



UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS EXATAS
DEPARTAMENTO DE ESTATÍSTICA
CURSO DE ESTATÍSTICA

Adriane Machado

Larissa Luciane Ferreira

**ANÁLISE DE SOBREVIDA DE CRIANÇAS DE MUITO BAIXO PESO
AO NASCER**

**CURITIBA
2019**



UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS EXATAS
DEPARTAMENTO DE ESTATÍSTICA
CURSO DE ESTATÍSTICA

Adriane Machado

Larissa Luciane Ferreira

ANÁLISE DE SOBREVIDA DE CRIANÇAS DE MUITO BAIXO PESO AO NASCER

Trabalho de Conclusão de Curso apresentado à disciplina Laboratório B do Curso de Estatística do Setor de Ciências Exatas da Universidade Federal do Paraná, como exigência parcial para obtenção do grau de Bacharel em Estatística.

Orientadora: Profa. Dra. Suely Ruiz Giolo

**CURITIBA
2019**

AGRADECIMENTOS

À Universidade Federal do Paraná, pela oportunidade de fazer o Curso de Estatística.

À nossa orientadora, Profa. Dra. Suely Ruiz Giolo, pelo suporte, correções e incentivos durante a elaboração deste trabalho.

À Profa. Dra. Stela Adami Vayego, que aceitou participar da banca deste trabalho.

Aos nossos familiares e amigos, que nos deram apoio e incentivo nas horas difíceis, de desânimo e cansaço.

A todos que direta ou indiretamente fizeram parte da nossa formação, o nosso muito obrigada.

"A persistência é o menor caminho para o êxito".

(Charles Chaplin)

RESUMO

O peso ao nascer é um indicador de grande importância no que se refere à saúde do recém-nascido, visto que o baixo peso (inferior a 2,5 kg) está associado à maior mortalidade e morbidade neonatal. O peso ao nascer é subdividido em categorias, sendo a de interesse neste trabalho a categoria muito baixo peso (MBP), onde se encontram os neonatos que apresentam peso inferior a 1,5 kg. No Brasil, segundo o Sistema de Informação sobre Nascidos Vivos (SINASC), foi registrado, apenas no ano de 2016, um total de 37.374 bebês nascidos com MBP em hospitais públicos. Diante deste cenário, identificar fatores associados à sobrevivência de recém-nascidos com MBP é de grande relevância. Com esse objetivo, o presente trabalho apresenta um estudo sobre crianças nascidas com MBP no centro médico da Universidade de Duke. Os dados compreendem uma série histórica de 7 anos (1981 a 1987), a qual contém informações sobre bebês nascidos com muito baixo peso (MBP) no centro médico citado, ou que foram transportados para este centro logo após o nascimento. No contexto de análise de sobrevivência, foram utilizados, para a análise exploratória, o estimador de Kaplan-Meier e o teste *logrank*. Na sequência, foi ajustado o modelo de regressão de Cox, que não apresentou resultado satisfatório, pois a suposição de taxas de falha proporcionais não foi atendida para uma das covariáveis. Em virtude disso, optou-se pelo uso do modelo de Cox estratificado. Além da variável "oxigênio suplementar", utilizada no modelo como variável estratificadora, outras cinco covariáveis foram identificadas como fatores que influenciaram o tempo de sobrevivência dos bebês com MBP: Apgar no 1º minuto, hemorragia intraventricular, pneumotórax, ventilação assistida e tempo gestacional em semanas. A partir do modelo ajustado, foi constatado que o risco de óbito foi maior para bebês provenientes de gestações mais curtas, bem como para os bebês que necessitaram de ventilação assistida. Além disso, pôde-se apurar que o risco de óbito foi maior para os recém-nascidos que apresentaram hemorragia intraventricular, asfixia grave e presença de pneumotórax.

Palavras-chave: Bebês prematuros; Fatores de risco; Modelo de Cox Estratificado.

Sumário

AGRADECIMENTOS.....	iii
RESUMO	v
1 INTRODUÇÃO	7
2 MATERIAL E MÉTODOS	10
2.1 Material.....	10
2.1.1 Conjunto de Dados.....	10
2.1.2 Recursos Computacionais	11
2.2 Métodos	12
2.2.1 Conceitos Básicos de Análise de Sobrevivência	12
2.2.2 Estimador de Kaplan-Meier e Teste <i>logrank</i>	13
2.2.3 Modelo de Regressão de Cox.....	13
2.2.4 Seleção de Covariáveis	16
2.2.5 Avaliação da Qualidade de Ajuste do Modelo	17
3 RESULTADOS E DISCUSSÃO	18
3.1 Análise Exploratória	18
3.2 Resultados do Modelo de Cox.....	24
4 CONCLUSÕES.....	30
APÊNDICE A	34

1 INTRODUÇÃO

Ao tomar conhecimento de uma possível gravidez, grande parte das mulheres passa a tomar diversos cuidados na tentativa de reduzir problemas de saúde para o bebê. Dentre os problemas que podem ocorrer, está o baixo peso ao nascer. Existem duas categorias principais de baixo peso ao nascer. São elas: *i*) baixo peso, bebês nascidos com menos de 2,5 kg, e *ii*) muito baixo peso (MBP), bebês nascidos com menos de 1,5 kg (BECKWITH; RODNING, 1991, p. 25-58).

O peso ao nascimento é considerado um dos mais importantes indicadores da qualidade de vida da criança, por contribuir significativamente para a mortalidade infantil. Recém-nascidos prematuros e de baixo peso representam, entre a população neonatal, os grupos mais vulneráveis ao óbito (CARNEIRO et al., 2015).

Além disso, crianças nascidas com muito baixo peso constituem um grupo de alto risco para apresentar problemas de desenvolvimento psicológico. Crianças pré-termo com baixo peso ao nascer apresentam risco aumentado de sete a dez vezes de desenvolver problemas de paralisia cerebral, surdez e deficiência mental em relação às crianças a termo com peso acima de 2,5 kg (LINHARES et al., 2000).

As principais causas de MBP são o nascimento prematuro (nascimentos com menos de 37 semanas de gestação) e restrição de crescimento intrauterino, geralmente devido a problemas de placenta, saúde materna ou defeitos congênitos (INTENSIVE CARE NURSERY HOUSE STAFF MANUAL, 2004, p. 65-68).

No Brasil, segundo o Sistema de Informação sobre Nascidos Vivos (SINASC), órgão criado em 1990 pelo Ministério da Saúde com o objetivo de declarar informações públicas sobre nascidos vivos no Sistema Único de Saúde (SUS), foi registrado, apenas no ano de 2016, um total de 37.374 bebês nascidos em hospitais com peso inferior a 1,5 kg, sendo 33.731 nascidos com menos de 37 semanas de gestação, como pode ser observado na Tabela 1. Isso implica que 90,25% dos nascidos de muito baixo peso são prematuros.

Qualquer bebê nascido prematuramente apresenta maior probabilidade de nascer com muito baixo peso. Contudo, outros fatores podem aumentar esta probabilidade. Dentre eles, podem ser citados: a raça, mães adolescentes (idade < 15 anos), gravidez múltiplas, mulheres expostas às drogas, ao álcool e cigarros durante a gravidez, mulheres com menor nível socioeconômico e complicações durante a gravidez (INTENSIVE CARE NURSERY HOUSE STAFF MANUAL, 2004, p. 65-68).

Tabela 1 - Distribuição do total de crianças nascidas em hospitais no Brasil em 2016, por região e tempo de gestação em semanas

Região	Semanas de gestação						Total
	< 22	22 a 27	28 a 31	32 a 36	37 a 41	≥ 42	
1 - Região Norte	84	837	1.019	622	451	33	3.046
2 - Região Nordeste	356	2.844	3.340	1.894	1.384	82	9.900
3 - Região Sudeste	337	5.136	6.714	3.276	1.061	31	16.555
4 - Região Sul	67	1.618	1.988	1.027	307	11	5.018
5 - Região Centro-Oeste	88	923	1.011	550	274	9	2.855
Total	932	11.358	14.072	7.369	3.477	166	37.374

Fonte: MS/SVS/DASIS - Sistema de Informações sobre Nascidos Vivos – SINASC.

Outro fator de risco para o nascimento de crianças com muito baixo peso é a falta de acompanhamento pré-natal. Um estudo publicado por Carneiro et al. (2015) revelou que a taxa de prematuros foi de 9,2% entre as gestantes que realizaram pelo menos uma consulta de pré-natal e de 25,6% entre as que não realizaram consultas de pré-natal.

A área de neonatologia tem se desenvolvido muito nos últimos anos; um estudo recente publicado por Norman et al. (2019) revelou que a taxa de sobrevivência de bebês extremamente prematuros aumentou exponencialmente nos últimos 40 anos. Sendo assim, bebês nascidos cada vez mais cedo e, conseqüentemente, com menor peso, apresentam atualmente maiores chances de sobrevivência quando comparado com anos atrás. O estudo revelou, ainda, que o país líder em viabilidade neonatal precoce é a Suécia: no período de 2014 a 2016, cerca de 77% dos bebês nascidos entre 22 e 26 semanas de gestação sobreviveram por pelo menos um ano. Um dos motivos para esta taxa elevada se deve a qualidade dos hospitais, assim como ao suporte tecnologicamente avançado à vida neonatal no país. Porém, tais tecnologias não se encontram disponíveis em todos os países, havendo discrepâncias na qualidade do atendimento neonatal em diferentes partes do globo. Ainda assim, o aumento da taxa de sobrevivência de prematuros é uma notícia animadora, que mostra uma tendência de avanço positivo para as próximas décadas.

Dessa forma, estudos envolvendo recém-nascidos com muito baixo peso têm sido objeto de interesse crescente (LINHARES et al., 2000; PICOLLI et al., 2012). Um desses interesses, por exemplo, diz respeito aos diversos fatores que podem estar envolvidos com a sobrevivência ou não desses bebês. Nesta perspectiva, o foco deste trabalho se concentra

em avaliar os fatores que influenciaram a sobrevivência de bebês nascidos com muito baixo peso (entre os anos de 1981 e 1987) no centro médico da Universidade de Duke, Carolina do Norte (EUA), ou que foram transportados para este centro logo após o nascimento.

Em linhas gerais, o trabalho apresenta a seguinte estrutura. No Capítulo 2, são descritos o conjunto de dados e os métodos estatísticos utilizados para a análise dos mesmos. No Capítulo 3, são expostos os resultados obtidos com base nos métodos descritos no Capítulo 2. Por fim, o Capítulo 4 apresenta considerações finais do estudo realizado.

