e realização de parto cesáreo (RP=1,265, p-valor<0,001) e baixo número de consultas de acompanhamento pré-natal também estiveram associadas à maior prevalência de BPN, assim como o recém-nascido ser do sexo feminino (RP=1,240, p-valor<0,001) e apresentar índice de Apgar baixo (RP=1,550, p-valor<0,001) ou muito baixo (RP=1,620, p-valor<0,001). Conclui-se que características sociodemográficas da mãe, do recém-nascido, bem como da gestação e do parto, estão significativamente associadas a uma maior prevalência de BPN. Destaca-se que é necessário um olhar mais atento por parte dos profissionais de saúde e dos formuladores de políticas públicas às gestantes com 35 anos ou mais, baixa escolaridade, sem companheiro, de raça/cor não branca, com pouca frequência nas consultas pré-natal e nulíparas, para que recebam apoio a fim de reduzir o número de recém-nascidos com baixo peso ao nascer.

Palavras-chaves: Modelo Log-linear de Poisson, Baixo Peso ao Nascer, Prevalência, Sistemas de Informação em Saúde, Estudos Transversais

Agradecimentos

Gostaria de agradecer à toda minha família e principalmente à minha namorada Luana por todo o apoio durante minha trajetória na graduação.

Ao meu orientador, José Rodrigo de Moraes, por todo o ensinamento, paciência e compreensão no decorrer deste trabalho.

Aos membros da banca Mauricio Franca Lila e Rafael Febrone Meloni, por terem aceitado fazer parte deste trabalho e por todas as sugestões que possibilitaram melhorar ainda mais.

Por fim, gostaria de agradecer aos amigos que fiz durante a graduação. Sem vocês, essa caminhada teria sido muito mais difícil.

Sumário

Resumo.....	5
Agradecimentos	6
Lista de Figuras.....	8
Lista de Quadros.....	9
Lista de Tabelas	10
1 Introdução.....	11
1.1 Objetivos	12
2 Material e Métodos.....	13
2.1 Sistema de Informações sobre Nascidos Vivos (SINASC).....	13
2.2 Instrumento de coleta de dados	14
2.3 População de estudo	15
2.4 Variáveis de estudo	15
2.5 Modelos Lineares Generalizados (MLG).....	18
2.5.1 Componentes de um MLG	18
2.5.2 Modelo log-linear de Poisson (com variância robusta).....	19
2.5.3 Estimador de variância robusta	21
2.5.4 Método de Máxima Verossimilhança (MV)	23
2.5.5 Inferência sobre os parâmetros do modelo.....	25
2.5.6 Teste de Wald de significância individual	25
2.5.7 Teste de Wald de significância geral.....	26
2.5.8 Teste de Razão de Verossimilhanças	27
2.5.9 Medidas para avaliação da qualidade do ajuste do modelo.....	29
3 Resultados	33
4 Discussão e Conclusão.....	42
Anexos.....	45
Referências.....	46

Lista de Figuras

Figura 1: Curva ROC (adaptada de HOSMER & LEMESHOW, 2000).....	31
Figura 2: Curva ROC do modelo selecionado	41

Lista de Quadros

Quadro 1: Características maternas e do bebê	16
Quadro 2: Dados numa tabela de contingência 2x2 (análise bivariada)	23
Quadro 3: Classificação das unidades da amostra segundo as categorias observadas (linha) e as categorias preditivas (coluna) da variável resposta.	30

Lista de Tabelas

Tabela 1: Distribuição (absoluta e percentual) dos recém-nascidos com ou sem baixo peso ao nascer segundo as características sociodemográficas maternas, da gestação e do parto e do recém-nascido.	33
Tabela 2: Resultados do ajuste do modelo log-linear de Poisson (com variância robusta) considerando as características sociodemográficas maternas, da gestação e do parto, e do recém-nascido, com as suas respectivas categorias, previamente listadas no Quadro 1 (modelo original).	35
Tabela 3: Resultados do ajuste do modelo log-linear de Poisson (com variância robusta) considerando as características sociodemográficas maternas, da gestação e do parto, e do recém-nascido, agregando as categorias não significativas (modelo selecionado).	37
Tabela 4: Medidas para avaliação da qualidade do ajuste do modelo log-linear de Poisson (original versus selecionado).	40

1 Introdução

De acordo com a Organização Mundial da Saúde (OMS), o baixo peso ao nascer (BPN) compreende a faixa de peso ao nascer inferior a 2500g (WHO, 2014). Em termos mundiais, estima-se que de 15 a 20% dos recém-nascidos apresentam BPN, o que representa mais de 20 milhões de nascimentos por ano. Desse modo, o BPN pode ser considerado mundialmente como um importante problema de saúde pública com consequências a curto e longo prazo (WHO, 2014).

A avaliação do estado de saúde do recém-nascido envolve o cálculo de medidas antropométricas, entre elas pode citar o peso, o comprimento e perímetro cefálico. O peso ao nascer é a primeira medida de peso do recém-nascido calculada, preferencialmente nas primeiras horas de vida. Contudo, algumas localidades podem não apresentar estruturas hospitalares adequadas, impossibilitando que o peso seja medido nas primeiras horas de vida. Tal fato pode acarretar uma distorção na proporção de recém-nascidos com BPN, pois segundo Cutland et al.(2017) recém nascidos a termo perdem de 3,5 a 6,6% do seu peso nos primeiros 2,5-2,7 dias após o nascimento, sendo que recém-nascidos amamentados exclusivamente com leite materno perdem mais peso e levam mais tempo para recuperá-lo, comparativamente aos recém-nascidos que consomem fórmulas ou misturas.

O BPN está associado à ocorrência de morbimortalidade neonatal (FERRAZ e NEVES, 2011), e é um importante indicador de saúde pública relacionado a fatores como saúde materna, nutrição, assistência à saúde e pobreza (CUTLAND et al., 2017). O risco de mortalidade de neonatos é 20 vezes maior que o de neonatos com peso ao nascer superior a 2500g (CUTLAND et al., 2017)

O BPN também está associado ao risco de mortalidade durante o primeiro ano e a problemas de desenvolvimento na infância (TIAN et al., 2019; WILCOX, 2001). Zerbeto et al. (2015) afirmam que quanto menor o peso ao nascer, maiores são os riscos de problemas no desenvolvimento da criança, podendo apresentar deficiências neurológicas, sensoriais e linguísticas. Além disso, o BPN está relacionado ao risco de aparecimento de doenças na fase adulta (TIAN et al., 20019; WILCOX, 2001), tais como doenças como diabetes, coronariopatias, hipertensão arterial e obesidade (BISMARCK-NASR et al.,2008).

A OMS tem como meta reduzir, até o ano de 2025, em 30% o número de recém-nascidos com BPN. Isso representa uma redução de 3% por ano até 2025 e de 6 milhões de recém-nascidos no total (WHO, 2014).

Segundo Cutland et al. (2017), o BPN é resultado de uma gestação pré-termo (duração inferior a 37 semanas completas), uma gestação com desenvolvimento intrauterino, ou ambos os casos. Porém as causas do BPN são multifatoriais e suas causas exatas são ainda desconhecidas.

Em 2019, a prevalência de BPN no Brasil foi de 8,70%. Neste ano, observaram-se diferenças na distribuição de BPN entre as grandes regiões brasileiras. A região Norte apresentou o menor percentual de recém-nascidos com BPN (7,74%), enquanto o Sudeste apresentou o maior percentual (9,33%), seguido das regiões Sul (8,86%), Centro-Oeste (8,53%) e Nordeste (8,19%). No Estado do Rio de Janeiro, a prevalência de BPN foi de 9,28%. (SECRETARIA DE VIGILANCIA EM SAÚDE, 2021).

1.1 Objetivos

Este trabalho tem como objetivo avaliar a associação das características sociodemográficas maternas, da gestação e do parto, e do recém-nascido com o baixo peso a nascer, no Estado do Rio de Janeiro, em 2019.

Entre os objetivos específicos, pode-se citar:

- Analisar a distribuição de recém-nascidos por desfecho de estudo (ter BPN *versus* não ter BPN), segundo as características sociodemográficas maternas, da gestação e do parto e do recém-nascido;
- Identificar e analisar as características sociodemográficas maternas, da gestação e do parto e do recém-nascido que estão significativamente associadas com a prevalência de BPN;
- Avaliar a qualidade do ajuste do modelo e sua capacidade preditiva.

2 Material e Métodos

2.1 Sistema de Informações sobre Nascidos Vivos (SINASC)

No Brasil, as características da mãe, da gestação, do parto e do recém-nascido podem ser avaliadas com base nos dados do Sistema de Informações sobre Nascidos Vivos (SINASC). O SINASC, criado em 1990, é um dos sistemas de informações de saúde, de abrangência nacional, desenvolvidos pelo Departamento de Informática do Sistema Único de Saúde (DATASUS), do Ministério da Saúde, em 1990, com o objetivo de coletar informações sobre nascimentos com vida ocorridos no território brasileiro, a fim de subsidiar a formulação, a implementação, o monitoramento e avaliação de políticas públicas (SZWARCOWALD et al., 2019), relacionadas a saúde do recém-nascido e da gestante para as três esferas de gestão do Sistema Único de Saúde (municipal, estadual e federal) (MS, 2019).

A gestão do SINASC é realizada, a partir de 2003, pela Secretaria de Vigilância em Saúde (SVS), do Ministério da Saúde, juntamente com as Secretarias Estaduais e Municipais de Saúde. As Secretarias Municipais de Saúde são responsáveis pela coleta das Declarações de Nascidos Vivos (DN), bem como pelo encaminhamento mensal das informações registradas nas DN para as Secretarias Estaduais, que, por sua vez, ficam encarregadas de enviar tais informações para a Secretaria de Vigilância em Saúde (MS, 2019; PEDRAZA, 2012; MS, 2004).

O SINASC abrange atualmente todas as Unidades da Federação do país. Szwarcwald et al. (2019) analisaram a cobertura das informações de nascidos vivos por UF provenientes do SINASC, e verificaram que em 2013 a cobertura foi alta e homogênea em todas as Unidades da Federação. Em particular, observou-se no Estado do Rio de Janeiro um percentual de cobertura de 99,7%, superior à cobertura registrada em nível nacional (94,8%).

A base de dados do SINASC é de uso público e compila um número significativo e relevante de variáveis relativas aos nascidos vivos (sexo, peso ao nascer, índice de Apgar, local de nascimento, etc), à gravidez/parto (idade gestacional, tipo de parto, número de consultas pré-natais, etc) e à mãe (idade, raça/cor, situação conjugal, nível de instrução, município de residência, etc). A partir dos dados do SINASC é possível avaliar a situação de saúde materno-infantil e as condições de assistência ao pré-natal e ao parto (STEVANATO et al., 2017), visando reduzir os riscos de morbidade e mortalidade para a gestante e para a criança (MS, 2018). Além disso, é possível calcular indicadores para

conhecer o perfil demográfico e epidemiológico dos nascidos vivos, com desagregabilidade municipal, estadual e nacional, a fim de subsidiar a formulação e avaliação de políticas voltadas para a promoção de saúde materno-infantil (SILVESTRIN et al., 2018).

2.2 Instrumento de coleta de dados

O SINASC utiliza a Declaração de Nascido Vivo (DN) como instrumento de coleta de dados. A DN é um documento padrão oficial de validade nacional preenchido por profissionais de saúde nos estabelecimentos de saúde públicos e privados onde os partos são realizados, ou por profissionais dos Cartórios de Registro Civil, no caso de partos ocorridos em unidades domiciliares (PEDRAZA, 2012).

O formulário de DN (Anexo 1) é dividido em oito partes referentes ao: 1) Recém-nascido; 2) Local da ocorrência; 3) Mãe; 4) Pai; 5) Gestação e parto; 6) Anomalia congênita; 7) Responsável pelo preenchimento; 8) Cartório. (SECRETARIA DE VIGILANCIA EM SAÚDE, 2021).

A parte do recém-nascido (Parte 1) refere-se às variáveis do recém-nascido, como nome, data/hora do nascimento, sexo e peso ao nascer (em gramas) e o índice de Apgar, medido no primeiro e no quinto minuto de vida. Com relação ao peso ao nascer, a indicação no formulário da DN é que o peso do recém-nascido seja mensurado nas primeiras 5 horas de vida.

