

Universidade Federal do Paraná

Ângela Luiza Cunha Legey

Leonardo de Marqui Marques

**Modelos de Riscos Competitivos com Variáveis  
Tempo Dependentes no Estudo do Impacto de  
Peritonites na Mortalidade Cardiovascular  
Tardia em Pacientes em Diálise Peritoneal**

**Curitiba**

**2017**

Ângela Luiza Cunha Legey  
Leonardo de Marqui Marques

**Modelos de Riscos Competitivos com Variáveis Tempo  
Dependentes no Estudo do Impacto de Peritonites na  
Mortalidade Cardiovascular Tardia em Pacientes em  
Diálise Peritoneal**

Trabalho de Conclusão de Curso apresentado  
à disciplina Laboratório B do Curso de Gra-  
duação em Estatística da Universidade Fede-  
ral do Paraná, como exigência parcial para  
obtenção do grau de Bacharel em Estatística.

Orientadora: Professora Dra. Silvia Emiko  
Shimakura

Curitiba  
2017

# Agradecimentos

Agradeço a Deus, por ser um Pai tão presente em todas as minhas conquistas, se hoje eu estou aqui é porque Ele nunca soltou da minha mão.

À minha família, pelo incentivo em todos momentos, pelo apoio nas horas difíceis, pelo conforto e segurança de uma família sem igual.

À orientadora, Profa. Dra. Silvia Emiko Shimakura, pelo incentivo e conhecimento transmitido. E a professora Dra. Suely Ruiz Giolo e ao professor Dr. Walmes Marques Zeviani pela disponibilidade em participar da banca deste trabalho.

As amizades feitas durante o curso, por de alguma forma fazerem parte dessa caminhada, pelas horas de estudo juntos e pelas risadas sem fim.

Aos meus amigos do Jesus na UFPR, pois sem vocês a universidade não teria sido a mesma. Vocês fizeram toda a diferença.

Ao parceiro de projeto Leonardo, pelo comprometimento, paciência e companheirismo durante todo esse período.

À todos que, mesmo não sendo citados aqui, tanto contribuíram para a conclusão desta etapa.

*Ângela Luiza Cunha Legey*

Agradeço a Deus, todo poderoso.

Aos bons professores, fonte de inspiração e motivação.

À família e aos amigos, refúgio e força.

À aleatoriedade da vida, também chamada de sorte. A mim mesmo.

*Leonardo de Marqui Marques*

*"E conhecereis a verdade, e a verdade vos libertará."  
(João 8:32)*

*"As boas ideias não têm idade, apenas têm futuro."  
(Robert Mallet)*

*"Todo o conhecimento é uma resposta a uma pergunta."  
(Thomas Henry Huxley)*

*"Uma boa vida é inspirada pelo amor e guiada pelo conhecimento."  
(Bertrand Russell)*

# Resumo

A diálise é o tratamento mais utilizado para pacientes com insuficiência renal. Em 2014, aproximadamente 50 mil brasileiros estavam em tratamento de diálise, dos quais 3400 faziam diálise peritoneal. A inflamação decorrente da diálise, a peritonite, tem surgido como um potencial fator para doença cardiovascular tardia. Com isso, surge a importância da verificação do impacto da peritonite na mortalidade por doença cardiovascular (DCV). Este trabalho apresenta um estudo de pacientes em diálise peritoneal tratados em 122 centros de diálise no Brasil. Os dados analisados foram coletados no período de seis anos, compreendendo de Dezembro de 2004 a Janeiro de 2011. Para a seleção das covariáveis, foi utilizado o método de seleção de Collett. Foram ajustados os modelos de riscos proporcionais de Cox e sua extensão tempo dependente; modelo multiestado e modelo de riscos competitivos. Para o modelo de Cox, as covariáveis: idade, diabetes, doença arterial coronariana e ter número de peritonites igual a três ou quatro foram identificadas como fatores que mais influenciaram na mortalidade por DCV. No modelo multiestado, as covariáveis: idade, diabetes e doença arterial coronariana foram as que mais contribuíram para a morte por DCV. O efeito de idade na transição da terceira peritonite para morte DCV foi o triplo, comparado com as transições partindo de nenhuma, uma e duas peritonites. No modelo de riscos competitivos, as covariáveis: sexo, idade, número de peritonites, diabetes, doença arterial coronariana e hemodiálise anterior foram significativas. O sexo masculino foi fator de proteção tanto no modelo multiestado quanto no modelo de riscos competitivos. A qualidade do ajuste para todos os modelos foi verificada através de análise de resíduos de Schoenfeld e esta foi considerada satisfatória.