2 MATERIAL E MÉTODOS

2.1 Material

2.1.1 Conjunto de Dados

O conjunto de dados utilizado neste trabalho foi coletado no centro médico da Universidade de Duke, localizado no Estado da Carolina do Norte, Estados Unidos da América (EUA). Os dados compreendem uma série histórica de 7 anos (1981 a 1987), a qual contém informações sobre 671 bebês nascidos com muito baixo peso no centro médico citado, ou que foram transportados para este centro logo após o nascimento. Dos 671 bebês no estudo, 144 foram a óbito durante o período de tempo em que permaneceram internados no centro médico, enquanto os demais deixaram o hospital com vida. O conjunto de dados, que está disponível em <http://biostat.mc.vanderbilt.edu/wiki/Main/DataSets>, foi analisado por O'Shea et al. (1992) com o objetivo de estudar os efeitos de fatores pré-natais sobre o risco de hemorragia cerebral intraventricular subclínica em recém-nascidos de muito baixo peso ao nascer.

Diferente do estudo de O'Shea et al. (1992), o objetivo deste trabalho se concentra em avaliar os fatores que influenciaram a sobrevivência desse grupo de bebês. Para tanto, foram considerados os dados dos bebês que permaneceram internados no centro médico por um período de no máximo 150 dias, bem como os que não apresentaram informações inconsistentes e/ou falta de informação. Do total de 671 bebês no estudo, foi observado que 23 deles (o que equivale a 3,43%) apresentaram tempo de internamento superior a 150 dias (discrepante em relação à maioria dos bebês), bem como 62 apresentaram informações inconsistentes e/ou falta de informação. Após a exclusão dos dados desses 85 bebês, o conjunto de dados ficou composto de 586 recém-nascidos, 17,75% deles com registro de óbito e 82,25% com registro de alta hospitalar no período de 150 dias após o nascimento.

O Quadro 1 apresenta a descrição das variáveis (informações) disponíveis no conjunto de dados dos bebês nascidos em Duke. Além das informações descritas, tem-se também o tempo (em dias) decorrido desde o nascimento até o óbito ou até a alta hospitalar dos bebês. O par de informação (t_i, δ_i) , com t_i denotando os tempos e $\delta_i = 1$, se ocorreu o óbito e $\delta_i = 0$, se ocorreu a alta hospitalar, para $i = 1, \dots, 586$, corresponde à variável resposta considerada nas análises realizadas neste trabalho.

Para a realização das análises, optou-se por categorizar as variáveis "quantidade de plaquetas do bebê" e "Apgar no 1º minuto", apresentadas no Quadro 1, objetivando maior facilidade na interpretação dos resultados.

Quadro 1 – Informações na base de dados dos bebês nascidos em Duke com muito baixo peso

Variável	Descrição	Categorias
apg_cat	Apgar no 1º minuto	1 - Boa vitalidade 2 - Asfixia moderada 3 - Asfixia grave
bwt	Peso ao nascer	De 400g a 1500g
cld	Oxigênio suplementar administrado no bebê	1 - Sim 2 - Não
delivery	Tipo do parto	1 - Normal 2 - Cesárea
gest	Idade gestacional no nascimento do bebê	De 22 a 38 semanas
inout	Local do nascimento do bebê	1 - Duke 2 - Transportado
ivh	Bebê com hemorragia intraventricular	1 - Sim 2 - Não
magsulf	Tratamento materno com sulfato de magnésio	1 - Sim 2 - Não
meth	Mãe tratada com beta-methasona no pré-natal	1 - Sim 2 - Não
pda	Persistência do canal arterial (defeito cardíaco congênito detectado no bebê)	1 - Sim 2 - Não
pltct_cat	Quantidade de plaquetas do bebê	1 - Regular 2 - Alterado
pneumo	Ocorreu pneumotórax (os pulmões do bebê entraram em colapso após o nascimento)	1 - Sim 2 - Não
race	Raça do bebê	1 - Branca 2 - Negra
sex	Sexo do bebê	1 - Feminino 2 - Masculino
toc	Tocólise: mãe tratada com medicamento beta-adrenérgico no pré-natal	1 - Sim 2 - Não
twn	Nascimento de gêmeos ou único	1 - Sim 2 - Não
vent	Ventilação assistida realizada no bebê após o nascimento	1 - Sim 2 - Não
year	Ano de nascimento do bebê	de 1981 a 1987

Fonte: O'Shea et al. (1992).

2.1.2 Recursos Computacionais

Para as análises e tratamento do banco de dados foi utilizado o *software* livre R, versão 3.5.0 (R CORE TEAM, 2018). Tal *software* possibilita a preparação, análise e visualização gráfica de dados. O principal pacote utilizado para as análises foi o *survival*.

2.2 Métodos

2.2.1 Conceitos Básicos de Análise de Sobrevida

Em análise de sobrevivência, a variável resposta é, geralmente, o tempo até ocorrência de um evento de interesse. Esse tempo é denominado tempo de falha, podendo ser o tempo até a morte do paciente, bem como até a cura ou a recidiva de uma doença (COLOSIMO; GIOLO, 2006).

Em muitos estudos em que se deseja avaliar o tempo de sobrevivência, é comum a presença de censuras. Censuras são observações parciais da resposta, decorrente da não ocorrência do evento de interesse durante o período de acompanhamento. Nesses casos, os tempos registrados para os indivíduos que apresentaram resposta parcial são denominados tempos censurados. Considerar tais tempos nas análises é importante para evitar conclusões viesadas.

Para os dados descritos na Seção 2.1.1, o tempo de interesse corresponde ao tempo (em dias) decorrido desde o nascimento até o óbito de bebês nascidos com muito baixo peso entre os anos de 1981 e 1987 no centro médico da Universidade de Duke, ou que foram transportados para este centro logo após o nascimento. As censuras consistem dos bebês que não apresentaram óbito (ou seja, que receberam alta hospitalar).

A variável aleatória, não-negativa e usualmente contínua T , que denota o tempo de falha, ou seja, o tempo até a ocorrência do evento de interesse, pode ser especificada pela função de sobrevivência, expressa por

$$S(t) = P(T > t),$$

que fornece a probabilidade do indivíduo não falhar até certo tempo t , ou seja, a probabilidade dele sobreviver ao tempo t . A variável T também pode ser especificada pela função taxa de falha, dada por

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}.$$

A função taxa de falha $\lambda(t)$ é bastante útil para descrever a distribuição do tempo de vida de pacientes. Ela descreve a forma em que a taxa instantânea de falha muda com o tempo (COLOSIMO; GIOLO, 2006).

2.2.2 Estimador de Kaplan-Meier e Teste *logrank*

Um passo importante de uma análise estatística consiste em explorar as informações contidas na base de dados, tornando-se imprescindível a realização de uma análise descritiva. No contexto de dados de sobrevivência, que, em geral, apresentam informações censuradas, as técnicas convencionais utilizadas para obtenção de medidas descritivas (tais como a média e a mediana, dentre outras) não são adequadas.

Para obtenção de tais medidas descritivas, foram propostos métodos que acomodam todas as informações, inclusive as censuras. Um desses métodos, o estimador não paramétrico de Kaplan-Meier (KAPLAN; MEIER, 1958), foi proposto com a finalidade de estimar a função de sobrevivência $S(t)$ e, a partir dela, obter estimativas para as medidas descritivas de interesse.

O estimador de Kaplan-Meier é definido por

$$\hat{S}(t) = \prod_{j:t_j < t} \left(1 - \frac{d_j}{n_j}\right),$$

em que $\hat{S}(t)$ corresponde à estimativa da probabilidade de sobreviver ao tempo t , d_j é o número de falhas no tempo t_j , e n_j é o número de indivíduos sob risco em t_j (ou seja, os indivíduos que não falharam e não censuraram até o tempo imediatamente anterior a t_j).

As principais propriedades do estimador de Kaplan-Meier são: *i*) é não viciado para amostras grandes; *ii*) é fracamente consistente; *iii*) converge assintoticamente para um processo gaussiano; e *iv*) é estimador de máxima verossimilhança de $S(t)$.

Neste trabalho, o estimador de Kaplan-Meier foi utilizado para avaliar a influência das variáveis (uma a uma), apresentadas no Quadro 1, sobre o tempo de sobrevivência dos bebês descritos na Seção 2.1.1. Para proceder a esta avaliação não apenas graficamente, mas também por meio de um teste estatístico, as curvas de sobrevivência associadas às respectivas categorias de cada uma das variáveis mostradas no Quadro 1 foram comparadas entre si por meio do teste *logrank* (MANTEL, 1966).

2.2.3 Modelo de Regressão de Cox

O modelo de regressão de Cox (COX, 1972) possibilita a análise do tempo até a ocorrência de determinado evento de interesse, levando-se em conta diversas covariáveis

simultaneamente. Para este modelo, as funções taxa de falha e de sobrevivência são dadas, respectivamente, por

$$\lambda(t | \mathbf{x}) = \lambda_0(t)\exp(\mathbf{x}'\boldsymbol{\beta})$$

e

$$S(t | \mathbf{x}) = [S_0(t)]^{\exp(\mathbf{x}'\boldsymbol{\beta})},$$

em que $S_0(t)$ corresponde à função de sobrevivência de base e $\lambda_0(t)$ à função taxa de falha de base, $\mathbf{x} = (x_1, \dots, x_p)$ denota o vetor de valores observados para as covariáveis em \mathbf{X} e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)$ o vetor de parâmetros desconhecidos associados às covariáveis.

O modelo de Cox também é denominado modelo de riscos proporcionais (ou de taxas de falha proporcionais) devido ao fato de a razão das taxas de falha de dois indivíduos quaisquer ser constante no tempo. Desse modo, para que o modelo seja considerado adequado aos dados, é necessário avaliar se a suposição de taxas de falha proporcionais está sendo atendida. Para tanto, existem diferentes métodos propostos na literatura. Neste trabalho, foram utilizados três métodos para avaliar a suposição mencionada. O primeiro deles corresponde a um método gráfico descritivo, que consiste em dividir os dados em m estratos de acordo com as categorias de uma dada covariável. Por exemplo, dividir os dados em dois estratos de acordo com a covariável sexo. Em seguida, deve-se estimar a função taxa de falha acumulada de base $\Lambda_0(t)$ para cada estrato j ($j = 1, \dots, m$). O estimador proposto por Breslow (1972) para estimar tal função é dado por

$$\hat{\Lambda}_0(t) = \sum_{j: t_j < t} \frac{d_j}{n_j},$$

em que d_j é o número de falhas em t_j e n_j o número de indivíduos sob risco em t_j . Em seguida, são obtidos, para cada covariável, os gráficos de $\log(\hat{\Lambda}_{0j}(t))$ versus t . Se a suposição for válida, as curvas devem apresentar diferenças aproximadamente constantes ao longo do tempo (COLOSIMO; GIOLO, 2006).