Segundo Schardosim et al (2018) o índice de Apgar foi proposto por Virginia Apgar e tem como objetivo avaliar as condições fisiológicas e de resposta dos neonatos. O índice de Apgar leva em consideração os seguintes parâmetros: frequência cardíaca, coloração da pele, esforço respiratório, tônus muscular e irritabilidade reflexa. Cada um desses parâmetros recebe uma pontuação entre 0 e 2 e o somatório será o índice de Apgar, variando de 0 a 10. Em 90% dos neonatos a pontuação desse índice está entre 7 e 10, que é considerada a faixa adequada pelos especialistas. Uma pontuação entre 4 e 6 é considerada como anóxia moderada e entre 0 e 3 anóxia grave.

A parte referente ao local da ocorrência (Parte 2) é a segunda parte do formulário da DN, que consta os quesitos relacionados ao local da ocorrência do parto (hospital ou outro estabelecimento de saúde, domicílio e outro local), além de informações sobre o bairro e município de ocorrência.

A parte referente à mãe (Parte 3) aborda basicamente as características sociodemográficas da mãe, como idade, sexo, naturalidade, situação conjugal, cor/raça e escolaridade materna.

A parte referente ao pai (Parte 4) aborda somente o nome e idade do pai do recém-nascido.

A parte referente à gestação e parto (Parte 5) aborda informações referentes a gestações anteriores: número de gestações anteriores, número de partos vaginais, número de cesáreas, número de nascidos vivos e número de perdas fetais/abortos. Também são abordadas informações sobre a gestação atual como o número de consultas pré-natal, tipo de gravidez e tipo de parto.

A parte sobre a Anomalia Congênita (Parte 6) é onde são descritas todas as anomalias congênitas que o recém-nascido apresentou. Caso não tenha apresentado nenhuma anomalia congênita, esta parte deve ficar em branco.

As duas últimas partes, denominadas preenchimento (Parte 7) e cartório (Parte 8), apresentam informações referentes ao funcionário que preencheu a Declaração de Nascido Vivo e do cartório onde a mesma foi registrada.

2.3 População de estudo

A população de estudo foi composta por todos os recém-nascidos sem anomalias congênitas, cujas mães possuem 15 anos ou mais de idade, com tipo de gravidez única (não múltipla) e com local de nascimento em hospital ou outro local. Foram excluídos da análise recém-nascidos sem informação sobre peso ao nascer, ou com valor de peso ao nascer extremo (inferior a 500 gramas ou superior a 6.000 gramas).

2.4 Variáveis de estudo

O Quadro 1 apresenta as características da mãe e do recém-nascido obtidas no formulário da Declaração de Nascido Vivo (Quadro 1).

Quadro 1: Características maternas e do bebê

Variáveis de estudo	Categorias
Características sociodemográficas maternas	
Faixa etária materna	15 a 19 anos 20 a 34 anos 35 anos ou mais
Situação conjugal da mãe	Sem Companheiro (solteira, viúva ou separada judicialmente/divorciada) Com Companheiro (casada/morando junto)
Escolaridade da Mãe	Sem instrução ou até 3 anos de estudo De 4 a 7 anos de estudo De 8 a 11 anos de estudo 12 anos de estudo ou mais
Características da gestação e do parto	
Paridade (Número total de filhos anteriores vivos ou mortos)	Nulípara De 1 a 3 filhos 4 filhos ou mais
Idade gestacional	Pré termo (menos que 37 semanas de gestação) A termo (de 37 a 41 semanas de gestação) Pós termo (42 semanas ou mais de gestação)
Tipo de Parto	Vaginal Cesáreo

Quadro 1: Características maternas e do bebê (**continuação**)

Variáveis de estudo	Categorias
Características da gestação e do parto	
Número de consultas pré-natal	Nenhuma De 1 a 3 consultas De 4 a 6 consultas 7 consultas ou mais
Mesorregião de nascimento do bebê	Baixada Centro Fluminense Metropolitana Noroeste Fluminense Norte Fluminense Sul Fluminense
Características do recém-nascido	
Sexo do bebê	Masculino Feminino
Raça/Cor da mãe	Branca Não branca
Índice de Apgar no quinto minuto	De 0 a 3 (muito baixo) De 4 a 6 (baixo) Escore de 7 a 10 (normalidade)
Baixo peso ao nascer (desfecho de estudo)	Sim Não

O desfecho de BPN foi obtido a partir da variável quantitativa “peso ao nascer”, medido em gramas, contida na parte do formulário referente ao recém-nascido. Trata-se de um desfecho binário, onde na categoria “sim” são considerados todos os recém-nascidos com baixo peso ao nascer (BPN), isto é, com peso inferior a 2500 gramas. Na categoria “não”, por sua vez, são considerados os recém-nascidos que não apresentaram baixo peso no momento do nascimento.

Com relação a variável cor/raça, optou-se por dicotomizá-la em branca versus não branca, assim como estudos de Moraes et al. (2011), Moreira et al. (2011) e Vale (2020).

2.5 Modelos Lineares Generalizados (MLG)

Modelos Lineares Generalizados (MLG) são uma classe de modelos estatísticos oriundos da regressão linear clássica que fornecem uma maneira sistemática de lidar com variáveis respostas categóricas (POWERS e XIE, 1999). Os modelos lineares padrões não conseguem lidar com respostas não-normais, como contagens ou proporções, o que motivou o desenvolvimento dos MLG, a fim de modelar variáveis respostas categóricas e binárias, entre outros tipos (FARAWAY, 2006).

2.5.1 Componentes de um MLG

Segundo Rencher e Schaalje (2008), os MLGs podem ser caracterizados por três componentes: um conjunto de n variáveis aleatórias independentes (componente aleatória), um preditor linear (componente determinística) e uma função de ligação. Essas componentes são descritas a seguir:

1. Sejam y_1, y_2, \dots, y_n variáveis aleatórias independentes com $E(y_i) = \mu_i \quad \forall i = 1, 2, \dots, n$ e função de probabilidade (caso discreto) ou função de densidade de probabilidade (caso contínuo) pertencente à família exponencial. Por exemplo, no caso discreto, uma função de probabilidade $P(y_i, \theta_i)$ pertence à família exponencial se puder ser escrita na forma $P(y_i, \theta_i) = e^{[a(y_i) \cdot b(\theta_i) + c(\theta_i) + d(y_i)]}$, onde a , b , c e d são funções conhecidas. A partir destas funções, pode-se conhecer a média e a variância da variável aleatória $a(y_i)$, por meio das seguintes expressões:

$$E[a(y_i)] = \frac{-c'(\theta_i)}{b'(\theta_i)}$$
$$VAR[a(y_i)] = \frac{b''(\theta_i)c'(\theta) - c''(\theta_i)b'(\theta_i)}{[b'(\theta_i)]^3}$$

2. Um preditor linear dado por: $\mathbf{x}'_i \boldsymbol{\beta} = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{i,k-1}$, onde $\boldsymbol{\beta}$ é o vetor, de dimensão $k \times 1$, dos parâmetros desconhecidos do modelo e \mathbf{x}'_i é o vetor das variáveis explicativas (covariáveis ou variáveis dummy para os níveis dos fatores), de dimensão $k \times 1$, referente a i -ésima unidade.

3. Uma função de ligação não-linear, monótona e diferenciável, que descreve como a média da variável dependente (ou desfecho) y_i , denotada por $E(y_i) = \mu_i$, se relaciona com o preditor linear $\mathbf{x}'_i \boldsymbol{\beta}$, isto é: $g(\mu_i) = \mathbf{x}'_i \boldsymbol{\beta}$.

2.5.2 Modelo log-linear de Poisson (com variância robusta)

O modelo de regressão log-linear de Poisson é um MLG, onde a distribuição de probabilidade da variável resposta y_i é a distribuição de Poisson e a função de ligação $g(\mu_i)$ é a função logarítmica (FARAWAY, 2006). A função de probabilidade da variável aleatória de Poisson é dada por:

$$P(y_i, \mu_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}; \quad y_i = 0, 1, 2, \dots$$

Como demonstrado abaixo, a função de probabilidade da variável aleatória de Poisson pode ser expressa na forma da família exponencial-

$$\begin{aligned} P(y_i, \mu_i) &= \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!} = \exp \left[\ln e^{-\mu_i} + \ln \mu_i^{y_i} - \ln(y_i!) \right] = \\ &= \exp \left[-\mu_i + y_i \ln \mu_i - \ln(y_i!) \right] \end{aligned}$$

onde: $a(y_i) = y_i$, $b(\mu_i) = \ln \mu_i$, $c(\mu_i) = -\mu_i$ e $d(y_i) = -\ln(y_i!)$

Com relação a sua especificação, o modelo de regressão log-linear de Poisson é dado por:

$$\ln(\mu_i) = \mathbf{x}'_i \boldsymbol{\beta} \quad ; \quad i = 1, 2, \dots, n \quad (1.1)$$

onde:

$\mathbf{x}'_i = [1 \quad x_{i1} \quad x_{i2} \quad \dots \quad x_{i,k-1}]$ é o vetor de variáveis explicativas, de dimensão $1 \times k$, referente a i -ésima unidade.

$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{k-1} \end{bmatrix}$ é o vetor de parâmetros desconhecidos, de dimensão $k \times 1$.

$\mu_i = E(y_i)$ é a média da variável dependente y_i da i -ésima unidade.

$\ln(\mu_i)$ é o logaritmo da média da variável dependente y_i da i -ésima unidade.

A partir da equação (1.1), pode-se obter a média da variável dependente referente a i -ésima unidade, dada por:

$$\mu_i = e^{x_i'\beta} \quad (1.2)$$

ou, equivalentemente, pode ser expressa como na equação (1.3):

$$\mu_i = e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_{k-1} x_{i,k-1}} \quad (1.3)$$

No caso da variável dependente y_i ser binária, a média μ_i corresponde a prevalência do evento de interesse (sucesso), podendo ser denotada, alternativamente, pela letra p_i , e dada por:

$$p_i = e^{x_i'\beta} \quad (1.4)$$

Tal modelo foi utilizado, no presente estudo, para analisar um desfecho binário (1=sucesso, 0=fracasso), o que possibilita a estimação de razões de prevalências (RP) para cada variável explicativa (variável *dummy* para os níveis dos fatores).

A Razão de Prevalência (RP) é uma medida de associação utilizada em estudos seccionais para determinar a prevalência de sucesso de um grupo (digamos Grupo 1) comparativamente a outro grupo, chamado grupo de referência e denotado, neste trabalho, por Grupo 0. Considerando, por exemplo, a situação em que a variável explicativa x_{ij} é categórica binária (uma variável *dummy*), a razão de prevalência é dada por:

$$RP_j = \frac{p_i(x_{ij} = 1)}{p_i(x_{ij} = 0)} \quad (1.5)$$

$$RP_j = \frac{e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_j(1) + \dots + \beta_{k-1} x_{i,k-1}}}{e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_j(0) + \dots + \beta_{k-1} x_{i,k-1}}} = e^{\beta_j(1) - \beta_j(0)} = e^{\beta_j} \quad (1.5.1)$$

Se $\beta_j > 0$ temos que $RP_j > 1$, e isso significa que a prevalência de sucesso do grupo 1 é $(e^{\beta_j} - 1) * 100\%$ maior do que a do grupo de referência (grupo 0). E caso $\beta_j < 0$ ($RP_j < 1$), a prevalência de sucesso no grupo 1 é $(e^{\beta_j} - 1) * 100\%$ menor do que a do grupo de referência. Se $\beta_j = 0$ ($RP_j = 1$), a prevalência de sucesso do grupo 1 é igual à do grupo de referência.

Quando o modelo possui variáveis categóricas com mais de dois níveis, a mesma expressão dada em (1.5.1) pode ser utilizada, devendo para tanto escolher um dos níveis da variável como referência (base de comparação). Supondo uma variável explicativa categórica com K categorias, a razão de prevalência pode ser definida do seguinte modo:

$$RP_j = e^{\beta_j}, j = 1, 2, \dots, K,$$

onde K indica o grupo de referência, sendo $\beta_K = 0$ e $RP_K = 1$.