**Palavras-chave:** Mortalidade cardiovascular; Diálise peritoneal; Variável tempo dependente; Modelo multiestado; Modelo de riscos competitivos

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>7</b>
<b>2</b>	<b>MATERIAL E MÉTODOS</b>	<b>9</b>
<b>2.1</b>	<b>Material</b>	<b>9</b>
2.1.1	Conjunto de Dados	9
2.1.2	Recursos Computacionais	10
<b>2.2</b>	<b>Métodos</b>	<b>10</b>
2.2.1	Funções de interesse	10
2.2.2	Estimador de Kaplan-Meier	11
2.2.3	Modelo de Regressão de Cox e suas extensões	11
2.2.3.1	Modelo de Cox	11
2.2.3.2	Modelo de Cox com variáveis tempo dependente	12
2.2.3.3	Modelo de Cox estratificado	12
2.2.3.4	Modelos Marginais	13
2.2.3.5	Método de seleção de covariáveis	13
2.2.3.5.1	Método de Collet	13
2.2.3.5.2	Teste de Wald	13
2.2.3.6	Adequação do Modelo de Cox	14
2.2.3.6.1	Resíduos de Schoenfeld	14
2.2.3.6.2	Resíduos Martingal e Deviance	15
2.2.3.6.3	Resíduos de Cox e Snell	15
2.2.3.6.4	Contrastes de Estimativas Kaplan-Meier vs Cox	16
2.2.4	Modelos Multiestado	17
2.2.5	Modelos de Riscos Competitivos	18
<b>3</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>22</b>
<b>3.1</b>	<b>Análise Descritiva</b>	<b>22</b>
<b>3.2</b>	<b>Modelo de Cox</b>	<b>24</b>
<b>3.3</b>	<b>Modelo de Cox Estendido</b>	<b>25</b>
<b>3.4</b>	<b>Modelo Multiestado</b>	<b>31</b>
3.4.1	Ajustes Marginais	32
3.4.2	Ajuste Conjunto	34
<b>3.5</b>	<b>Modelo de Riscos Competitivos</b>	<b>39</b>
<b>4</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>40</b>

REFERÊNCIAS . . . . . 42

APÊNDICES 44

# 1 Introdução

A insuficiência renal é uma das mais graves complicações que acontece em pacientes internados em unidades de terapia intensiva (UTI) (YU, 1995). Esta incidência varia de 10% a 25% dos casos e estes pacientes apresentam uma alta taxa de morbimortalidade. Apesar dos avanços tecnológicos, a mortalidade por insuficiência renal entre os pacientes com necessidades dialíticas fica em torno de 50% dos casos, podendo chegar a 90% (SILVA; THOMÉ, 2009).

A terapia para a substituição da função renal – a diálise – é o tratamento mais utilizado nos pacientes que possuem disfunção renal. Em 2014, aproximadamente 50 mil pacientes estavam em diálise nas 312 unidades brasileiras, sendo que 8.6% estavam em diálise peritoneal (DP) – sendo a diálise peritoneal automatizada (APD), a modalidade mais predominante (SESSO et al., 2016). A diálise peritoneal é um processo no qual um líquido especial, denominado dialisato, é injetado na cavidade peritoneal e extrai do sangue do paciente, por difusão e por osmose, as substâncias tóxicas ou em excesso que se encontram no organismo (BARBOSA et al., 2012). Dentro da diálise peritoneal, existem duas modalidades: a diálise peritoneal automatizada (APD), realizada à noite durante o sono do paciente, este ficando conectado a uma máquina cicladora automática que periodicamente substitui a solução de diálise por uma nova; e a diálise peritoneal ambulatorial contínua (CAPD), a qual permite a realização em domicílio pelo paciente e/ou responsável, sendo quatro trocas diárias, com tempo de permanência de 4 a 6 horas (SILVA; SILVA, 2006). A base fundamental desse tratamento é manter o paciente metabolicamente estável, sem problemas de uremia (acúmulo no sangue de substâncias tóxicas as quais são normalmente eliminadas pelos rins na urina). Entretanto, há uma preocupação para prevenção de complicações infecciosas, nutricionais, respiratórias, digestivas e cardiovasculares, para que haja uma total recuperação da função renal do paciente (SILVA; THOMÉ, 2009).