Outro método gráfico utilizado para avaliar a suposição de riscos proporcionais do modelo de Cox, faz uso dos resíduos padronizados de Schoenfeld (SCHOENFELD, 1982), denotados por s_{iq} , para $i = 1, \dots, d$ e $q = 1, \dots, p$, com d sendo o número de falhas e p o número de covariáveis. Grambsch e Therneau (1994) sugeriram o gráfico de $s_{iq} + \hat{\beta}_q$ versus t , que deve apresentar uma linha horizontal (reta com inclinação nula) para que haja evidências a favor da suposição de proporcionalidade dos riscos.

Além dos métodos gráficos mencionados, os quais muitas vezes são subjetivos, um terceiro método utilizado foi a obtenção do coeficiente de correlação de Pearson (ρ) entre os resíduos padronizados de Schoenfeld e os tempos t , para cada covariável. Valores de ρ próximos de zero mostram não haver evidências para a rejeição da suposição de riscos proporcionais (COLOSIMO; GIOLO, 2006).

Como nem sempre a suposição de riscos proporcionais é válida, foram propostas extensões do modelo de Cox. O modelo de Cox estratificado é considerado uma dessas extensões, sendo útil quando a suposição de riscos proporcionais apresenta violação para uma das covariáveis presentes no modelo. No modelo de Cox estratificado, os dados são estratificados pelas categorias da variável que violou a suposição, de modo que em cada estrato a suposição continue válida para as demais covariáveis. Por exemplo, as taxas de falha podem não ser proporcionais entre homens e mulheres, mas esta suposição pode valer no estrato formado somente por homens e naquele formado somente por mulheres (COLOSIMO; GIOLO, 2006). O modelo de Cox estratificado fica, então, expresso por

$$\lambda(t | \mathbf{x}_{ij}) = \lambda_{0j}(t) \exp(\mathbf{x}'_{ij} \boldsymbol{\beta}),$$

para $j = 1, \dots, m$ e $i = 1, \dots, n_j$, em que n_j é o número de observações no j -ésimo estrato.

A estratificação não cria complicações para a estimação do vetor de parâmetros $\boldsymbol{\beta}$, que é estimado utilizando o método da máxima verossimilhança parcial proposto por Cox (1975). Considerando que existam $k \leq n$ falhas distintas nos tempos $t_1 < t_2 < \dots < t_k$, a função de verossimilhança parcial é expressa por

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n \left(\frac{\exp(\mathbf{x}'_i \boldsymbol{\beta})}{\sum_{j \in R(t_i)} \exp(\mathbf{x}'_j \boldsymbol{\beta})} \right)^{\delta_i},$$

em que $R(t_i)$ denota o conjunto de indivíduos sob risco em t_i .

Considerando que há m estratos, tem-se associado a cada estrato j ($j = 1, \dots, m$) uma função de verossimilhança parcial similar à apresentada acima, sendo a estimação do vetor $\boldsymbol{\beta}$ baseada na soma dos logaritmos das funções de verossimilhança parciais, isto é, em

$$l(\boldsymbol{\beta}) = [l_1(\boldsymbol{\beta}) + \dots + l_m(\boldsymbol{\beta})],$$

com $l_j(\boldsymbol{\beta}) = \log(L_j(\boldsymbol{\beta}))$ obtida utilizando-se os dados dos indivíduos no j -ésimo estrato. As derivadas para $l(\boldsymbol{\beta})$ são encontradas por meio da soma das derivadas obtidas para cada estrato e, então, $l(\boldsymbol{\beta})$ é maximizada com respeito a $\boldsymbol{\beta}$ (COLOSIMO; GIOLO, 2006).

No modelo de Cox estratificado é necessário estimar, além dos coeficientes de regressão β , os componentes $\lambda_{0_j}(t)$, $j = 1, \dots, m$. Para tanto, Breslow (1972) propôs um estimador para a função taxa de falha acumulada de base $\Lambda_{0_j}(t)$, o qual é expresso por

$$\hat{\Lambda}_{0_j}(t) = \sum_{k: t_k < t} \frac{d_k}{\sum_{l \in R_k} \exp(x_l' \hat{\beta})}$$

para $j = 1, \dots, m$, em que d_k corresponde ao número de falhas em t_k e R_k ao conjunto de indivíduos sob risco em t_k .

2.2.4 Seleção de Covariáveis

No que se refere à seleção de covariáveis, métodos computacionais foram propostos para auxiliar nesse processo, dentre eles, os métodos denominados *backward*, *forward* e *stepwise* (CHARNET et al., 2008). O método *forward* inicia o procedimento de seleção sem nenhuma covariável no modelo, apenas o intercepto. Dessa forma, passa a adicionar uma covariável por vez, avaliando sua correlação com a variável resposta e o nível de significância α adotado.

Diferente do método *forward*, o método *backward* inicia o procedimento de seleção com todas as covariáveis no modelo. Seguindo raciocínio contrário ao do método *forward*, as covariáveis são retiradas do modelo uma por vez, também considerando o nível de significância α adotado.

Em contrapartida, o método *stepwise* pode ser visto como uma combinação dos dois métodos apresentados. Similar ao método *forward*, tal método inicia o procedimento de seleção sem nenhuma covariável no modelo. Nos passos iniciais, covariáveis são incluídas no modelo. Após a inclusão de pelo menos duas covariáveis no modelo, pode-se ter a inclusão ou exclusão de covariáveis no modelo (uma por vez, de acordo com a relevância de cada covariável para o modelo). Nesse método são fixados dois níveis de significância, um para a inclusão e outro para a exclusão de covariáveis do modelo.

Vale destacar, que o teste da razão de verossimilhanças (TRV), bem como o Critério de Informação de Akaike (AIC), são, em geral, utilizados para auxiliar no processo tanto de inclusão quanto de exclusão de covariáveis (uma a uma) do modelo.

2.2.5 Avaliação da Qualidade de Ajuste do Modelo

Além da suposição de proporcionalidade, há a necessidade de verificar outros aspectos relacionados com a adequação e a qualidade de ajuste do modelo. Para tanto, pode-se analisar diferentes resíduos, dependendo do aspecto que se deseja avaliar.

Os resíduos de Cox-Snell, por exemplo, podem ser utilizados para avaliar o ajuste global do modelo. Para os modelos de Cox e de Cox estratificado, eles são definidos, respectivamente, por

$$\hat{e}_i = \hat{\Lambda}(t_i | \mathbf{x}_i) = \hat{\Lambda}_0(t_i) \exp \left\{ \sum_{k=1}^p x_{ik} \hat{\beta}_k \right\}, \quad i = 1, \dots, n$$

e

$$\hat{e}_i = \hat{\Lambda}(t_i | \mathbf{x}_i) = \hat{\Lambda}_{0j}(t_i) \exp \left\{ \sum_{k=1}^p x_{ik} \hat{\beta}_k \right\}, \quad j = 1, \dots, m \text{ e } i = 1, \dots, n_j$$

em que $\hat{\Lambda}(\cdot)$ é a função taxa de falha acumulada obtida a partir do modelo ajustado.

Para que o modelo seja considerado adequado, os resíduos de Cox-Snell devem seguir a distribuição exponencial padrão (LAWLESS, 2011). Para verificar este fato, pode-se utilizar técnicas gráficas. Uma delas, é a construção do gráfico das probabilidades de sobrevivência dos resíduos, $\hat{S}(\hat{e}_i)$, obtidas via o estimador de Kaplan-Meier, versus essas mesmas probabilidades obtidas a partir do modelo exponencial padrão, que deve ser próxima a uma reta com inclinação 1 para que o modelo seja considerado adequado.

Por outro lado, os resíduos *martingal* são úteis para auxiliar a determinar a forma funcional de uma covariável (em geral contínua) sendo incluída no modelo de regressão. Além disso, pode ser realizada uma avaliação dos resíduos *deviance*, que auxiliam a examinar a acurácia do modelo para cada indivíduo sob estudo (por exemplo, verificar se há indivíduos atípicos que podem estar influenciando na qualidade do modelo ajustado).

Tanto para os resíduos *deviance* quanto para os resíduos *martingal*, os resíduos devem apresentar um comportamento aleatório em torno de zero para que o modelo seja considerado apropriado aos dados sob análise. Os resíduos *martingal* e *deviance* são expressos, para $i = 1, \dots, n$, respectivamente, por

$$\hat{m}_i = \delta_i - \hat{e}_i$$

e

$$\hat{d}_i = \text{sin}(\hat{m}_i) [-2(\hat{m}_i + \delta_i \log(\delta_i - \hat{m}_i))]^{1/2},$$

em que $\delta_i = 1$ se falha e 0 se censura, e \hat{e}_i corresponde ao i -ésimo resíduo de Cox-Snell.

3 RESULTADOS E DISCUSSÃO

3.1 Análise Exploratória

Em uma abordagem inicial, adotou-se a realização de uma análise exploratória do conjunto de dados descritos na Seção 2.1.1. A Tabela 2 mostra as frequências absolutas e os percentuais de bebês em cada categoria das covariáveis registradas no estudo, além dos percentuais de falha (óbito) e de censuras (alta hospitalar). As variáveis idade gestacional (*gest*) e peso ao nascer (*bwt*) foram categorizadas para proceder às análises iniciais.