2.5.3 Estimador de variância robusta

Na seção 2.5.2 foi apresentado o modelo de regressão log-linear de Poisson, que é um MLG que adota a função de ligação logarítmica e a distribuição de Poisson para o desfecho de estudo. O modelo de regressão de Poisson é geralmente adotado em estudos longitudinais, onde o desfecho representa uma contagem (um número inteiro positivo) num determinado intervalo de tempo, entretanto também pode ser utilizado para modelar desfechos binários ou binomiais (FARAWAY, 2006). Coutinho et al. (2008) mencionam que quando o modelo de regressão de Poisson é aplicado a desfechos binários (ou binomiais) ocorre uma superestimação do erro-padrão da medida de associação¹ uma vez que a variância da distribuição de Poisson cresce progressivamente, enquanto a variância da distribuição Binomial atinge seu maior valor quando a prevalência é igual a 0,5. Para desfechos raros, o erro padrão da distribuição Poisson convergirá para o erro padrão da distribuição binomial (CUMMINGS, 2009). Neste caso, segundo Hirakata (2009), a medida de razão de chances (odds ratio) é numericamente similar ao risco relativo (ou razão de prevalência) e pode ser interpretada de maneira semelhante. Entretanto, no caso de desfechos frequentes (não raros), Hirakata (2009) destaca que a razão de chances superestima o verdadeiro valor do risco relativo. Desse modo, o uso do modelo de regressão de Poisson para estimar razões de prevalências em estudos transversais, no caso de desfechos não raros, acarretará em valores elevados para os erros padrão dos estimadores, p-valores dos testes de significância e intervalos de confiança para as medidas de associação (CUMMINGS, 2009).

Este problema pode ser resolvido ajustando um modelo de regressão de Poisson utilizando um estimador robusto de variância, o que significa que o referido modelo pode ser utilizado mesmo em situações em que os dados não são satisfatoriamente aproximados por uma distribuição de Poisson (ZOU, 2004; CUMMINGS, 2009). O estimador robusto de variância é também chamado de estimador de Huber, White, Huber–White, e de “estimador sanduíche” (CUMMINGS, 2009). O emprego deste estimador permite, portanto, obter estimativas robustas de variâncias (ou erros padrão) dos estimadores dos

¹ Razão de prevalência (RP) para estudos transversais, e Razão de Risco (ou Risco relativo) para estudos de coorte e ensaios clínicos randomizados.

parâmetros do modelo ou das medidas de associação de interesse, bem como possibilita a obtenção de intervalos de confiança e p-valores adequados.

Segundo Zou (2004) e Hiraakata (2009) não há justificativas para utilizar o modelo de regressão logística para desfechos binários quando a medida de associação de interesse entre as variáveis explicativas e o desfecho é o risco relativo, ou a razão de prevalências² em estudos de natureza transversal, sobretudo ao levar em conta que a regressão de Poisson com variância robusta também está implementada em vários pacotes estatísticos padrões. No programa R, o modelo de regressão log-linear de Poisson pode ser ajustado usando a função “glm” (generalized linear models). Para a obtenção de estimativas robustas para as variâncias (ou erros padrão) foi adicionalmente empregado o pacote sandwich do programa R, versão 3.5.1.

Considerando uma única variável explicativa (ou de exposição) e um desfecho, ambas de natureza binária, Zou (2004) apresentou em seu artigo as expressões dos estimadores não corrigido e corrigido da variância para o risco relativo, este último empregado para resolver o problema de superestimação da variância que ocorre quando o modelo de regressão de Poisson é aplicado para desfecho binário.

No caso da distribuição de Poisson a função de log-verossimilhança é dada por

$$\ln L(\alpha, \beta) = C \cdot \sum_{i=1}^n [y_i(\alpha + \beta x_i) - \exp(\alpha + \beta x_i)]$$

Onde C é uma constante. O estimador da variância do risco relativo estimado é dado por:

$$\widehat{VAR}(\widehat{RR}) = \frac{1}{a} + \frac{1}{c}$$

Já a expressão da variância corrigida do risco relativo estimado é dada por:

$$VAR(\widehat{RR}) = \frac{1}{a^2} \sum_{i=1}^{n_i} [y_i - \exp(\alpha + \beta)]^2 + \frac{1}{c^2} \sum_{i=1}^{n_o} [y_i - \exp(\alpha)]^2$$

que é estimada pela seguinte expressão:

$$\widehat{VAR}(\widehat{RR}) = \frac{1}{a} - \frac{1}{n_1} + \frac{1}{c} - \frac{1}{n_0}$$

² Matematicamente, o cálculo da razão de prevalências (RP) em estudos transversais é idêntico ao cálculo do risco relativo em estudos de coorte (McNutt et al., 2003).

onde os valores de a, b, c, n_0 e n_1 são obtidos através das informações de uma tabela de contingência ilustrada no Quadro 3:

Quadro 2: Dados numa tabela de contingência 2x2 (análise bivariada)

Variável de exposição	Desfecho		Total
	$y = 1$ (Sucesso)	$y = 0$ (Fracasso)	
$x = 1$ (Exposto)	a	b	$n_1 = a + b$
$x = 0$ (não exposto)	c	d	$n_0 = c + d$

2.5.4 Método de Máxima Verossimilhança (MV)

De acordo com Rencher e Schaalje (2008), para estimar o parâmetro β pode-se utilizar o método de máxima verossimilhança (MV). Segundo Kutner (2005), o método de MV consiste em escolher como estimativas dos valores dos parâmetros aquelas que são mais consistentes com os dados da amostra. Considere uma amostra aleatória y_1, y_2, \dots, y_n de observações independentes de uma população com função de probabilidade $P(y, p)$, sendo p um parâmetro desconhecido que descreve alguma característica da distribuição de y .

Como y_1, y_2, \dots, y_n são variáveis aleatórias independentes e cada uma segue uma distribuição de Poisson com parâmetro p_i , a função de probabilidade de y_i pode ser expressa por:

$$P(y_i, p_i) = \frac{e^{-p_i} p_i^{y_i}}{y_i!}; \quad y_i = 0, 1, 2, \dots$$

onde: $p_i = p_i(\beta) = e^{x_i' \beta}$.

A função de máxima verossimilhança da amostra, denotada por $L(\beta)$, é obtida do seguinte modo:

$$L(\beta) = \prod_{i=1}^n P(y_i, p_i) = \prod_{i=1}^n \frac{e^{-p_i} p_i^{y_i}}{y_i!}; \quad y_i = 0, 1, 2, \dots \quad (2.1)$$

Aplicando o logaritmo da função de verossimilhança da amostra, obtém-se a seguinte expressão:

$$\begin{aligned} \ln L(\boldsymbol{\beta}) &= \ln \prod_{i=1}^n P(y_i, p_i) = \sum_{i=1}^n \ln P(y_i, p_i) \\ &= \sum_{i=1}^n (-p_i + y_i \ln p_i - \ln y_i!) = \sum_{i=1}^n (-e^{x_i' \boldsymbol{\beta}} + y_i x_i' \boldsymbol{\beta} - \ln y_i!) \end{aligned} \quad (2.2)$$

Para maximizar o logaritmo da função de verossimilhança amostral, basta derivar $\ln L(\boldsymbol{\beta})$ em relação a $\boldsymbol{\beta}$ e igualar a $\mathbf{0}$, obtendo um sistema de equações de verossimilhança:

$$\frac{\partial}{\partial \boldsymbol{\beta}} \ln L(\boldsymbol{\beta}) = \sum_{i=1}^n x_i' [y_i - e^{x_i' \boldsymbol{\beta}}] = \mathbf{0} \quad (2.3)$$

A solução do sistema de equações (2.3), obtida por meio de métodos iterativos, é o estimador $\hat{\boldsymbol{\beta}}_{MV}$ de máxima verossimilhança do vetor $\boldsymbol{\beta}$ de parâmetros do modelo log-linear de Poisson. Segundo Cordeiro e Lima Neto (2004), os programas computacionais sempre utilizam o método de Escore de Fisher para obter as estimativas de MV dos parâmetros de MLGs.

Para grandes amostras, a distribuição de $\hat{\boldsymbol{\beta}}_{MV}$ é assintoticamente normal multivariada (POWERS e XIE, 1999), com média $E(\hat{\boldsymbol{\beta}}_{MV}) = \boldsymbol{\beta}$ e matriz de variância-covariância, denotada por $\widehat{VAR}(\hat{\boldsymbol{\beta}}_{MV})$, sendo esta obtida pelo inverso da matriz de informação de Fisher, isto é:

$$\widehat{VAR}(\hat{\boldsymbol{\beta}}_{MV}) = [\mathfrak{I}(\hat{\boldsymbol{\beta}}_{MV})]^{-1} = \begin{bmatrix} \widehat{VAR}(\hat{\beta}_0) & \cdots & \widehat{COV}(\hat{\beta}_0, \hat{\beta}_{p-1}) \\ \vdots & \ddots & \vdots \\ \widehat{COV}(\hat{\beta}_{p-1}, \hat{\beta}_0) & \cdots & \widehat{VAR}(\hat{\beta}_{p-1}) \end{bmatrix}$$

onde:

$\mathfrak{I}(\hat{\boldsymbol{\beta}}_{MV}) = \frac{\partial^2}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \ln L(\boldsymbol{\beta})$ é uma matriz simétrica, de dimensão $p \times p$, com as segundas derivadas parciais de $\ln L(\boldsymbol{\beta})$ em relação a $\boldsymbol{\beta}$.

A partir da distribuição normal assintótica de $\hat{\boldsymbol{\beta}}_{MV}$, pode-se calcular intervalos de confiança e realizar testes de hipóteses para $\boldsymbol{\beta}$, ou então, para funções de $\boldsymbol{\beta}$, como por exemplo, razões de prevalência (RP) em estudo seccionais (ou transversais).

2.5.5 Inferência sobre os parâmetros do modelo

A inferência sobre os parâmetros do modelo envolve o cálculo de intervalos de confiança e a realização de testes de hipóteses para β , ou ainda, para as razões de prevalência (RP). Através da normalidade assintótica de $\hat{\beta}_{MV}$, pode-se obter um intervalo de confiança (IC) para o parâmetro β_j ($j = 0, 1, \dots, k - 1$) do modelo. com nível de confiança de $100(1 - \alpha)\%$, sendo dado por:

$$IC(\beta_j, 100(1 - \alpha)\%) = \left[\hat{\beta}_j - z_{(1-\frac{\alpha}{2})} \sqrt{VAR(\hat{\beta}_j)}, \hat{\beta}_j + z_{(1-\frac{\alpha}{2})} \sqrt{VAR(\hat{\beta}_j)} \right] \quad (3.1)$$

onde $-z_{(1-\frac{\alpha}{2})}$ e $+z_{(1-\frac{\alpha}{2})}$ são valores (designados por “quantis” ou “valores críticos”) da distribuição normal padrão (ou reduzida) no percentil $(1 - \frac{\alpha}{2})$.

O intervalo de confiança (IC) para a razão de prevalência RP_j ($j = 0, 1, \dots, k - 1$), é obtida aplicando o exponencial dos limites do intervalo de confiança para β_j , como mostrado em (3.2):

$$IC(RP_j, 100(1 - \alpha)\%) = \left[e^{\hat{\beta}_j - z_{(1-\frac{\alpha}{2})} \sqrt{VAR(\hat{\beta}_j)}}, e^{\hat{\beta}_j + z_{(1-\frac{\alpha}{2})} \sqrt{VAR(\hat{\beta}_j)}} \right] \quad (3.2)$$

2.5.6 Teste de Wald de significância individual

O teste de Wald é utilizado para verificar a significância dos parâmetros do modelo. Os parâmetros são testados individualmente sob as seguintes hipóteses:

$$\begin{cases} H_0: \beta_j = 0 \\ H_1: \beta_j \neq 0 \end{cases}$$

onde β_j é o efeito principal do j -ésimo nível da variável explicativa categórica. Já a estatística de teste utilizada, sob a hipótese nula H_0 , é dada por:

$$Z = \frac{\hat{\beta}_j}{\sqrt{VAR(\hat{\beta}_j)}} \sim N(0,1) \quad (3.3)$$

onde $\widehat{\beta}_j$ é o estimador de MV do parâmetro β_j e $\widehat{VAR}(\widehat{\beta}_j)$ é o estimador da variância de $\widehat{\beta}_j$.

A região crítica do teste é dada por:

$$RC = \left\{ z \in \mathbb{R} \mid |z| > z_{\left(1-\frac{\alpha}{2}\right)} \right\}$$

onde: $z_{\left(1-\frac{\alpha}{2}\right)}$ é o valor crítico da distribuição normal padrão no percentil $\left(1 - \frac{\alpha}{2}\right)$.

Por sua vez, o critério de decisão é: Ao nível de significância de $100 * \alpha\%$, rejeita-se H_0 caso o valor observado da estatística de teste (Z) pertença a região crítica (RC), e neste caso conclui-se que o j -ésimo nível da variável explicativa é significativamente diferente de zero. Caso o valor observado não pertença a RC, não se rejeita H_0 ao nível de significância de $100 * \alpha\%$, ou seja, o j -ésimo nível da variável explicativa não é significativamente diferente de zero.