Tabela 2 – Totais de bebês nas categorias de cada covariável, frequências absolutas e percentuais de falha (óbitos) e de censuras (alta hospitalar) dos 586 bebês nascidos em Duke

(continua)

Covariável	Categoria	Total	Censura (alta = 0)	Falha (óbito = 1)	Total %	Censura %	Falha %
apg_cat	Asfixia grave	200	137	63	34%	69%	32%
	Asfixia moderada	176	146	30	30%	83%	17%
	Boa vitalidade	210	199	11	36%	95%	5%
bwt_cat	Inferior a 1 kg	182	111	71	31%	61%	39%
	De 1 kg a 1,5 kg	404	371	33	69%	92%	8%
cld	Oxigênio suplem.: não	423	374	49	72%	88%	12%
	Oxigênio suplem.: sim	126	108	18	22%	86%	14%
	NI - Não informado	37	0	37	6%	0%	100%
delivery	Cesárea	283	239	44	48%	84%	16%
	Normal	303	243	60	52%	80%	20%
gest	De 22 a 27 semanas	153	85	68	26%	56%	44%
	De 28 a 31 semanas	341	307	34	58%	90%	10%
	De 32 semanas ou mais	92	90	2	16%	98%	2%
inout	Nascido em Duke	492	414	78	84%	84%	16%
	Transportado	94	68	26	16%	72%	28%
ivh	Ausência	394	345	49	67%	88%	12%
	Presença	65	31	34	11%	48%	52%
	NI - Não informado	127	106	21	22%	83%	17%
magsulf	Ausência	327	273	54	56%	83%	17%
	Presença	52	45	7	9%	87%	13%
	NI - Não informado	207	164	43	35%	79%	21%
meth	Ausência	287	228	59	49%	79%	21%
	Presença	237	209	28	40%	88%	12%
	NI - Não informado	62	45	17	11%	73%	27%
pda	Ausência	476	406	70	81%	85%	15%
	Presença	110	76	34	19%	69%	31%

Tabela 2 – Totais de bebês nas categorias de cada covariável, frequências absolutas e percentuais de falha (óbitos) e de censuras (alta hospitalar) dos 586 bebês nascidos em Duke

(conclusão)

Covariável	Categoria	Total	Censura (alta = 0)	Falha (óbito = 1)	Total %	Censura %	Falha %
pltct_cat	Alterado	488	405	83	83%	83%	17%
	Regular	63	57	6	11%	90%	10%
	NI - Não informado	35	20	15	6%	57%	43%
pneumo	Ausência	473	427	46	81%	90%	10%
	Presença	113	55	58	19%	49%	51%
race	Negra	327	261	66	56%	80%	20%
	Branca	235	201	34	40%	86%	14%
	NI - Não informado	24	20	4	4%	83%	17%
sex	Feminino	298	248	50	51%	83%	17%
	Masculino	288	234	54	49%	81%	19%
toc	Ausência	400	335	65	68%	84%	16%
	Presença	124	102	22	21%	82%	18%
	NI - Não informado	62	45	17	11%	73%	27%
twn	Gravidez única	462	381	81	79%	82%	18%
	Gravidez múltipla	124	101	23	21%	81%	19%
vent	Ausência	257	254	3	44%	99%	1%
	Presença	329	228	101	56%	69%	31%
year	1981	27	23	4	5%	85%	15%
	1982	83	72	11	14%	87%	13%
	1983	97	71	26	17%	73%	27%
	1984	115	91	24	20%	79%	21%
	1985	119	103	16	20%	87%	13%
	1986	103	88	15	18%	85%	15%
	1987	42	34	8	7%	81%	19%

Fonte: Os autores (2019).

A partir da Tabela 2, observa-se, para a covariável raça (*race*), que bebês negros apresentaram frequência superior à dos bebês brancos, 56% e 40%, respectivamente. Quanto ao tipo de parto (*delivery*), foram registrados percentuais não muito distintos para as categorias cesárea e normal, 48% e 52% respectivamente. Percentuais similares também foram observados para o sexo dos bebês, 51% feminino e 49% masculino. Para ventilação assistida, as categorias "sem" e "com" apresentaram, respectivamente, os percentuais 44% e 56%. Ainda, 40% das mães foram tratadas com o medicamento beta-methasona (*meth*) durante o pré-natal, enquanto 49% não.

A covariável Apgar no 1º minuto (*apg_cat*) refere-se a um teste realizado para avaliar a vitalidade do recém-nascido. O teste é denominado teste de Apgar, podendo ser

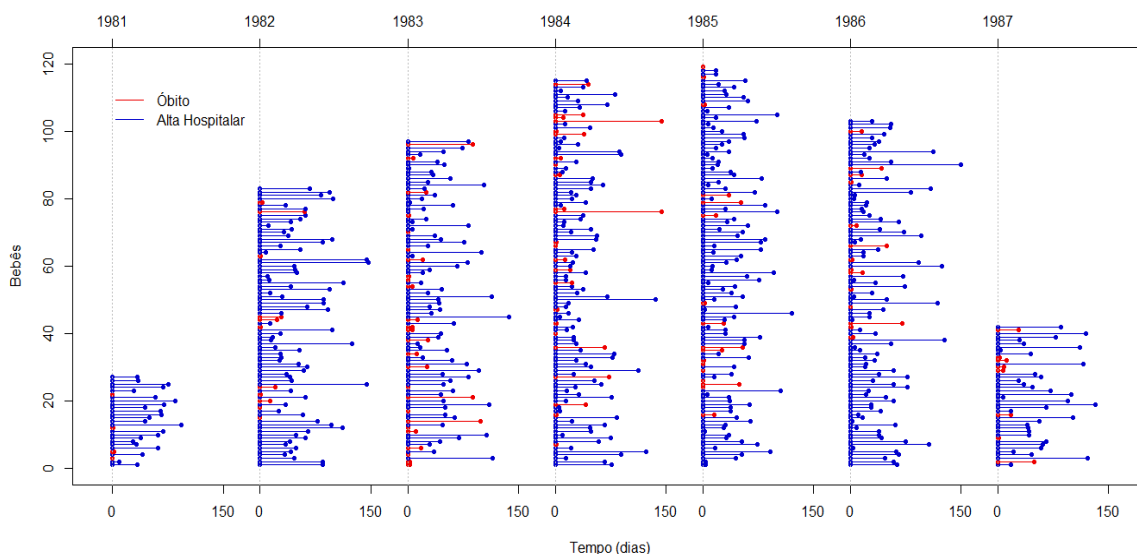
avaliado em vários momentos desde o nascimento do bebê. Nesse caso, foi avaliado no primeiro minuto de vida dos recém-nascidos. Em suma, o teste avalia características como frequência cardíaca, respiração, tônus muscular, reflexos e cor da pele do recém-nascido. Para cada característica é atribuída uma pontuação, totalizando o valor na escala Apgar. A escala Apgar foi desenvolvida pela Dra. Virgínia Apgar na década de 1950 apresentando basicamente três categorias: se o valor obtido estiver *a*) entre 0 e 3, é considerado asfixia grave; *b*) entre 4 e 6, asfixia moderada; e *c*) entre 7 e 10, boa vitalidade (APGAR, 1953). Para o estudo dos bebês de Duke, a Tabela 2 apresenta os seguintes percentuais para essas três categorias: 34%, 30% e 36%, nesta ordem.

As covariáveis idade gestacional (*gest_cat*), pulmões em colapso após o nascimento (*pneumo*), hemorragia intraventricular (*ivh*) e peso ao nascer (*bwt_cat*), se destacaram quanto ao percentual de óbitos para pelo menos uma de suas categorias. A idade gestacional destacou-se na categoria de 22 a 27 semanas de idade gestacional, que apresentou um total de 44% de falhas, ou seja, de bebês que apresentaram óbito. Já a covariável *pneumo*, que indica se os pulmões do bebê entraram em colapso após o nascimento (pneumotórax), se destacou na categoria que indica a ocorrência de pneumotórax no bebê, que apresentou 51% de óbitos. De maneira similar, a covariável que indica a presença de hemorragia intraventricular (*ivh*) apresentou um total de 52% de óbitos na presença dessa hemorragia. Por outro lado, a covariável *bwt_cat*, que indica se o bebê nasceu com peso extremamente baixo (inferior a 1 kg) ou com muito baixo peso (entre 1 kg e 1,5 kg), mostrou que a taxa de óbitos entre os bebês nascidos com menos de 1 kg foi de 39%, enquanto para os nascidos com peso entre 1 kg e 1,5 kg foi de 8%.

Quanto à quantidade de plaquetas dos recém-nascidos (*pltct_cat*), classificada nas categorias alterado e regular, com a faixa de valores de plaquetas de um recém-nascido entre 300.000 e 600.000/mm³ considerada normal (podendo variar conforme o laboratório), notou-se que o percentual de alta hospitalar entre os recém-nascidos na categoria de valores regular foi de 90%, ou seja, 90% dos bebês nesta categoria não apresentou óbito.

No que se refere à variável resposta de interesse, tempo (em dias) até a ocorrência do desfecho (óbito ou alta hospitalar), tem-se na Figura 1 a representação desses tempos para cada um dos 586 recém-nascidos por ano de nascimento (1981 a 1987). Esta figura mostra um percentual elevado de alta hospitalar (82,25% dos recém-nascidos deixaram o hospital com vida). Além disso, a Figura 1 evidencia que os bebês que permaneceram por mais de 30 dias internado formaram a maioria dentre os que receberam alta hospitalar (cerca de 67%), destacando-se alguns deles que ficaram 150 dias no hospital.

Figura 1 – Tempo até a ocorrência de óbito ou alta hospitalar por ano de nascimento dos bebês



Fonte: Os autores (2019).

A seguir, foi utilizado o estimador de Kaplan-Meier, para cada uma das covariáveis descritas na Tabela 2, com o objetivo de verificar a associação de cada uma delas com o tempo de sobrevivência dos bebês, auxiliando, assim, a identificar as covariáveis candidatas a entrar no modelo de Cox. A Figura 2 mostra as curvas de sobrevivência, estimadas por meio do estimador de Kaplan-Meier, associadas as oito covariáveis que mostraram indícios acentuados de diferenças entre suas respectivas categorias. São elas: Apgar no 1º minuto (*apg_cat*), peso no nascimento (*bwt_cat*), oxigênio suplementar (*cld*), idade gestacional (*gest_cat*), hemorragia intraventricular (*ivh*), quantidade de plaquetas do bebê (*pltc_cat*), pneumotórax (*pneumo*) e ventilação assistida (*vent*). As curvas de sobrevivência para as demais covariáveis encontram-se no Apêndice A.

Figura 2 – Curvas de sobrevivência estimadas por Kaplan-Meier para as covariáveis no estudo

(continua)

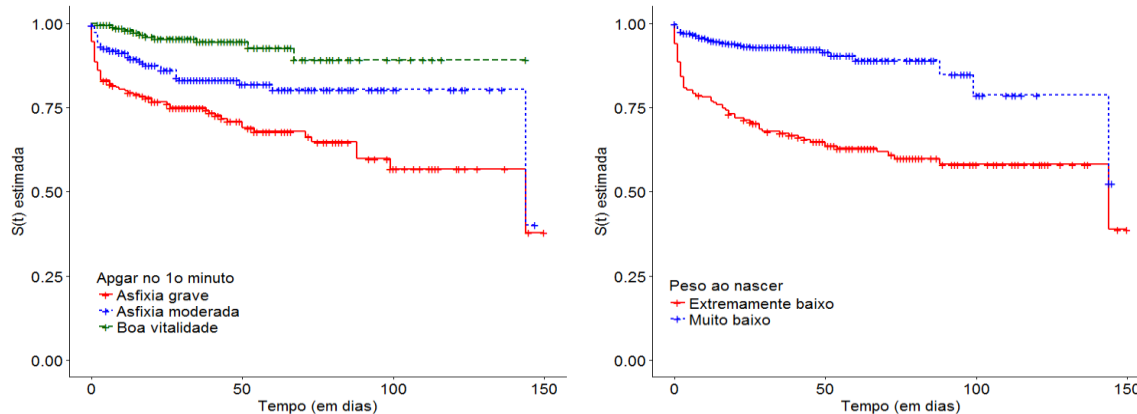
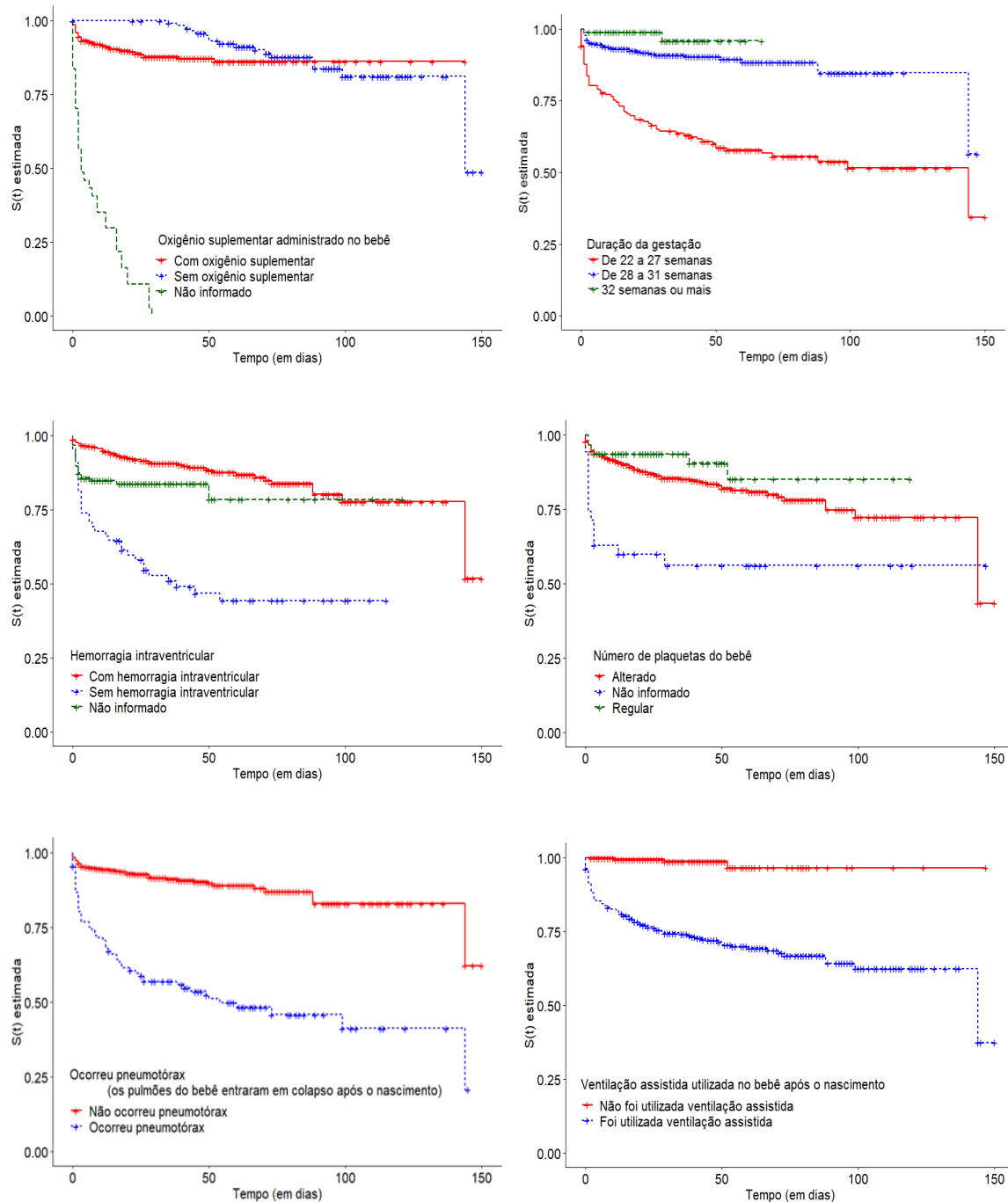


Figura 2 – Curvas de sobrevivência estimadas por Kaplan-Meier para as covariáveis no estudo
(conclusão)



Fonte: Os autores (2019).

Para as covariáveis mostradas na Figura 2, pode-se notar que há curvas distantes umas das outras, o que sugere que elas diferem entre si ao longo do tempo e, em consequência, que a probabilidade de óbito difere entre as categorias distintas de cada uma dessas covariáveis. Por exemplo, espera-se que bebês com a presença de pneumotórax apresentem probabilidade de óbito maior do que aqueles com ausência de pneumotórax.

A fim de testar se as diferenças observadas entre as curvas de sobrevivência são significativas, foi utilizado o teste *logrank*. Os resultados do teste, exibidos na Tabela 3, indicaram 13 covariáveis candidatas a entrar no modelo de Cox (valores $p \leq 0,10$). São elas: Apgar no 1º minuto (*apg_cat*), peso ao nascer (*bwt_cat*), oxigênio suplementar (*cld*), idade gestacional (*gest*), local de nascimento (*inout*), hemorragia intraventricular (*ivh*), mãe tratada com beta-methasona (*meth*), persistência do canal arterial (*pda*), quantidade de plaquetas (*pltct_cat*), pneumotórax (*pneumo*), tocolise (*toc*), ventilação assistida (*vent*) e ano do nascimento (*year*).

Tabela 3 – Resultados do teste *logrank* associado às covariáveis no estudo dos bebês de Duke

Covariável	Descrição	Estatística de Teste	Valor p
apg_cat	Apgar no 1º minuto	39,7	< 0,0001
bwt_cat	Peso ao nascer	498,0	< 0,0001
cld	Oxigênio suplementar administrado no bebê	378,0	< 0,0001
delivery	Tipo de parto	1,7	0,2
gest	Idade gestacional no nascimento	345,0	< 0,0001
inout	Local do nascimento	7,2	0,007
ivh	Hemorragia intraventricular	67,8	< 0,0001
magsulf	Mãe tratada com sulfato de magnésio	1,5	0,5
meth	Mãe tratada com beta-methasona	9,6	0,008
pda	Persistência do canal arterial	4,7	0,03
pltct_cat	Quantidade de plaquetas do bebê	19,0	< 0,0001
pneumo	Ocorrência de pneumotórax	99,7	< 0,0001
race	Raça do bebê	2,5	0,3
sex	Sexo do bebê	0,0	0,8
toc	Tocolise	4,9	0,09
twn	Nascimento de gêmeos ou único	0,4	0,5
vent	Ventilação assistida	63,9	< 0,0001
year	Ano de nascimento do bebê	10,0	0,1

Fonte: Os autores (2019).

Embora o estimador de Kaplan-Meier e o teste *logrank* sejam ferramentas úteis para a análise das covariáveis (uma a uma), eles não são adequados para avaliar o efeito conjunto de duas ou mais covariáveis. Diante de tal limitação, se faz necessário o uso de modelos de regressão, dentre eles, o modelo de Cox e o modelo de Cox estratificado, cujos resultados são apresentados a seguir.

3.2 Resultados do Modelo de Cox

Tendo em vista o modelo de Cox e o método de seleção de covariáveis *stepwise*, detalhados na Seção 2.2, foi selecionado o modelo com as covariáveis: idade gestacional, ventilação assistida, pneumotórax, Apgar no 1º minuto (três categorias), hemorragia intraventricular e oxigênio suplementar.

No processo de seleção de covariáveis notou-se a presença de correlação elevada entre as covariáveis peso ao nascer e idade gestacional. A justificativa é que a idade gestacional influencia no peso do bebê, pois quanto mais tempo durar a gestação mais nutrientes o bebê absorverá de sua mãe, como mencionado no Capítulo 1. Optou-se, assim, por manter a covariável tempo gestacional, por se tratar de uma variável com maior relevância para o estudo, visto que todos os bebês analisados são classificados como de muito baixo peso. Além disso, a covariável quantidade de plaquetas também apresentou correlação com outras covariáveis, sendo, por isso, excluída do modelo.

Para proceder o ajuste do modelo de Cox, as categorias de referência consideradas para as covariáveis categóricas foram as seguintes: categoria "não" para a covariável *cld*, categoria "ausência de ventilação assistida" para a covariável *vent*, categoria "asfixia grave" para a covariável *apg_cat*, categoria "ausência de hemorragia" para a covariável *ivh* e categoria "ausência de pneumotórax" para a covariável *pneumo*. A Tabela 4 apresenta as estimativas dos parâmetros, obtidas a partir do ajuste do modelo de Cox, para todas as covariáveis que apresentaram significância estatística.

Tabela 4 – Estimativas associadas ao modelo de Cox ajustado aos dados dos bebês de Duke

Covariável		Coefficiente	Exp(Coefficiente)	Valor p
apg_cat	Asfixia moderada	-0,3517	0,7035	0,1303
	Boa vitalidade	-0,9527	0,3857	0,0061
cld	Oxigênio: sim	-1,8712	0,1539	< 0,0001
	Não informado	0,9207	2,5112	0,0023
gest		-0,2993	0,7413	< 0,0001
ivh	Presença	0,4463	1,5624	0,0825
	Não informado	0,6429	1,9021	0,0219
pneumo	Sim	1,0021	2,7241	< 0,0001
vent	Sim	2,3562	10,551	< 0,0001

Fonte: Os autores (2019).