Além da abordagem tradicional de teste de hipóteses, pode-se aplicar a abordagem baseada no p-valor do teste de Wald. O p-valor do teste de Wald de significância individual, dado por $p\text{-valor} = 2 \cdot P(Z > |z|)$, é geralmente incluído nas saídas de programas estatísticos quando se ajustam MLGs. Equivalentemente, rejeita-se H_0 se o p-valor do teste for menor ou igual ao nível de significância α , isto é, $p\text{-valor} \leq \alpha$. Caso contrário, não se rejeita H_0 ao nível de significância α .

2.5.7 Teste de Wald de significância geral

O teste de significância geral de Wald é um teste mais amplo que possibilita testar se múltiplos parâmetros do modelo são iguais a zero, frequentemente utilizado para avaliar a significância do efeito de variáveis qualitativas com mais de dois níveis em modelos lineares generalizados (POWERS e XIE, 1999; HEERINGA et al., 2010). Suponha que o vetor β , de dimensão $p \times 1$, é particionado em r subvetores $(\beta_1, \beta_2, \dots, \beta_r)$, onde β_1 tem dimensão $p_1 \times 1$, β_2 tem dimensão $p_2 \times 1$, até β_r com dimensão $p_r \times 1$, tal que: $p_1 + p_2 + \dots + p_r = p$. Se o objetivo é testar, por exemplo, hipóteses sobre β_r , onde as hipóteses são dadas por:

$$\begin{cases} H_0: \beta_r = \mathbf{0} \\ H_1: \beta_r \neq \mathbf{0} \end{cases}$$

A estatística de teste (W) é dada por

$$W = (\hat{\beta}_r)' \hat{V}_r^{-1} (\hat{\beta}_r) \sim \chi_r^2 \quad (3.4)$$

onde W é a estatística de Wald que, sob a hipótese nula H_0 , tem distribuição Qui-quadrada com r graus de liberdade (r é a dimensão de β_r) e $\hat{V}_r = \widehat{VAR}(\hat{\beta}_r)$ é uma submatriz da matriz de variância-covariância de $\hat{\beta}$.

A região crítica do teste de Wald geral é dada por:

$$RC = \{w \in \mathbb{R} \mid w > \chi_{\alpha;r}^2\} \quad (3.5)$$

Com relação ao critério de decisão, rejeita-se H_0 ao nível de significância $100 * \alpha\%$ caso o valor observado w_{obs} da estatística de teste (W) pertença a região crítica (RC), isto é, β_r é significativamente diferente de zero. Em outras palavras, existe associação significativa entre a variável explicativa categórica em questão e a variável resposta. Caso o valor observado não pertença a RC , não se rejeita H_0 ao nível de significância de $100 * \alpha\%$, ou seja, β_r não é significativamente diferente de zero. Neste caso, todos os parâmetros que compõem o subvetor β_r não são significativamente diferentes de zero, indicando que não existe associação significativa entre a variável explicativa categórica em questão e a variável resposta do modelo.

Equivalentemente, rejeita-se H_0 se $p\text{-valor} \leq \alpha$ onde $p\text{-valor} = P(W > w_{obs})$. Caso contrário, não se rejeita H_0 ao nível de significância α .

2.5.8 Teste de Razão de Verossimilhanças

Segundo Powers e Xie (1999), o teste de razão de verossimilhanças (no inglês, *likelihood-ratio test*), é o teste estatístico mais comumente utilizado para testar múltiplos parâmetros, ou seja, é extremamente útil para avaliar a significância do efeito de variáveis explicativas categóricas, representadas por múltiplas variáveis dummy em MLGs. O teste de razão de verossimilhanças trata-se de um teste de comparabilidade de modelos aninhados, onde se compara o logaritmo da função de verossimilhança do modelo completo com o logaritmo da função de verossimilhança de outro modelo, designado como modelo reduzido.

Em termos de notação, suponha que o objetivo é comparar os dois modelos:

$$\begin{cases} H_0: \ln(\mu) = \mathbf{x}'_1 \beta_1 \\ H_1: \ln(\mu) = \mathbf{x}' \beta \end{cases} \quad (3.6)$$

onde na hipótese nula H_0 representa-se o modelo reduzido $\ln(\mu) = \mathbf{x}'_1 \boldsymbol{\beta}_1$, onde \mathbf{x}_1 é um vetor com p_1 variáveis explicativas. Na hipótese alternativa H_1 , por sua vez, representa-se o modelo completo $\ln(\mu) = \mathbf{x}' \boldsymbol{\beta}$, onde $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$ é um vetor com $p = (p_1 + p_2)$ variáveis explicativas.

Neste caso, testar as hipóteses definidas em (3.6) equivale a testar se a hipótese de que os parâmetros referentes às p_2 variáveis explicativas são ou não todos iguais a zero, isto é:

$$\begin{cases} H_0: \boldsymbol{\beta}_2 = 0 \\ H_1: \boldsymbol{\beta}_2 \neq 0 \end{cases}$$

A estatística de teste, utilizada nesse caso, é conhecida como estatística de razão de verossimilhanças (POWERS e XIE, 1999), ou ainda, como estatística ΔD (DOBSON, 2018), sendo esta última definida pela diferença (3.7) entre as deviances dos modelos reduzido (D_0) e completo (D_1). Deviance é um indicador de qualidade do ajuste que mede o desvio do modelo de interesse em relação ao modelo saturado (ou maximal) (POWERS E XIE, 1999).

$$\Delta D = D_0 - D_1 \quad (3.7)$$

onde:

$$D_0 = 2[\ln L(\widehat{\boldsymbol{\beta}}_{MAX}) - \ln L(\widehat{\boldsymbol{\beta}}_0)] \sim \chi^2_{n-(p_1+1)} \quad (3.7.1)$$

$$D_1 = 2[\ln L(\widehat{\boldsymbol{\beta}}_{MAX}) - \ln L(\widehat{\boldsymbol{\beta}}_1)] \sim \chi^2_{n-(p+1)} \quad (3.7.2)$$

onde $\ln L(\widehat{\boldsymbol{\beta}}_{MAX})$ é o logaritmo da função de verossimilhança para o modelo maximal (onde o número de parâmetros é igual ao número de observações), $\ln L(\widehat{\boldsymbol{\beta}}_0)$ é o logaritmo da função de verossimilhança do modelo reduzido e, $\ln L(\widehat{\boldsymbol{\beta}}_1)$ é o logaritmo da função de verossimilhança do modelo completo.

Substituindo (3.7.1) e (3.7.2) em (3.7), a estatística ΔD pode ser expressa da seguinte forma:

$$\Delta D = 2[\ln L(\widehat{\boldsymbol{\beta}}_1) - \ln L(\widehat{\boldsymbol{\beta}}_0)] \sim \chi^2_{p-p_1} \quad (3.8)$$

A estatística ΔD tem distribuição Qui-quadrada com $v = p - p_1$ graus de liberdade, onde v é dada pela diferença entre o número de variáveis explicativas (ou de parâmetros) dos dois modelos.

A região crítica do teste é dada por:

$$RC = \{\Delta D \in \mathbb{R} \mid \Delta D > \chi_{\alpha, p-p_1}^2\}$$

Quanto ao critério de decisão, rejeita-se H_0 , ao nível de significância $100 * \alpha\%$, se o valor observado da estatística de teste (ΔD) pertencer a região crítica (RC), isto é, os parâmetros referentes às p_2 variáveis explicativas não são todos iguais a zero. Caso o valor observado não pertença a RC, não se rejeita H_0 ao nível de significância de $100 * \alpha\%$, ou seja, os parâmetros referentes às p_2 variáveis explicativas são todos iguais a zero.

Equivalentemente, utilizando a abordagem baseada no p-valor do teste, rejeita-se H_0 se p-valor $\leq \alpha$, sendo p-valor = $P(\Delta D > \Delta D_{obs})$. Caso contrário, não se rejeita H_0 ao nível de significância α .

2.5.9 Medidas para avaliação da qualidade do ajuste do modelo

Para o modelo sob consideração, pode-se estimar a probabilidade (prevalência) do evento de interesse para cada uma das n unidades da amostra. Escolhendo-se um ponto de corte δ , os indivíduos podem ser classificados como tendo ou não tendo a característica de interesse. Assim, se a probabilidade estimada \hat{p}_i for maior que δ ($\hat{p}_i > \delta$), a unidade é classificada como tendo a característica de interesse (sucesso), caso contrário a unidade é classificada como não tendo a característica de interesse (fracasso). Em geral, o ponto de corte é $\delta = 0,5$, mas qualquer outro ponto de corte pode ser adotado para a classificação das unidades. Segundo Medronho et al. (2009), o melhor ponto de corte é aquele onde existe o menor erro possível, tanto de falsos positivos quanto de falsos negativos.

A capacidade preditiva do modelo pode ser avaliada através do cálculo da taxa global de classificações corretas e do cálculo das medidas de sensibilidade e especificidade. Essas medidas são calculadas com base numa tabela de contingência, onde as unidades da amostra são classificadas segundo as categorias observadas e previstas da variável resposta, a partir do modelo selecionado.

O Quadro 4 indica os possíveis resultados do ajuste de um modelo de regressão linear de Poisson para uma variável resposta binária, ou seja, ilustra a classificação das unidades da amostra segundo as categorias observadas e as categorias previstas da variável resposta.

Quadro 3: Classificação das unidades da amostra segundo as categorias observadas (linha) e as categorias preditivas (coluna) da variável resposta.

Categorias observadas da variável resposta	Categorias preditas da variável resposta		Total
	Sucesso ($\hat{Y} = 1$)	Fracasso ($\hat{Y} = 0$)	
Sucesso ($Y = 1$)	O_{11} (Verdadeiros Positivos)	O_{12} (Falsos Negativos)	$O_{11} + O_{12}$
Fracasso ($Y = 0$)	O_{21} (Falsos Positivos)	O_{22} (Verdadeiros Negativos)	$O_{21} + O_{22}$
Total	$O_{11} + O_{21}$	$O_{12} + O_{22}$	$\sum_{i=1}^2 \sum_{j=1}^2 o_{ij}$

A taxa global de classificações corretas (TG) é dada pela razão entre o número de classificações corretas pelo modelo e o número total de observações. O valor é apresentado em forma percentual e calculado usando a expressão (4.1):

$$TG = \left(\frac{o_{11} + o_{22}}{\sum_{i=1}^2 \sum_{j=1}^2 o_{ij}} \right) * 100 \quad (4.1)$$

A sensibilidade (S) é uma medida quantitativa que avalia a capacidade do modelo em classificar a unidade corretamente quando a mesma tem a característica de interesse, isto é, classificar a unidade como tendo a característica de interesse quando a unidade de fato possui essa característica. É calculada pela proporção de verdadeiros positivos entre o total de unidades que possuem a característica, como indicado em (4.2).

$$S = P(\hat{Y} = 1 | Y = 1) = \left(\frac{o_{11}}{o_{11} + o_{12}} \right) * 100 \quad (4.2)$$

Já a especificidade (E) avalia a capacidade do modelo em classificar a unidade corretamente quando a mesma não tem a característica de interesse, isto é, classificar a unidade como não tendo a característica de interesse quando ela de fato não possui essa característica. É calculada pela proporção de verdadeiros negativos entre o total de unidades que não possuem a característica, através da expressão (4.3).

$$E = P(\hat{Y} = 0 | Y = 0) = \left(\frac{o_{22}}{o_{21} + o_{22}} \right) * 100 \quad (4.3)$$

De acordo com Marôco (2010), um modelo com boa capacidade preditiva apresenta valores de sensibilidade e especificidade maiores que 80%. Para valores entre 50% e 80%, o modelo tem capacidade preditiva considerada razoável. Já para valores inferiores a 50%, a capacidade preditiva do modelo é classificada como ruim.

Marôco (2010) ainda cita outra medida de capacidade do modelo para discriminar os sujeitos (unidades) com a característica de interesse *versus* sem a característica de interesse, a área sob a curva ROC (*Receiver Operating Characteristic*). A área da curva ROC varia entre 0 e 1, e quanto mais próxima de 1 melhor a capacidade preditiva do modelo. Se a área sob a curva for igual a 0.5 o modelo não consegue discriminar os sujeitos melhor do que a escolha por mero acaso.

A curva ROC é um gráfico onde temos os valores da sensibilidade no eixo das ordenadas e os valores referentes ao complemento da especificidade ($1 - \text{especificidade}$) no eixo das abscissas (Figura 1) (HOSMER & LEMESHOW, 2000). A curva ROC pode ser utilizada para auxiliar na identificação do melhor ponto de corte (MEDRONHO et al., 2009).