Conforme apresentado na Seção 2.2, é necessário avaliar a adequação do modelo de Cox mostrado na Tabela 4, bem como a validade da suposição de riscos proporcionais para este modelo. Quanto à suposição mencionada, o modelo apresentou violação para a

covariável oxigênio suplementar (*cld*), sendo possível observar na Tabela 5 valor relativamente elevado do coeficiente de correlação de Pearson (ρ), bem como valor p significativo ao nível de 5% para a categoria "sim" desta covariável.

Tabela 5 – Coeficiente de correlação de Pearson entre os resíduos padronizados de Schoenfeld e os tempos para o modelo de Cox ajustado aos dados de bebês com muito baixo peso

Covariável		ρ	Qui-quadrado	Valor p
apg_cat	Asfixia moderada	0,0639	0,4258	0,514
	Boa vitalidade	0,0547	0,3204	0,571
cld	Oxigênio suplementar: sim	0,5499	18,820	< 0,0001
	Não informado	0,1538	3,9985	0,045
gest		0,0237	0,0774	0,781
ivh	Presença	-0,0268	0,0884	0,766
	Não informado	-0,1324	1,7552	0,185
pneumo	Sim	0,0562	0,3684	0,544
vent	Sim	-0,1227	1,5201	0,218
GLOBAL		-	24,235	0,004

Fonte: Os autores (2019).

Como a suposição de riscos proporcionais foi violada para uma das covariáveis (oxigênio suplementar), foi ajustado o modelo de Cox estratificado pelas categorias desta covariável, resultando nas estimativas apresentadas na Tabela 6.

Tabela 6 – Estimativas associadas ao modelo de Cox estratificado ajustado aos dados dos bebês

Covariável		Coefficiente	Exp(Coefficiente)	Valor p
apg_cat	Asfixia moderada	-0,4971	0,6083	0,0416
	Boa vitalidade	-0,8685	0,4196	0,0131
gest		-0,2842	0,7526	< 0,0001
ivh	Presença	0,3811	1,4639	0,1396
	Não informado	0,7833	2,1887	0,0054
pneumo	Sim	1,0157	2,7614	< 0,0001
vent	Sim	2,7548	15,718	0,0002

Fonte: Os autores (2019).

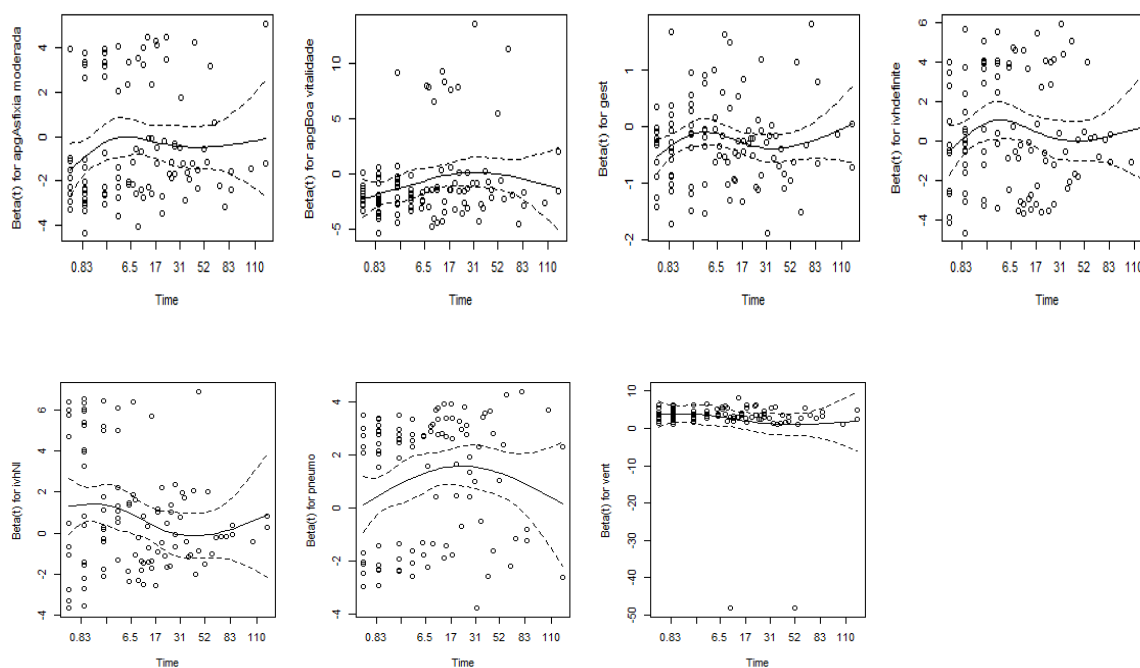
A fim de avaliar a adequação do modelo de Cox estratificado pela covariável oxigênio suplementar, nota-se, na Tabela 7, que os valores obtidos para os coeficientes de correlação de Pearson estão próximos de zero, indicando a não violação da suposição de riscos proporcionais para as demais cinco covariáveis. Além disso, foram analisados os gráficos dos resíduos padronizados de Schoenfeld (mostrados na Figura 3), que também indicaram evidências a favor da suposição de riscos proporcionais (inclinação \approx nula).

Tabela 7 – Coeficiente de correlação de Pearson entre os resíduos padronizados de Schoenfeld e os tempos para o modelo de Cox estratificado ajustado aos dados dos bebês de Duke

Covariável		ρ	Qui-quadrado	P-valor
apg_cat	Asfixia moderada	0,0593	0,3771	0,539
	Boa vitalidade	0,1047	1,1713	0,279
gest		0,0193	0,0528	0,818
ivh	Presença	-0,0135	0,0222	0,881
	Não informado	-0,1218	1,4611	0,227
pneumo	Sim	0,0185	0,0409	0,840
vent	Sim	-0,0965	0,9101	0,340
GLOBAL		---	4,0696	0,772

Fonte: Os autores (2019).

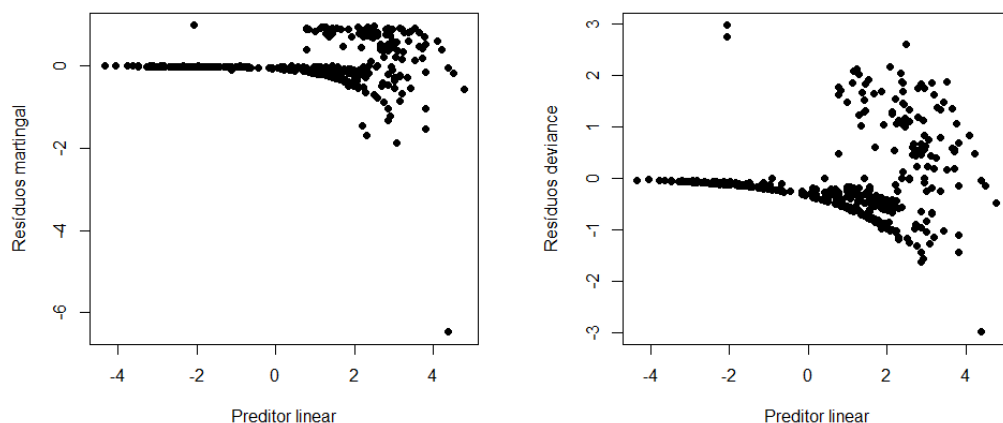
Figura 3 – Gráficos dos resíduos padronizados de Schoenfeld *versus* tempos para o modelo de Cox estratificado ajustado aos dados dos bebês de Duke



Fonte: Os autores (2019).

Por fim, outros gráficos foram analisados para avaliar a adequação do modelo ajustado. A partir dos gráficos dos resíduos *martingal* e *deviance*, mostrados na Figura 4, pode-se notar que os resíduos *deviance* não evidenciaram a presença de pontos atípicos, uma vez que grande parte dos resíduos se concentrou próximo de zero e entre -3 e 3.

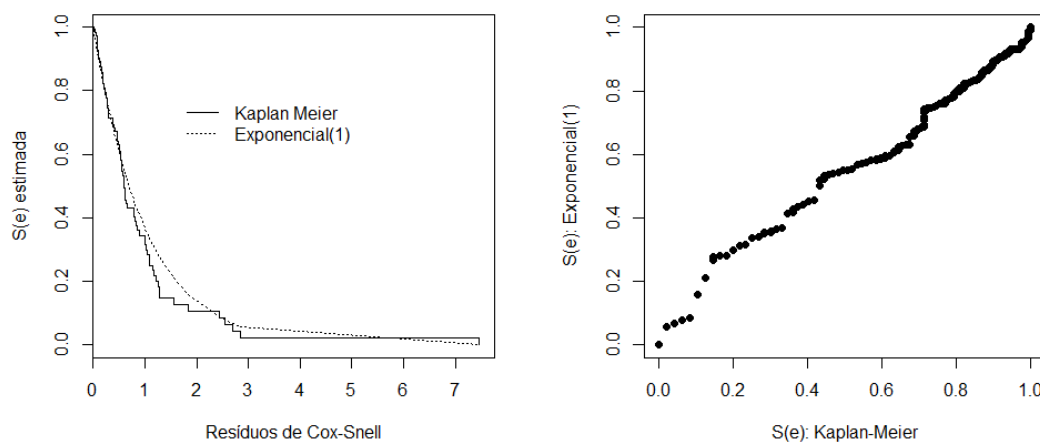
Figura 4 – Resíduos *martingal* e *deviance* associados ao modelo de Cox estratificado



Fonte: Os autores (2019).

Quanto aos resíduos de Cox-Snell associados ao modelo ajustado, nota-se a partir da Figura 5 que os mesmos seguem distribuição exponencial padrão, o que indica que o modelo apresentou um bom ajuste aos dados sob análise.

Figura 5 – Análise dos resíduos de Cox-Snell associados ao modelo de Cox estratificado



Fonte: Os autores (2019).