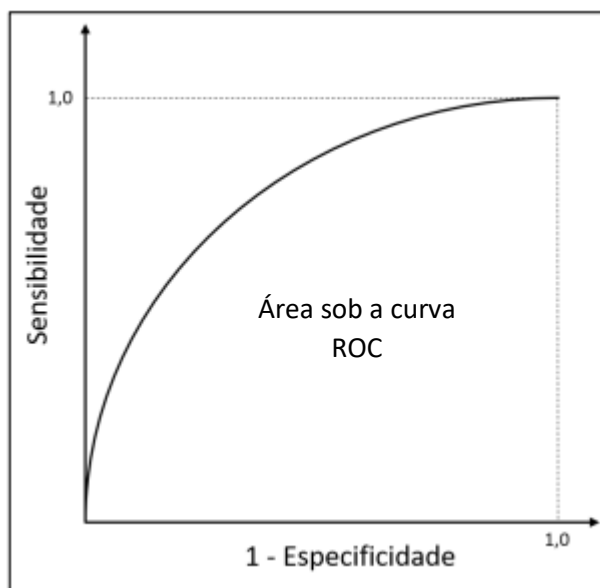


Figura 1: Curva ROC (adaptada de HOSMER & LEMESHOW, 2000).

Outra medida quantitativa utilizada para avaliar a qualidade de ajuste de MLGs é o Pseudo- R^2 proposto por McFadden. Essa medida necessariamente está compreendida no intervalo $[0,1]$, sendo 0 quando o modelo selecionado não oferece nenhuma melhora sobre o modelo nulo e 1 quando o modelo selecionado oferece um ajuste perfeito aos dados observados. Segundo Marôco (2010), o pseudo R de McFadden é uma medida que

fornece uma melhor interpretabilidade, quando comparadas a outras medidas de Pseudo- R^2 , e representa o ganho de informação estimada pelo modelo sob consideração em comparação ao modelo nulo. Em outras palavras, o pseudo R^2 de McFadden indica a melhoria proporcional no logaritmo da função de verossimilhança devido aos termos do modelo de interesse, comparativamente ao modelo nulo (DOBSON E BARNETT, 2018).

A medida de pseudo R^2 de McFadden é dada pela seguinte expressão:

$$R_{MF}^2 = 1 - \frac{\ln L(\hat{\beta})}{\ln L(\hat{\beta}^{(0)})} \quad (4.4)$$

onde $\ln L(\hat{\beta}^{(0)})$ é o logaritmo da função de verossimilhança do modelo nulo, e $\ln L(\hat{\beta})$ é o logaritmo da função de verossimilhança do modelo sob consideração.

3 Resultados

A Tabela 1 apresenta a distribuição dos recém-nascidos com ou sem baixo peso ao nascer segundo as variáveis de estudo.

Tabela 1: Distribuição (absoluta e percentual) dos recém-nascidos com ou sem baixo peso ao nascer segundo as características sociodemográficas maternas, da gestação e do parto e do recém-nascido.

Variáveis de estudo	Baixo peso ao nascer (BPN)				Total
	Não (n=156.888)		Sim (n=16.385)		
	n	%	n	%	
Mesorregião					
Baixada	7484	91,8	671	8,2	8155
Centro Fluminense	4551	91,7	414	8,3	4965
Metropolitana	121082	90,5	12727	9,5	133809
Noroeste Fluminense	3146	92,0	273	8,0	3419
Norte Fluminense	10535	90,0	1172	10,0	11707
Sul Fluminense	10090	89,9	1128	10,1	11218
Faixa etária materna					
15 a 19 anos	21410	89,3	2556	10,7	23966
20 a 34 anos	108783	91,2	10482	8,8	119265
35 anos ou mais	26695	88,9	3347	11,1	30042
Situação conjugal					
Sem companheiro	102712	90,2	11112	9,8	113824
Com companheiro	54176	91,1	5273	8,9	59449
Escolaridade materna					
Sem instrução ou até 3 anos de estudo	1950	87,3	284	12,7	2234
De 4 a 7 anos de estudo	23948	89,6	2791	10,4	26739
De 8 a 11 anos de estudo	97375	90,8	9904	9,2	107279
12 anos de estudo ou mais	33615	90,8	3406	9,2	37021
Paridade					
Nulípara	60519	89,6	7041	10,4	67560
De 1 a 3 filhos	86741	91,5	8031	8,5	94772
4 filhos ou mais	9628	88,0	1313	12,0	10941
Idade gestacional					
Pré termo (menos que 37 semanas)	9169	47,0	10338	53,0	19507
A termo (de 37 a 41 semanas)	145423	96,1	5963	3,9	151386
Pós termo (42 semanas ou mais)	2296	96,5	84	3,5	2380
Tipo de parto					
Vaginal	68004	92,3	5709	7,7	73713
Cesáreo	88884	89,3	10676	10,7	99560

Continua

Tabela 1: Continuação

Variáveis de estudo	Baixo peso ao nascer (BPN)				Total
	Não (n=156.888)		Sim (n=16.385)		
	n	%	n	%	
Número de consultas pré-natal					
Nenhuma	1998	81,0	469	19,0	2467
De 1 a 3 consultas	7041	80,0	1760	20,0	8801
De 4 a 6 consultas	28365	84,8	5079	15,2	33444
7 consultas ou mais	119484	92,9	9077	7,1	128561
Sexo do bebê					
Masculino	79570	91,3	7573	8,7	87143
Feminino	77318	89,8	8812	10,2	86130
Raça/Cor da mãe					
Branca	53789	91,1	5260	8,9	59049
Não branca	103099	90,3	11125	9,7	114224
Índice de Apgar no quinto minuto					
De 0 a 3 (muito baixo)	201	47,9	219	52,1	420
De 4 a 6 (baixo)	714	55,1	583	44,9	1297
Escore de 7 a 10 (normalidade)	155973	90,9	15583	9,1	171556

Aproximadamente 9,5% dos recém-nascidos no estado do Rio de Janeiro apresentaram BPN. Observa-se também maior percentual de BPN entre recém-nascidos de sexo feminino e entre os recém-nascidos que apresentaram índices de Apgar abaixo do normal (com escores inferiores a 7). Maior percentual de BPN foi encontrado entre os recém-nascidos prematuros, isto é, os bebês pré termo apresentaram um maior índice de BPN comparados aos bebês a termo e pós termo (Tabela 1).

No que se refere às consultas de acompanhamento pré-natal, maiores percentuais de BPN foram observados entre os bebês que compareceram a poucas (1 a 3 consultas) ou nenhuma consultas pré-natais.

Com relação as mesorregiões do estado do Rio de Janeiro, observa-se que proporções ligeiramente maiores de recém-nascidos com BPN estão nas regiões Norte Fluminense (10,0%) e Sul Fluminenses (10,1%), seguido da região Metropolitana (9,5%).

Quanto as características sociodemográficas das mães, verifica-se que as mães nas faixas etárias de 15 a 19 anos e 35 anos ou mais apresentaram uma proporção de recém-nascidos com BPN maiores do que as mães na faixa etária de 20 a 34 anos. As mães sem companheiros apresentaram uma proporção maior de filhos com BPN em comparação as mães com companheiros. Com relação a escolaridade materna, observa-se, de modo

geral, que quanto menor o nível de escolaridade da mãe, maior o percentual de recém-nascidos com BPN. Além disso, maior percentual de BPN foi observado entre as mães de cor não branca.

Quanto a variável paridade, as mães nulíparas ou com 4 filhos ou mais apresentam uma proporção maior de BPN, comparativamente às mães com 1 a 3 filhos.

As mães que realizaram parto cesáreo apresentaram uma proporção maior de recém-nascidos com BPN em relação as mães que realizaram parto vaginal.

A Tabela 2 apresenta os principais resultados encontrados ao ajustar o modelo log-linear de Poisson (com variância robusta) considerando todas as variáveis explicativas com as suas respectivas categorias previamente listadas no Quadro 1. Entre os resultados do ajuste do modelo, estão as razões de prevalência de BPN com seus intervalos de 95% de confiança e os p-valores obtidos no teste de Wald de significância individual e geral.

Tabela 2: Resultados do ajuste do modelo log-linear de Poisson (com variância robusta) considerando as características sociodemográficas maternas, da gestação e do parto, e do recém-nascido, com as suas respectivas categorias, previamente listadas no Quadro 1 (modelo original).

Variáveis de estudo	Modelo log-linear de Poisson (modelo original)		
	RP	IC (RP, 95%)	p-valor (Wald)
Mesorregião			< 0,001
Baixada	0,964	[0,837 ; 1,111]	0,551
Centro Fluminense	1,173	[1,007 ; 1,367]	0,016
Metropolitana	1,223	[1,084 ; 1,379]	< 0,001
Noroeste Fluminense	1	-	-
Norte Fluminense	1,186	[1,040 ; 1,354]	0,002
Sul Fluminense	1,180	[1,034 ; 1,347]	0,003
Faixa Etária			< 0,001
15 a 19 anos	1,001	[0,953 ; 1,050]	0,975
20 a 34 anos	1	-	-
35 anos ou mais	1,107	[1,063 ; 1,154]	< 0,001
Situação Conjugal			0,007
Sem companheiro	1,041	[1,004 ; 1,080]	0,007
Com companheiro	1	-	-

Continua

Tabela 2: Continuação

Variáveis de estudo	Modelo log-linear de Poisson (modelo original)		
	RP	IC (RP, 95%)	p-valor (Wald)
Escolaridade Materna			< 0,001
Sem instrução ou até 3 anos de estudo	1,157	[1,020 ; 1,312]	0,006
De 4 a 7 anos de estudo	1,079	[1,017 ; 1,145]	0,001
De 8 a 11 anos de estudo	1,021	[0,977 ; 1,067]	0,238
12 anos de estudo ou mais	1	-	-
Idade gestacional			< 0,001
Pré termo (menos que 37 semanas)	11,802	[11,416 ; 12,201]	< 0,001
A termo (de 37 a 41 semanas)	1	-	-
Pós termo (42 semanas ou mais)	0,915	[0,738 ; 1,135]	0,410
Paridade			< 0,001
Nulípara	1	-	-
De 1 a 3 filhos	0,793	[0,766 ; 0,821]	< 0,001
4 filhos ou mais	0,835	[0,782 ; 0,892]	< 0,001
Tipo de Parto			< 0,001
Vaginal	1	-	-
Cesáreo	1,268	[1,225 ; 1,312]	< 0,001
Número de consultas pré-natal			< 0,001
Nenhuma	1	-	-
De 1 a 3 consultas	0,944	[0,852 ; 1,045]	0,122
De 4 a 6 consultas	0,817	[0,743 ; 0,899]	< 0,001
7 consultas ou mais	0,567	[0,515 ; 0,623]	< 0,001
Sexo do bebê			< 0,001
Masculino	1	-	-
Feminino	1,240	[1,202 ; 1,279]	< 0,001
Raça/Cor da mãe			< 0,001
Branca	1	-	-
Não branca	1,075	[1,038 ; 1,113]	< 0,001
Índice de Apgar no quinto minuto			< 0,001
De 0 a 3 (muito baixo)	1,618	[1,415 ; 1,851]	< 0,001
De 4 a 6 (baixo)	1,548	[1,423 ; 1,684]	< 0,001
Escore de 7 a 10 (normalidade)	1	-	-

De acordo com os resultados do teste de Wald de significância geral realizado observa-se que todas as variáveis apresentaram associação significativa com a prevalência de BPN ao nível de significância de 5%. Entretanto, a partir do teste de Wald individual, observa-se que apenas algumas categorias de certas variáveis explicativas não tiveram efeito significativo considerando o nível de significância de 5%. São elas: da

variável Mesorregião a categoria “Baixada” (p-valor=0,551), da variável Faixa Etária a categoria “15 a 19 anos” (p-valor=0,975), da variável Escolaridade Materna a categoria “de 8 a 11 anos de estudo” (p-valor=0,238), da variável Idade gestacional a categoria “Pós termo” (p-valor=0,410) e da variável Número de consultas pré-natal a categoria “de 1 a 3 consultas” (p-valor=0,122). Os intervalos de 95% de confiança das razões de prevalência de BPN relativas as categorias mencionadas incluem a unidade, indicando que os efeitos destas categorias não são estatisticamente significativos (Tabela 2).

Do exposto, optou-se por agregar as categorias não significativas das variáveis explicativas categóricas ordinais com as suas respectivas categorias de referência. A Tabela 3 fornece os novos resultados do ajuste do modelo log-linear de Poisson, isto é, as novas medidas de razão de prevalência de BPN (RP), com seus intervalos de 95% de confiança, e os novos p-valores para o teste de Wald.

Tabela 3: Resultados do ajuste do modelo log-linear de Poisson (com variância robusta) considerando as características sociodemográficas maternas, da gestação e do parto, e do recém-nascido, agregando as categorias não significativas (modelo selecionado).

Variáveis de estudo	Modelo log-linear de Poisson (modelo selecionado)		
	RP	IC (RP, 95%)	p-valor (Wald)
Mesorregião			< 0,001
Baixada	0,963	[0,836 ; 1,109]	0,536
Centro Fluminense	1,174	[1,007 ; 1,368]	0,015
Metropolitana	1,221	[1,083 ; 1,377]	< 0,001
Noroeste Fluminense	1	-	-
Norte Fluminense	1,187	[1,040 ; 1,354]	0,002
Sul Fluminense	1,179	[1,033 ; 1,346]	0,004
Faixa Etária			< 0,001
15 a 34 anos	1	-	-
35 anos ou mais	1,104	[1,060 ; 1,150]	< 0,001
Situação Conjugal			0,003
Sem companheiro	1,045	[1,008 ; 1,082]	0,003
Com companheiro	1	-	-
Escolaridade Materna			< 0,001
Sem instrução ou até 3 anos de estudo	1,137	[1,008 ; 1,281]	0,011
De 4 a 7 anos de estudo	1,060	[1,015 ; 1,106]	0,001
8 anos de estudo ou mais	1	-	-
Prematuridade			< 0,001
Sim (menos que 37 semanas)	11,812	[11,428 ; 12,210]	< 0,001
Não (37 semanas ou mais)	1	-	-

Continua

Tabela 3: Continuação

Variáveis de estudo	Modelo log-linear de Poisson (modelo selecionado)		
	RP	IC (RP, 95%)	p-valor (Wald)
Paridade			< 0,001
Nulípara	1	-	-
De 1 a 3 filhos	0,794	[0,769 ; 0,821]	< 0,001
4 filhos ou mais	0,838	[0,787 ; 0,893]	< 0,001
Tipo de Parto			< 0,001
Vaginal	1	-	-
Cesáreo	1,265	[1,223 ; 1,309]	< 0,001
Número de consultas pré-natal			< 0,001
Até 3 consultas	1	-	-
De 4 a 6 consultas	0,856	[0,814 ; 0,900]	< 0,001
7 consultas ou mais	0,593	[0,564 ; 0,623]	< 0,001
Sexo do bebê			< 0,001
Masculino	1	-	-
Feminino	1,240	[1,202 ; 1,279]	< 0,001
Raça/Cor da mãe			< 0,001
Branca	1	-	-
Não branca	1,079	[1,042 ; 1,116]	< 0,001
Índice de Apgar no quinto minuto			< 0,001
De 0 a 3 (muito baixo)	1,620	[1,416 ; 1,853]	< 0,001
De 4 a 6 (baixo)	1,550	[1,425 ; 1,685]	< 0,001
Escore de 7 a 10 (normalidade)	1	-	-

Na análise do modelo selecionado observa-se que os recém-nascidos do sexo feminino apresentaram uma prevalência de BPN 24,0% maior em comparação aos do sexo masculino (RP=1,240; p-valor<0,001). Os recém-nascidos que apresentaram índices de Apgar muito baixo e baixo apresentaram prevalências de BPN 62,0% (RP=1,620; p-valor<0,001) e 55,0% (RP=1,240; p-valor<0,001), respectivamente maiores em comparação aos recém-nascidos que apresentaram índice de Apgar normal. Os bebês prematuros apresentaram uma prevalência de BPN quase 12 vezes maior (RP=11,812; p-valor < 0,001) em relação aos não prematuros.

Com relação ao número de consultas pré-natal, observa-se que as mães que realizaram de 4 a 6 consultas apresentaram uma prevalência de BPN 14,4% menor do que as mães que realizaram até 3 consultas (RP=0,856; p-valor<0,001). Já as mães com 7 ou mais consultas apresentaram uma prevalência de BPN 40,7% menor em comparação as mães que realizaram até 3 consultas (RP=0,593; p-valor<0,001). Deste modo, à medida

que aumenta o número de consultas pré-natal, observa-se uma redução na prevalência de BPN.

Comparativamente ao Noroeste Fluminense nota-se, com exceção da Baixada, que todas as mesorregiões apresentaram maior prevalência de BPN. A maior diferença (prevalência de BPN 22,1% maior) foi encontrada na região Metropolitana versus Noroeste Fluminense. Com relação à Baixada, não se observou uma diferença estatisticamente significativa entre a sua prevalência de BPN e a da região Noroeste Fluminense (p-valor > 0,05).

Quanto as características sociodemográficas das mães, observa-se que as mães com 35 anos ou mais de idade apresentaram uma prevalência de BPN 10,4% maior em comparação as mães com 15 a 34 anos (RP=1,104; p-valor<0,001). As mães sem companheiros apresentaram uma prevalência de BPN 4,5% maior em relação as mães com companheiros (RP=1,045; p-valor=0,003). Quanto a escolaridade materna, observa-se que as mães sem instrução ou com até 3 anos de estudo apresentaram uma prevalência de recém-nascidos com BPN 13,7% maior em comparação as mães com 8 anos ou mais de estudo (RP=1,137; p-valor=0,011); e as mães com 4 a 7 anos de estudo apresentaram uma prevalência de BPN 6% maior em relação as mães com 8 anos ou mais de estudo (RP=1,060; p-valor=0,001). Também pode-se notar que as mães de cor não branca apresentaram uma prevalência de recém-nascidos com BPN quase 8% maior em comparação com as mães de cor branca (RP=1,079; p-valor<0,001).

Em comparação as mães nulíparas, as mães com 1 a 3 filhos e as mães com 4 filhos ou mais apresentaram prevalências de BPN 20,6% e 16,2% menores, respectivamente. Por fim, as mães que realizaram parto cesáreo apresentaram uma prevalência de recém-nascidos com BPN 26,5% maior em comparação as mães que realizaram parto vaginal (RP=1,265; p-valor<0,001).

Tabela 4: Medidas para avaliação da qualidade do ajuste do modelo log-linear de Poisson (original versus selecionado).

Medidas	Modelo original (categorias não agregadas)	Modelo selecionado (categorias agregadas)
Taxa global de classificações corretas*	91,092%	91,093%
Sensibilidade*	63,399%	63,375%
Especificidade*	93,984%	93,983%
Pseudo R ² McFadden	23,980%	23,977%
Área sob a curva ROC	0,8288	0,8286
Teste de Razão de Verossimilhanças	$\Delta D=26.392$; p-valor<0,001	$\Delta D=26.390$; p-valor<0,001

*O ponto de corte ótimo em ambos os modelos foi de 0,083

Na Tabela 4 observa-se que tanto o modelo original (sem as categorias agregadas) quanto o modelo selecionado (com as categorias agregadas) possuem a mesma capacidade preditiva ou discriminatória (Figura 2), isto é, agregando ou não as categorias não significativas das variáveis explicativas ordinais não há alteração da capacidade preditiva ou discriminatória dos modelos. Além disso, com base no teste de razão de verossimilhanças, observa-se que ambos os modelos são mais adequados do que o modelo nulo (incluindo somente a constante).

A Figura 2 apresenta a curva ROC do modelo selecionado e a respectiva área sob a curva ROC. Para uma área sob a curva ROC igual a 0,829, pode-se dizer que o modelo selecionado apresenta uma boa capacidade para discriminar os bebês que tem BPN e os que não tem BPN.

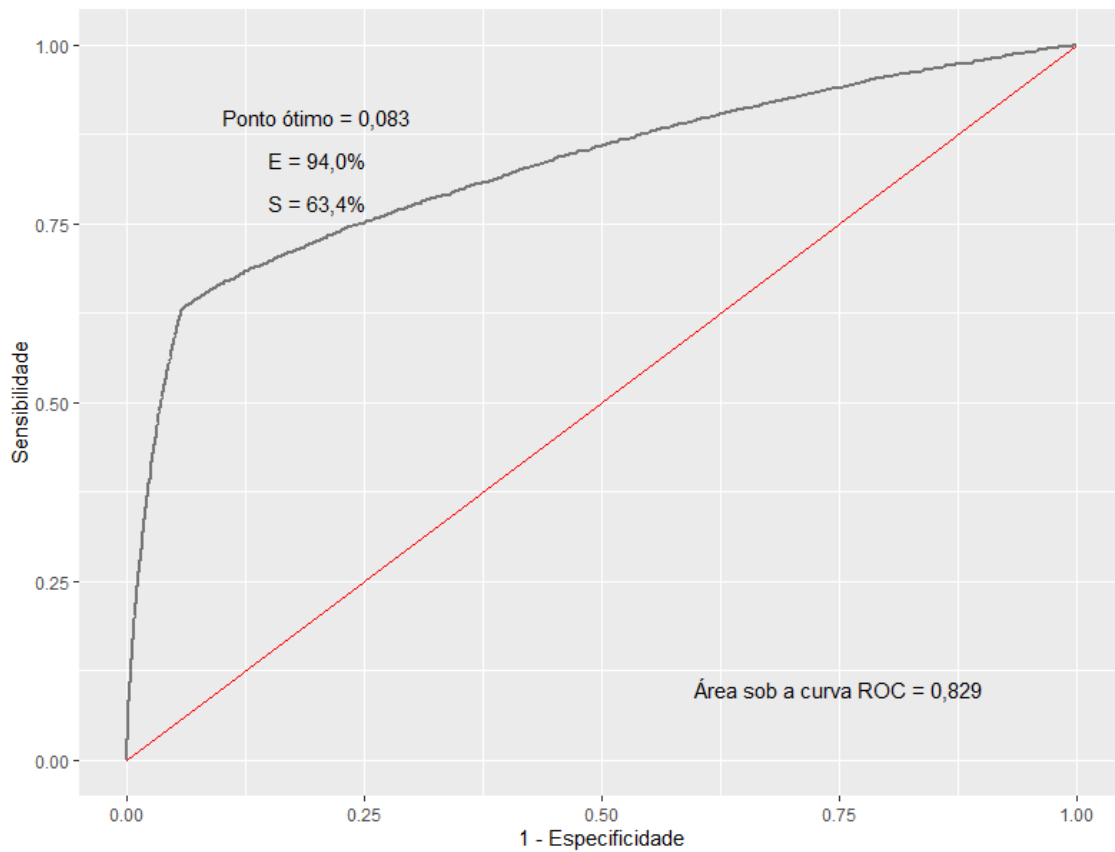


Figura 2: Curva ROC do modelo selecionado

Considerando o valor 0,083 como ponto de corte ótimo obteve-se uma especificidade de 94%, sensibilidade 63,4% e uma taxa global de classificações corretas de 91,1%. Desse modo, 91,1% do total de recém-nascidos são classificados corretamente pelo modelo selecionado. Além disso, 63,4% dos recém-nascidos com BPN foram classificados corretamente pelo modelo selecionado, enquanto 94% dos recém-nascidos sem BPN foram classificados corretamente (Tabela 4).

4 Discussão e Conclusão

O presente estudo encontrou uma maior prevalência de baixo peso ao nascer entre os recém-nascidos de mães não brancas, com baixa escolaridade, que realizaram um número abaixo do recomendado de consultas pré-natal e prematuros. Da mesma forma, no estudo transversal desenvolvido por Belfort et al.(2018), com 751 adolescentes atendidas em maternidade pública do Rio de Janeiro, verificou-se, por meio de modelos de regressão logística, que o parto prematuro e o menor número de consultas de pré-natal estão significativamente associadas ao BPN. Com relação a raça/cor (não branca *versus* branca) e a situação conjugal (sem *versus* com companheiro), Belford et al.(2018) observaram na análise bivariada apenas associação entre viver sem companheiro e a maior chance de BPN, mas esta associação não foi significativa na análise multivariada. Além da associação da prematuridade e do baixo número de consultas pré-natal com o BPN, no estudo de Santos et al.(2008), observou-se ainda, via regressão logística múltipla, associação entre a menor escolaridade materna e o BPN, mas não entre a situação conjugal das mães e este desfecho. Diferentemente do presente trabalho, Santos et al. (2008) consideraram na análise apenas os recém-nascidos filhos de mães na faixa etária de 10 a 34 anos.