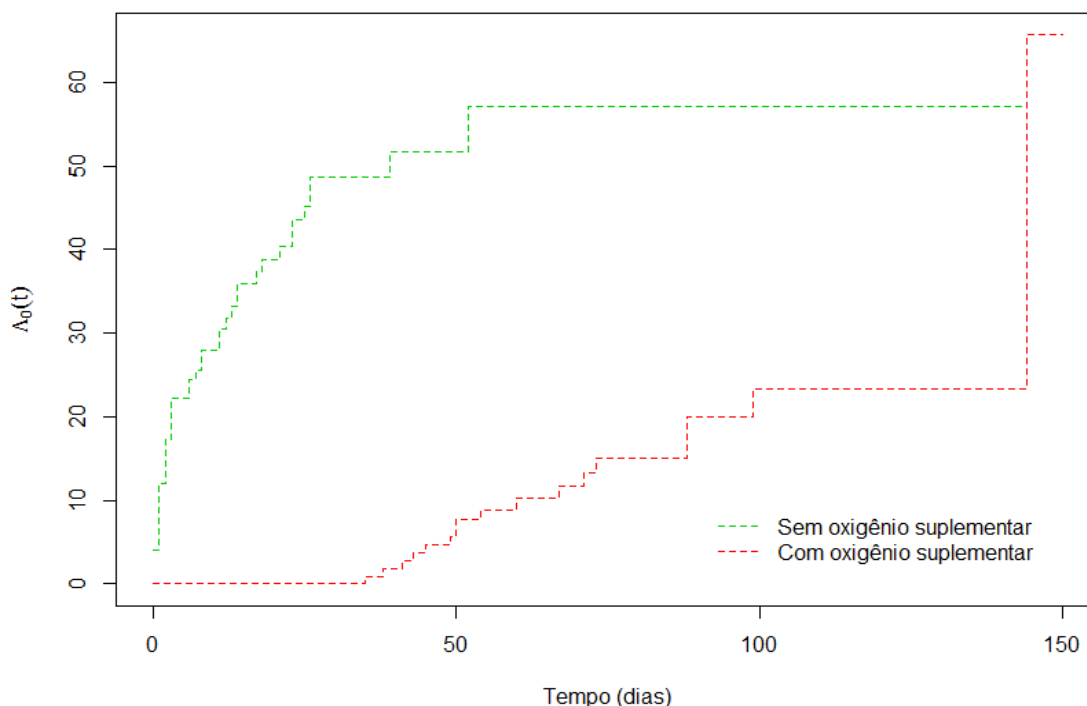
Uma vez que o modelo selecionado apresentou ajuste satisfatório, pode-se concluir, a partir das estimativas exibidas na Tabela 6, que o risco de óbito dos recém-nascidos que apresentaram pneumotórax foi cerca de 2,76 vezes o dos que não apresentaram a ocorrência desse distúrbio pulmonar. Notou-se, também, que o risco de óbito dos recém-nascidos que necessitaram de ventilação assistida foi cerca de 15,71 vezes o dos que não necessitaram de ventilação assistida. Ainda, para os recém-nascidos que apresentaram asfixia grave, o risco de óbito foi $\exp(0,8685) = 2,38$ vezes o dos com boa vitalidade, enquanto para os que

apresentaram asfixia moderada, o risco de óbito foi $\exp(-0,4971 + 0,8685) = 1,45$ vezes o dos recém-nascidos com boa vitalidade.

A duração da gestação também foi considerada um fator de risco para os recém-nascidos. Por exemplo, para as gestações em que a diferença de duração foi de 3 semanas, estima-se que o risco de óbito seja 57,37% menor entre os recém-nascidos que tiveram gestação com duração superior (3 semanas a mais). De modo análogo, para as gestações em que a diferença de duração foi de 5 semanas, espera-se que o risco de óbito seja 75,85% menor entre os recém-nascidos que tiveram gestação com duração de 5 semanas a mais. Logo, pode-se concluir que gestações com duração mais longas reduzem significativamente o risco de óbito do recém-nascido. Ademais, o risco estimado de óbito dos recém-nascidos que apresentaram hemorragia intraventricular foi cerca de 46,39% maior do que o dos recém-nascidos que não apresentaram essa complicação médica.

Quanto à covariável oxigênio suplementar administrado no bebê (utilizada como variável estratificadora no modelo de Cox estratificado), tem-se na Figura 6 as curvas de risco de óbito acumulado de base $\Lambda_{0j}(t)$, $j = 1, \dots, m$, estimadas para a duas principais categorias desta covariável: "sim" e "não" (optou-se por não apresentar a curva associada à categoria "não informado").

Figura 6 – Curvas do risco de óbito acumulado associadas à covariável oxigênio suplementar



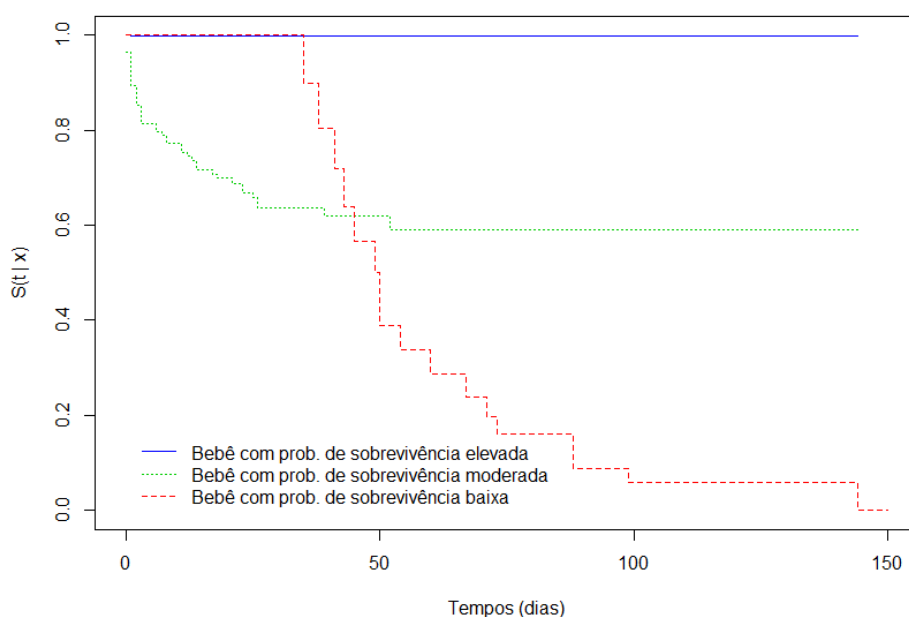
Fonte: Os autores (2019).

Com base nas curvas exibidas na Figura 6, fica evidente, para a maior parte do eixo dos tempos, que o risco de óbito foi maior para os recém-nascidos em que o oxigênio suplementar não foi administrado, com tal risco apresentando crescimento expressivo até por volta dos 55 dias, quando então se estabilizou. Por outro lado, pode-se notar que o risco de óbito cresceu de forma mais lenta até por volta dos 100 dias para os recém-nascidos em que o oxigênio suplementar foi administrado. Após 100 dias, o risco de óbito permaneceu estável por um período, e, então, voltou a crescer significativamente por volta dos 140 dias, tornando-se inclusive superior ao risco dos recém-nascidos em que o oxigênio suplementar não foi administrado. Vale ressaltar que a oferta ou não de oxigênio suplementar se baseia em protocolos de avaliação e monitorização contínua da função respiratória do recém-nascido nas primeiras horas de vida.

A título de ilustração do modelo de Cox estratificado ajustado aos dados dos bebês de Duke, são mostradas na Figura 7 as curvas de sobrevivência $\hat{S}(t | \mathbf{x})$ para três bebês. As características desses três bebês, que apresentaram, respectivamente, probabilidades de sobrevivência elevada, moderada e baixa, são as descritas a seguir.

Bebês	Oxigênio suplementar	Apgar no 1º minuto	Semanas de gestação	Hemorragia intraventricular	Pneumotórax	Ventilação assistida
1	não	boa vitalidade	32	não	não	não
2	não	asfixia moderada	28	não	sim	sim
3	sim	asfixia grave	22	sim	sim	sim

Figura 7 – Curvas de sobrevivência estimadas para três bebês nascidos em Duke



Fonte: Os autores (2019).

4 CONCLUSÕES

Neste trabalho, os dados de 586 bebês nascidos com muito baixo no centro médico da Universidade de Duke, Estados Unidos da América (EUA), foram analisados com o objetivo de identificar fatores associados à sobrevida desses bebês.

Inicialmente, foi realizada uma análise exploratória com o objetivo de conhecer melhor o comportamento das variáveis contidas na base de dados. Nesta análise inicial, as curvas de sobrevivência foram estimadas pelo método de Kaplan-Meier para todas as categorias das covariáveis, aplicando-se, a seguir, o teste *logrank* a fim de identificar se as curvas de sobrevivência entre as categorias de cada covariável diferiam entre si.

Na sequência, foi ajustado o modelo de Cox, que não apresentou ajuste satisfatório por não ter atendido à suposição de taxas de falha proporcionais para uma das covariáveis. Devido à violação citada, foi então ajustado o modelo de Cox estratificado, que apresentou ajuste satisfatório e identificou as seguintes covariáveis com efeito significativo sobre o tempo de sobrevida dos bebês: idade gestacional, ventilação assistida, pneumotórax, Apgar no 1º minuto, hemorragia intraventricular e oxigênio suplementar.

Embora o modelo de Cox estratificado tenha apresentado ajuste satisfatório para os dados dos bebês analisados, é importante mencionar que os resultados obtidos para esses dados não podem ser inferidos para toda a população de bebês nascidos com muito baixo peso, pois como mencionado no Capítulo 1, os resultados podem variar principalmente devido à tecnologia neonatal disponível em cada região.