Moreira et al.(2018), como neste trabalho, também verificaram uma associação positiva entre o parto cesáreo e o desfecho de BPN, e salientaram que esta associação é complexa, tendo em vista que este tipo parto pode ser recomendado devido à complicações relacionadas a gestação e ao feto, que, por sua vez, levaram a recém-nascidos com peso abaixo de 2.500 gramas.

Com relação a idade da mãe, outros estudos já mostraram associação entre a idade materna e o BPN, utilizando diferentes faixas-etárias. Por exemplo, no estudo de Maia e Souza (2010), via modelo de regressão de Poisson (com variância robusta), observou-se que idades maternas entre 12 e 13 anos, 16 e 17 anos, 18 e 19 anos, 35 anos ou mais, em comparação a faixa de 20 a 24 anos, estão associadas a uma maior prevalência de BPN. No estudo de Tolio et al. (2020), por sua vez, foi incluída na faixa etária os extremos da idade reprodutiva (menos de 20 anos, 20 a 34 anos, 35 anos ou mais), e os autores constataram que a idade materna maior ou igual a 35 anos está associada a maior chance de BPN, considerando a faixa de 20 a 34 anos como referência.

O sexo do recém-nascido feminino também se mostrou associado a maior ocorrência de BPN no estudo de Nascimento et al.(2019), considerando recém-nascidos atendidos em um hospital público de alta complexidade, no município de Vitória-ES. No estudo desenvolvido por Uchimura et al. (2007) com 30.383 crianças nascidas vivas de gestação única em Maringá – PR, no período de 1996 a 2002, o índice de Apgar de 5º minuto esteve associado ao BPN, indicando um maior “risco” de BPN para as crianças com escore de Apgar no 5º minuto mais baixo, corroborando com o resultado obtido neste estudo.

Com relação à potencialidade deste trabalho, destaca-se o uso do modelo log-linear de Poisson com estimação de variância robusta. O modelo log-linear de Poisson é comumente utilizado na área epidemiológica onde o desfecho é o número de sucessos que ocorreram durante um certo intervalo, mas também pode e deve ser utilizado em estudos de natureza transversal com variável resposta binária, onde considera-se que todos os indivíduos tem o mesmo tempo de seguimento. Contudo, quando este modelo de regressão é aplicado à desfechos binários ocorre em geral uma superestimação da variância (ou erro-padrão) da razão de prevalência (COUTINHO et al, 2008; BARROS & HIRAKATA, 2003). Este problema é solucionado com o ajuste do modelo utilizando um estimador de variância robusta, possibilitando a obtenção adequada de intervalos de confiança e de p-valores dos testes de significância dos parâmetros de interesse (ZOU, 2004).

Quanto às limitações do estudo destacam-se o sub-registro e/ou o registro incorreto das informações maternas e do bebê na Declaração de Nascido Vivo disponível no SINASC. A literatura aponta uma variação entre 75,8% e 99,5% nos casos de sub-registro, a depender da região analisada. A pouca clareza metodológica do manual de preenchimento da Declaração de Nascidos Vivos (DN), a heterogeneidade de profissionais responsáveis pelo preenchimento da DN, a diferença de interesse entre uma variável e outra, o mau preenchimento e os erros de grafia são algumas das causas dos registros incorretos no SINASC, destacando-se as variáveis paridade, onde registra-se indistintamente tanto o número de partos anteriores ou o número de gestações anteriores, e a variável instrução materna, que muitas vezes não é preenchida ou é preenchida de maneira incorreta por não ser considerada importante pelo responsável pelo preenchimento (PEDRAZA, 2011), além da falta de capacitação para o preenchimento da declaração e do elevado número de informações preenchidas com a opção "ignorado".

Uma outra limitação deve-se a própria natureza do estudo de corte de transversal baseado nos dados do SINASC, que impossibilita o estabelecimento de qualquer tipo de relação causal entre as características investigadas e o desfecho de BPN.


Também é importante destacar que, embora seja mundialmente utilizado e um grande marco na história da neonatologia, a variável utilizada no estudo “índice de Apgar” recebe críticas da comunidade científica com relação à acurácia do índice para determinar o bem-estar do recém-nascido. Ainda, estudos indicam que esse índice pode ser acabar subestimado em partos vaginais e superestimado em partos cesáreos (SCHARDOSIM et al, 2018). Contudo, a utilidade estatística deste índice é reconhecida quando necessário analisar o estado de saúde dos recém-nascidos nos primeiros minutos de vida.

O BPN, em especial o extremo baixo peso, tem posição de destaque como fator associado ao óbito neonatal e durante o primeiro ano, além de problemas de desenvolvimento durante a infância (TIAN et al., 2019; WILCOX, 2001). É importante destacar também a associação entre o baixo escore no índice de Apgar e a ocorrência de óbito neonatal. Como bem apontado por GAÍVA (2018), os fatores que aumentam o risco de morte na população com BPN podem ser combatidos com melhores práticas assistenciais, em especial nos primeiros dias de vida.

Do exposto, conclui-se que é necessário um olhar mais atento por parte dos profissionais de saúde e dos formuladores de políticas públicas às gestantes com 35 anos ou mais, baixa escolaridade, sem companheiro, de raça/cor não branca, com pouca frequência nas consultas pré-natal e nulíparas, para que recebam apoio a fim de reduzir o número de recém-nascidos com baixo peso ao nascer. Assim como, recomenda-se a realização de mais estudos com bebês que nascem prematuramente, por meio de parto cesáreo, de sexo feminino e que apresentam índice de Apgar abaixo do ideal, a fim de compreender melhor como se estabelece a relação de tais fatores com o baixo peso ao nascer.

Anexos

ANEXO A - Modelo da Declaração de Nascido Vivo



República Federativa do Brasil
Ministério da Saúde
1ª VIA - SECRETARIA DE SAÚDE

Declaração de Nascido Vivo

I Identificação do recém-nascido	1 Nome do Recém-nascido																	
	2 Data e hora do nascimento Data: _____ Hora: _____						3 Sexo <input type="checkbox"/> M - Masculino <input type="checkbox"/> F - Feminino <input type="checkbox"/> I - Ignorado											
II Local de ocorrência	4 Peso ao nascer em gramas			5 Índice de Apgar 1º minuto _____ 5º minuto _____			6 Detectada alguma anomalia congênita? Caso afirmativo, usar o bloco anomalia congênita para descrevê-las 1 <input type="checkbox"/> Sim 2 <input type="checkbox"/> Não 9 <input type="checkbox"/> Ignorado											
	7 Local da ocorrência 1 <input type="checkbox"/> Hospital 2 <input type="checkbox"/> Outros estab. saúde 3 <input type="checkbox"/> Domicílio 4 <input type="checkbox"/> Outros			8 Estabelecimento Ignorado <input type="checkbox"/> 9			Código CNES											
	9 Endereço da ocorrência, se fora do estab. ou da resid. da Mãe (rua, praça, avenida, etc) Número Complemento 10 CEP																	
III Mãe	11 Bairro/Distrito						Código			12 Município de ocorrência			Código			13 UF		
	14 Nome da Mãe												15 Cartão SUS					
	16 Escolaridade (última série concluída) Nível 0 <input type="checkbox"/> Sem escolaridade 1 <input type="checkbox"/> Fundamental I (1ª a 4ª série) 2 <input type="checkbox"/> Fundamental II (5ª a 8ª série) 3 <input type="checkbox"/> Médio (antigo 2º grau) 4 <input type="checkbox"/> Superior incompleto 5 <input type="checkbox"/> Superior completo Ignorado <input type="checkbox"/> 9						Série			17 Ocupação habitual (Informar anterior, se aposentada/desempregada) Código CBO 2002								
18 Data nascimento da Mãe			19 Idade (anos)			20 Naturalidade da Mãe Município / UF (se estrangeiro informar País)			21 Situação conjugal 1 <input type="checkbox"/> Solteira 2 <input type="checkbox"/> Casada 3 <input type="checkbox"/> Viúva 4 <input type="checkbox"/> Separada judicialmente/divorçada 5 <input type="checkbox"/> União estável 9 <input type="checkbox"/> Ignorada			22 Raça / Cor da Mãe 1 <input type="checkbox"/> Branca 2 <input type="checkbox"/> Preta 3 <input type="checkbox"/> Amarela 4 <input type="checkbox"/> Parda 5 <input type="checkbox"/> Indígena						
IV Pai	Residência da Mãe 16 Logradouro Número Complemento 19 CEP																	
	23 Bairro/Distrito			Código			26 Município			Código			27 UF					
	28 Nome do Pai												29 Idade do Pai					
V Gestação e parto	30 Gestações anteriores 30 Histórico gestacional ■ Nº gestações anteriores _____ ■ Nº de partos vaginais _____ ■ Nº de cesáreas _____ ■ Nº de nascidos vivos _____ ■ Nº de perdas fetais / abortos _____																	
	31 Idade Gestacional 31 Data da Última Menstruação (DUM) ____/____/____						33 Número de consultas de pré-natal			34 Mês de gestação em que iniciou o pré-natal			35 Tipo de gravidez 1 <input type="checkbox"/> Única 2 <input type="checkbox"/> Dupla 3 <input type="checkbox"/> Tripla ou mais 9 <input type="checkbox"/> Ignorado					
	32 Nº de semanas de gestação, se DUM Ignorada _____ Método utilizado para estimar 1 <input type="checkbox"/> Exame Físico 2 <input type="checkbox"/> Outro método 9 <input type="checkbox"/> Ignorado						36 Apresentação 1 <input type="checkbox"/> Cefálica 2 <input type="checkbox"/> Pélvica ou Podálica 3 <input type="checkbox"/> Transversa 9 <input type="checkbox"/> Ignorado			37 O Trabalho de parto foi induzido? 1 <input type="checkbox"/> Sim 2 <input type="checkbox"/> Não 9 <input type="checkbox"/> Ignorado			38 Tipo de parto 1 <input type="checkbox"/> Vaginal 2 <input type="checkbox"/> Cesáreo 9 <input type="checkbox"/> Ignorado					
VI Anomalia congênita	39 Cesárea ocorreu antes do trabalho de parto iniciar? 1 <input type="checkbox"/> Sim 2 <input type="checkbox"/> Não 3 <input type="checkbox"/> Não se aplica 9 <input type="checkbox"/> Ignorado																	
	40 Nascimento assistido por 1 <input type="checkbox"/> Médico 2 <input type="checkbox"/> Enfermeira/Ostetiz 3 <input type="checkbox"/> Parteira 4 <input type="checkbox"/> outros 9 <input type="checkbox"/> Ignorado																	
VII Preenchimento	41 Descrever todas as anomalias congênitas observadas																	
	42 Data do preenchimento			43 Nome do responsável pelo preenchimento			44 Função 1 <input type="checkbox"/> Médico 2 <input type="checkbox"/> Enfermeiro 3 <input type="checkbox"/> Parteira 4 <input type="checkbox"/> Func. Cartório 5 <input type="checkbox"/> Outros (descrever)											
VIII Cartório	45 Tipo documento 1 <input type="checkbox"/> CNES 2 <input type="checkbox"/> CRM 3 <input type="checkbox"/> COREN 4 <input type="checkbox"/> RG 5 <input type="checkbox"/> CPF			46 Nº do documento			47 Órgão emissor											
	48 Cartório			Código			49 Registro			50 Data								
51 Município												52 UF						

ATENÇÃO: ESTE DOCUMENTO NÃO SUBSTITUI A CERTIDÃO DE NASCIMENTO
O Registro de Nascimento é obrigatório por lei.
Para registrar esta criança, o pai ou responsável deverá levar este documento ao cartório de registro civil.

Versão 01/10 - 2ª Impressão 11/2010

Referências

- ALLEN J WILCOX. *International Journal of Epidemiology*, Volume 30, Issue 6, December 2001, Pages 1233–1241, <https://doi.org/10.1093/ije/30.6.1233>.
- BARROS AJ, HIRAKATA VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC Med Res Methodol*. 2003 Oct;3:21.
- BELFORT, Gabriella Pinto et al . Determinantes do baixo peso ao nascer em filhos de adolescentes: uma análise hierarquizada. *Ciênc. saúde coletiva*, Rio de Janeiro , v. 23, n. 8, p. 2609-2620, ago. 2018.
- BISMARCK-NASR, Elizabeth Maria; FRUTUOSO, Maria Fernanda Petrolí; GAMABARDELLA, Ana Maria Dianezi. Efeitos tardios do baixo peso ao nascer. *Rev. bras. crescimento desenvolv. hum.*, São Paulo, v. 18, n. 1, p. 98-103, abr. 2008. Disponível em <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0104-12822008000100013&lng=pt&nrm=iso>. acessos em 06 jun. 2020.
- CORDEIRO, G.M.; LIMA NETO, E.A. *Modelos Paramétricos*. São Paulo: Associação Brasileira de Estatística, 2004.
- COUTINHO LMS, SCAZUFCA M, MENEZES PR. Métodos para estimar razão de prevalência em estudos de corte transversal. *Rev Saúde Pública*. 2008 dez;42(6):992-8.
- CUMMINGS, Peter. Methods for estimating adjusted risk ratios. *The Stata Journal*, 2009, 9, Number 2, pp. 175–196.
- CUTLAND CL, LACKRITZ EM, MALLETT-MOORE T, ET AL. Low birth weight: Case definition & guidelines for data collection, analysis, and presentation of maternal immunization safety data. *Vaccine*. 2017;35(48 Pt A):6492-6500. doi:10.1016/j.vaccine.2017.01.049
- DOBSON, A. J., 1945- *An introduction to generalized linear models* / Annette J. Dobson.—2nd ed. p. cm.— (Chapman & Hall/CRC texts in statistical science series)
- FARAWAY, JULIAN J. (2006). *Extending the Linear Model With R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Boca Raton, FL: Chapman & Hall/CRC

FERRAZ TR, NEVES ET. Fatores de risco para baixo peso ao nascer em maternidades públicas: um estudo transversal. Rev Gaúcha Enferm., Porto Alegre (RS) 2011 mar;32(1):86-92.

GAÍVA, M. A. M., Lopes, F. S. P., Ferreira, S. M. B., & Mufato, L. F. (2018). Óbitos neonatais de recém-nascidos de baixo peso ao nascer. Revista Eletrônica De Enfermagem, 20. <https://doi.org/10.5216/ree.v20.47222>

HEERINGA, S. G.; WEST, B. T.; BERGLUND, P. A. Applied Survey Data Analysis. Chapman & Hall/CRC, 2010.

HIRAKATA, V. N. ESTUDOS TRANSVERSAIS E LONGITUDINAIS COM DESFECHOS BINÁRIOS: QUAL A MELHOR MEDIDA DE EFEITO A SER UTILIZADA?. Rev HCPA 2009;29(2):174-176.

HOSMER, D. W.; LEMESHOW, S. Applied Logistic Regression. 2 ed. New York, Wiley-Interscience, 2000.

Kelly; Edna Maria de Araújo; Djanilson Barbosa dos Santos. DESIGUALDADE ÉTNICO-RACIAL NA OCORRÊNCIA DA PREMATURIDADE NO INTERIOR DA BAHIA. In: ANAIS DO CONGRESSO BRASILEIRO DE SAÚDE COLETIVA, 2018, Rio de Janeiro. Anais eletrônicos... Campinas, Galoá, 2018. Disponível em: <<https://proceedings.science/saude-coletiva-2018/papers/desigualdade-etnico-racial-na-ocorrencia-da-prematuridade-no-interior-da-bahia>> Acesso em: 05 Maio. 2021.

KUTNER, MICHAEL H. Applied linear statistical models.-5th ed. Michael H Kutner ... [et al]. p. cm. - (McGraw-Hill/Irwin series Operations and decision sciences) Rev. ed. of: Applied linear regression models. 4th ed. c2004.

MAIA, Raquel da Rocha Paiva; SOUZA, José Maria Pacheco de. Fatores associados ao baixo peso ao nascer em município do norte do Brasil. Rev. bras. crescimento desenvolv. hum., São Paulo, v. 20, n. 3, p. 735-744, 2010.

MCNUTT, L. A., et al. Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes. American Journal of Epidemiology, Vol. 157, No. 10, DOI: 10.1093/aje/kwg074.

MARÔCO, J. Análise estatística com o PASW Statistics. Report Number, 2010.

MEDRONHO, R. A.; BLOCH, K. V.; LUIZ, R. R.; WERNECK, G. L. Epidemiologia. 2.ed. Rio de Janeiro: Atheneu; 2009.

MORAES, José Rodrigo de; MOREIRA, Jessica Pronestino de Lima; LUIZ, Ronir Raggio. Associação entre o estado de saúde autorreferido de adultos e a área de localização do domicílio: uma análise de regressão logística ordinal usando a PNAD 2008. Ciênc. saúde coletiva, Rio de Janeiro , v. 16, n. 9, p. 3769-3780, Sept. 2011 . Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232011001000013&lng=en&nrm=iso>. access on 13 May 2021. <https://doi.org/10.1590/S1413-81232011001000013>.

Moreira AI, Sousa PR, Sarno F. Baixo peso ao nascer e seus fatores associados. Einstein (São Paulo).2018;16(4):eAO4251.

http://dx.doi.org/10.31744/einstein_journal/2018AO4251

MOREIRA, Jessica Pronestino de Lima; MORAES, José Rodrigo de; LUIZ, Ronir Raggio. Utilização de consulta médica e hipertensão arterial sistêmica nas áreas urbanas e rurais do Brasil, segundo dados da PNAD 2008. Ciênc. saúde coletiva, Rio de Janeiro , v. 16, n. 9, p. 3781-3793, Sept. 2011 . Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232011001000014&lng=en&nrm=iso>. access on 13 May 2021. <https://doi.org/10.1590/S1413-81232011001000014>.

MINISTÉRIO DA SAÚDE. Secretaria de Vigilância em Saúde. Departamento de Vigilância de Doenças e Agravos não Transmissíveis e Promoção da Saúde. Saúde Brasil 2017: uma análise da situação de saúde e os desafios para o alcance dos objetivos de desenvolvimento sustentável [recurso eletrônico] / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância de Doenças e Agravos não Transmissíveis e Promoção da Saúde. Brasília: Ministério da Saúde, 2018.

MINISTÉRIO DA SAÚDE. Sistemas de Informações sobre Mortalidade (SIM) e Nascidos Vivos (Sinasc) para os profissionais do Programa Saúde da Família / Ministério da Saúde. – 2. ed. rev. atual. – Brasília: Ministério da Saúde, 2004;

MINISTÉRIO DA SAÚDE (BR). Secretaria de Vigilância em Saúde. Vigilância em saúde no Brasil 2003|2019: da criação da Secretaria de Vigilância em Saúde aos dias

atuais. Bol Epidemiol [Internet]. 2019 [data da citação]; 50(n.esp.):1-154. Disponível em: <http://www.saude.gov.br/boletins-epidemiologicos>

NASCIMENTO, R. C., BARBOSA, M. C. R., CORRÊA, M. M. (2019) Baixo-peso ao nascer: estudo de fatores associados em um hospital terciário da grande Vitória, ES, Brasil. DEMETRA, Rio de Janeiro, v.14: e43508, out-2019, 1-18, DOI: 10.12957/demetra.2019.43508.

PEDRAZA D. F. Qualidade do Sistema de Informações sobre Nascidos Vivos (Sinasc): análise crítica da literatura. Ciência & Saúde Coletiva, v. 17, p. 2729-2737, 2012.

PEREIRA, Veronica Aparecida et al. Desenvolvimento de bebês nos dois primeiros meses de vida: variáveis maternas e sociodemográficas. Pensando Famílias, v. 18, n. 1, p. 64-77, 2014. Disponível em: <<http://hdl.handle.net/11449/133834>>.

POWERS, D.; XIE, Y. Statistical Methods for Categorical Data Analysis. San Diego: Academic Press, 1999.

RENCHEER, A. C., and G. B. SCHAALJE. 2008. Linear models in statistics, 2nd ed. Hoboken, NJ: John Wiley & Sons.

SANTOS, Graciete Helena Nascimento dos; MARTINS, Marília da Glória and SOUSA, Márcia da Silva. Gravidez na adolescência e fatores associados com baixo peso ao nascer. Rev. Bras. Ginecol. Obstet. [online]. 2008, vol.30, n.5, pp.224-231.

Santos LPS, Bastos Neto BC, Pimenta RMC, Miranda SS. Análise do papel da raça/cor da pele no prognóstico do câncer de cavidade oral e orofaringe. J Dent Pub H. 2018;9(2):170-179.

SCHARDOSIM, Juliana Machado. A influência do ambiente e das práticas empregadas no trabalho de parto e parto no bem-estar do recém-nascido. 2018. 160 f., il. Tese (Doutorado em Saúde Coletiva) -Universidade de Brasília, Brasília, 2018.

SECRETARIA DE VIGILÂNCIA EM SAÚDE, Análise de Saúde e Vigilância de Doenças Não Transmissíveis, 2021. Painel de Monitoramento de Nascidos Vivos. Disponível em <<http://svs.aids.gov.br/dantps/centrais-de-conteudos/paineis-de-monitoramento/natalidade/nascidos-vivos/?s=MSQyMDE5JDlKMiQzNSQzJDEkMSQwJDYwMDAkMCQwJDUkMjEkM>>

CQwJDAkMCQwJDAkMCQwJDAkMCQwJDAkMCQw>. Acesso em: 03 de maio de 2021.

SECRETARIA DE VIGILÂNCIA EM SAÚDE, Análise de Saúde e Vigilância de Doenças Não Transmissíveis, 2021. Documentação. Disponível em <<http://svs.aids.gov.br/dantps/cgiae/sinasc/documentacao/>>. Acesso em: 03 de maio de 2021.

SILVESTRIN, Sonia; BURIOL, Viviane Costa de Souza; SILVA, Clécio Homrich da GOLDANI, Marcelo Zubaran. Avaliação da incompletude da variável escolaridade materna nos registros das Declarações de Nascidos Vivos nas capitais brasileiras - 1996 a 2013. *Cad. Saúde Pública*, Rio de Janeiro, v. 34, n. 2, 2018.

STEVANATO, Jaqueline Monte; GAIVA, Maria Aparecida Munhoz; SILVA, Ageo Mário Cândido da. Tendência da cobertura do Sistema de Informações sobre Nascidos Vivos em Mato Grosso, 2000 a 2012. *Epidemiol. Serv. Saúde*, Brasília, v. 26, n. 2, p. 265-274, 2017.

SZWARCWALD, Célia Landmann; LEAL, Maria do Carmo; ESTEVES-PEREIRA, Ana Paula; ALMEIDA, Wanessa da Silva de; FRIAS, Paulo Germano de; DAMACENA, Giseli Nogueira, SOUZA JÚNIOR, Paulo Roberto Borges de; ROCHA, Narayani Martins; MULLACHERY, Priscila Melissa Honorato. Avaliação das informações do Sistema de Informações sobre Nascidos Vivos (SINASC), Brasil. *Cad. Saúde Pública*, Rio de Janeiro, v. 35, n. 10, 2019

TIAN M, REICHETZEDER C, LI J, HOCHER B. Low birth weight, a risk factor for diseases in later life, is a surrogate of insulin resistance at birth. *J Hypertens*. 2019;37

TOLIO, F. B., MORAES, A. B., JACOBI, L. F. (2020) IDENTIFICAÇÃO DE FATORES DE RISCO PARA O BAIXO PESO AO NASCER – UMA ANÁLISE DE REGRESSÃO LOGÍSTICA. *Ciência e Natura*, Santa Maria, v. 42, Commemorative Edition: Statistic, e23, 2020 DOI: 10.5902/2179460X40497, ISSN 2179-460X

UCHIMURA, T. T., PELISSARI, D. M., SOARES, D. F. P. de P., UCHIMURA, N. S., SANTANA, R. G., & MORAES, C. M. S. (2008). Fatores de risco para o baixo peso ao nascer segundo as variáveis da mãe e do recém-nascido, em Maringá - PR, no período de 1996 a 2002. *Ciência, Cuidado E Saúde*, 6(1), 51. <https://doi.org/10.4025/ciencucuidaude.v6i1.4973>

Vale, C. C. R. Associação entre índices de adequação do cuidado pré-natal e baixo peso ao nascer no estado do rio de janeiro, 2015-2016 – COPPE, UFRJ. Rio de Janeiro, p. 91. 2020.

WORLD HEALTH ORGANIZATION (WHO). Global nutrition targets 2025: low birth weight policy brief [Internet]. Geneva: WHO; 2014. Available from: https://apps.who.int/iris/bitstream/handle/10665/149020/WHO_?jsessionid=58F175021A5C4AF4891BF477AF8211CE?sequence=2

ZERBETO AB, CORTELO FM, FILHO EB. Association between gestational age and birth weight on the language development of Brazilian children: a systematic review. J Pediatr (Rio J). 2015;91:326---32

ZOU G. A modified Poisson regression approach to prospective studies with binary data. Am J Epidemiol. 2004 Apr;159(7):702-6.