REFERÊNCIAS

- ANDRADE C. L. T.; CASTILHO E. A.; SZWARCOWALD C. L. Baixo peso ao nascer no Brasil de acordo com as informações sobre nascidos vivos do Ministério da Saúde, 2005. **Cadernos de Saúde Pública**, v. 24, n. 11, p. 2564-2572, 2008.
- APGAR V. A proposal for a new method of evaluation of the newborn infant. **Curr Res Anesth Analg.**, v. 32, n. 4, p. 260-267, 1953.
- BECKWITH, L.; RODNING, C. **Intellectual functioning in children born preterm: recent research**. New Jersey: Lawrence Erlbaum, 1991.
- BRESLOW, N. E. Discussion of the paper by D. R. Cox. **Journal of the Royal Statistical Society, B**, v. 34, p. 216-217, 1972.
- CARNEIRO, J. A. et al. Características de recém-nascidos de muito baixo peso admitidos em unidade de terapia intensiva neonatal. **Revista de Enfermagem UFPE on line**, v. 9, n. 4, p. 7207-7212, 2015.
- CHARNET, R.; FREIRE, C. A. L.; CHARNET, E.; BONVINO, H. **Análise de modelos de regressão linear com aplicações**. Campinas: Editora da Unicamp, 2.ed., 2008. 356p.
- COLOSIMO, E. A.; GIOLO, S. R. **Análise de sobrevivência aplicada**. São Paulo: Blucher, 2006. 392 p.
- COX, D. R. Regression models and life tables. **Journal Royal Statistical Society, B**, v. 34, p. 187-220, 1972.
- COX, D. R. Partial Likelihood. **Biometrika**, v. 62, p. 269-276, 1975.
- GRAMBSCH, P. M; THERNEAU, T. M. Proportional hazards tests and diagnostics based on weighted residuals. **Biometrika**, v. 81, n. 3, p. 515-526, 1994.
- INTENSIVE CARE NURSERY HOUSE STAFF MANUAL. **Very Low and Extremely Low Birthweight Infants**. UCSF Children's Hospital, 2004. Disponível em: < <http://the-eye.eu/public/Books/BioMed/Intensive%20Care%20Nursery%20-%20House%20Staff%20Manual%208th%20ed%20%282003%29%20WW.pdf> > Acesso em: 08 jun. 2019.
- KAPLAN, E. L.; MEIER, P. Nonparametric estimation from incomplete observations. **Journal of the American Statistical Association**, v. 53, n. 282, p. 457-481, 1958.
- LAWLESS, J. F. **Statistical models and methods for lifetime data**. New York: John Wiley & Sons, 2011. 362 p.
- LINHARES M. B. M. et al. Prematuridade e muito baixo peso como fatores de risco ao desenvolvimento da criança. **Paidéia**, v.10, n. 18, p. 60-69, 2000.

MANTEL, N. Evaluation of survival data and two new rank order statistics arising in its consideration. **Cancer Chemotherapy Reports**, v. 50, p. 163-170, 1966.

NORMAN, M. et al. Association between year of birth and 1-year survival among extremely preterm infants in Sweden during 2004-2007 and 2014-2016. **Journal of the American Medical Association**, v. 321, n. 12, p. 1188-1199, 2019.

O'SHEA, M. et al. Prenatal events and the risk of subependymal/intraventricular haemorrhage in very low birth weight neonates. **Paediatric and Perinatal Epidemiology**, v. 6, p. 352-362, 1992.

PICCOLI, A. et al. Perfil clínico de neonatos de muito baixo peso internados em uma Unidade de Tratamento Intensivo Neonatal. **Rev. HCPA**, v. 32, n. 4, p. 412-419, 2012.

R CORE TEAM. **R: A language and environment for statistical computing**. Vienna, Austria, 2018. Disponível em: <<https://www.R-project.org/>>. Acesso em: 18 fev. 2019.

SCHOENFELD, D. Partial residuals for the proportional hazard regression model. **Biometrika**, v. 69, p. 239-241, 1982.

APÊNDICES

APÊNDICE A

Figura A1 – Curvas de sobrevivência estimadas pelo estimador de Kaplan-Meier

(continua)

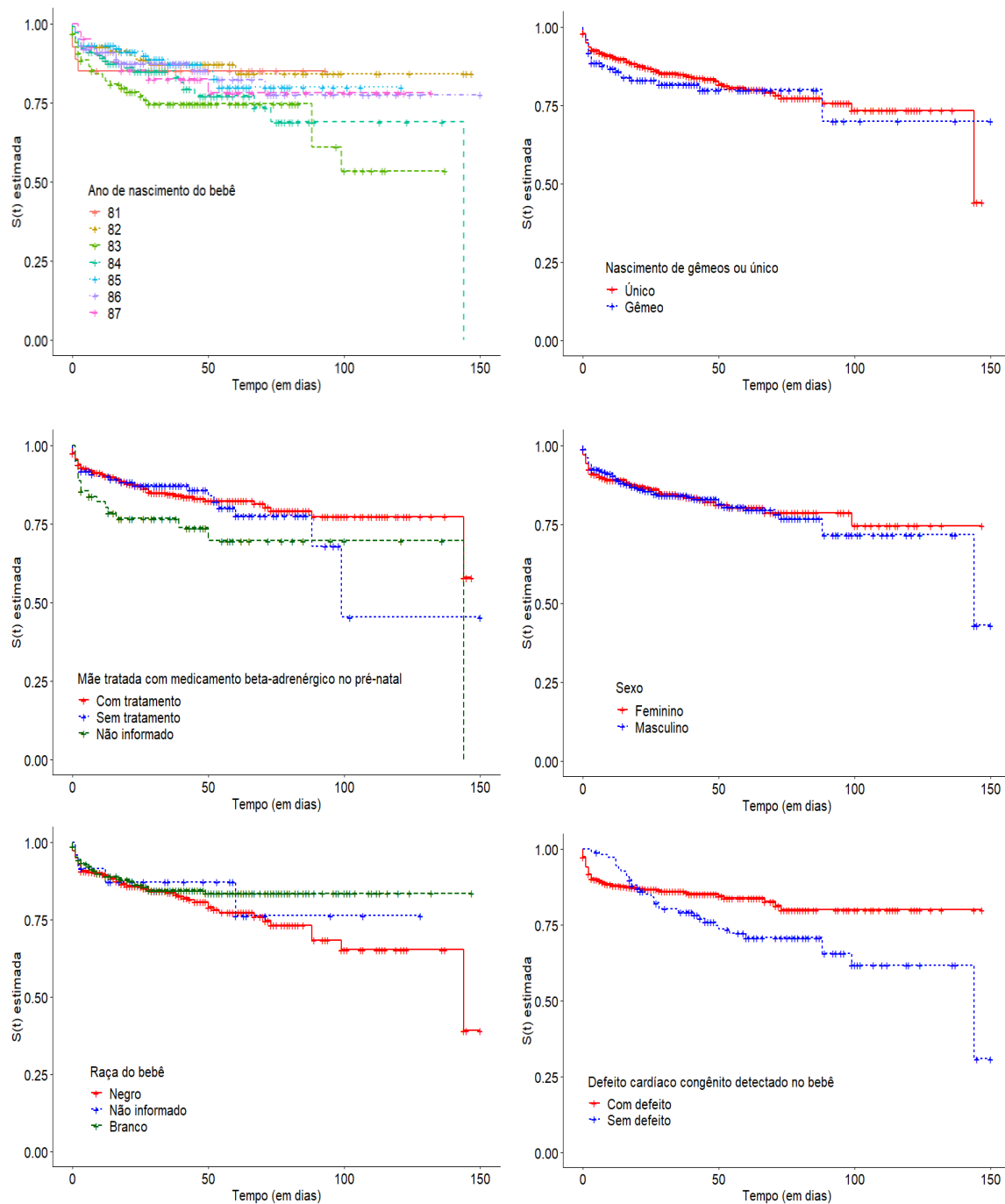
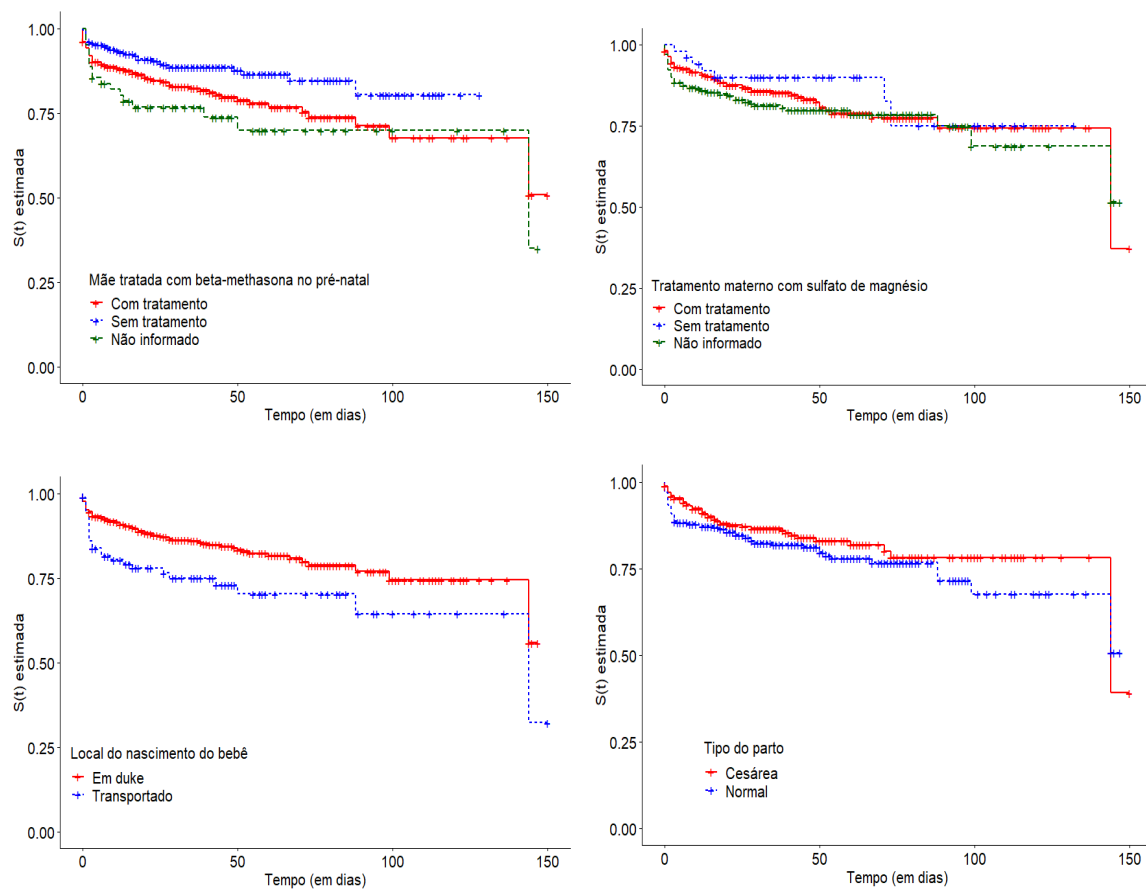


Figura A1 – Curvas de sobrevivência estimadas pelo estimador de Kaplan-Meier

(conclusão)



Fonte: Os autores (2019).