Dentre diversos fatores que podem explicar o risco de doenças cardiovasculares (DCV), o estado bem caracterizado e multifatorial de inflamação crônica sistêmica, observado em pacientes com doença renal crônica (DRC), tem emergido como um fator de risco em potencial para a DCV (BARBOSA et al., 2006). Essa inflamação tem sido associada a um aumento no risco de DCV em diferentes populações de pacientes (HONDA et al., 2006). O risco pode ser ainda maior quando os pacientes iniciam terapias de reposição renal, como a diálise peritoneal (CHO; JOHNSON, 2014).

A peritonite é uma inflamação do peritônio, uma membrana serosa que reveste parte da cavidade abdominal e recobre a maioria dos órgãos do abdômen. A peritonite é definida pela presença do dialisato (a solução purificadora) na cor turva, e acompanhada de febre e/ou dor abdominal (BARRETI; MONTELLI et al., 2001). A associação positiva



dos episódios de peritonite e mortalidade por doenças cardiovasculares (DCV) sugere uma ligação entre a inflamação aguda e desfechos cardiovasculares tardios (PECOITS-FILHO et al., 2016). Portanto há necessidade de estudos que verifiquem a relação entre inflamações relacionadas à diálise peritoneal e o risco de mortalidade tardia por DCV.

A abordagem mais intuitiva para problemas que envolvem saúde e modelagem de sobrevida (tempo até cura remissão da doença, dentre outras) é a análise de sobrevivência. Essa, como várias outras técnicas estatísticas, tem suas variações e extensões que foram desenvolvidas para atender às situações específicas de cada problema de forma direcionada a trazer resultados mais práticos e coerentes com a realidade. Uma abordagem é a análise de riscos competitivos. Esta técnica tem como objetivo modelar os tempos até a ocorrência de cada um dos possíveis desfechos competitivos/causas de falha. Os desfechos são pré-determinados pelos pesquisadores e os seus respectivos tempos são condicionados às covariáveis. Uma área pouco explorada em análise de sobrevivência é a junção de riscos competitivos com covariáveis mudando no tempo.

Quando se acompanha indivíduos por períodos relativamente longos é de se esperar que os valores de algumas covariáveis possam mudar ao longo do tempo. São exemplos: o tempo, os índices de colesterol, a glicose, a massa corpórea, os hábitos e, até mesmo, o tratamento de um indivíduo. O regime de funcionamento, bem como outras variáveis de uma máquina de diálise também podem ser alterados durante o período de acompanhamento. Logo, é importante levar em consideração essas mudanças e como elas incidem nas probabilidades de ocorrência de cada um dos possíveis desfechos. Tendo em mãos o modelo devidamente ajustado, é possível fazer inferências usuais, interpretações e a escolha de tratamentos adequados.

O objetivo deste trabalho foi revisar a literatura de análise de dados de sobrevivência no tratamento de covariáveis tempo-dependentes com riscos competitivos e aplicá-los nos dados de pacientes em diálise peritoneal. Também discutir a relevância da estruturação do banco de dados com variável tempo dependente e riscos competitivos e explorar uma área pouco estudada, que é a de variável tempo dependente e riscos competitivos.

Este trabalho está estruturado em três seções. Na Seção 2, serão descritos o conjunto de dados e a metodologia estatística utilizada para as análises. Na Seção 3, são apresentados os resultados obtidos a partir do modelo de regressão de Cox e de sua extensão do modelo multiestado e do modelo de riscos competitivos considerando a variável peritonite como tempo dependente. Na Seção 4, são apresentadas as considerações finais das análises realizadas neste trabalho.

## 2 Material e Métodos

### 2.1 Material

#### 2.1.1 Conjunto de Dados

Para as análises, utilizou-se um banco de dados financiado e coletado por *Baxter Healthcare Brasil*. O conjunto refere-se a uma coorte registrada a partir de Dezembro de 2004 a Janeiro de 2011. As informações são de pacientes adultos em diálise peritoneal com pelo menos 90 dias de tratamento em 122 centros de diálise no Brasil. Foram observados 5691 pacientes, contendo as variáveis descritas na Tabela 1.

Os pacientes desse estudo não sofreram transplante. Para estes pacientes, os possíveis desfechos são: morte (podendo ser por doença cardiovascular ou outras) ou sobreviver até o final do estudo. Todas as variáveis do conjunto de dados são ou foram categorizadas, com exceção da variável episódios de peritonites, que é quantitativa do tipo discreto.

Tabela 1 – Descrição das variáveis do conjunto de dados

Variável	Descrição	Categoria
Sexo	Sexo do paciente	Masculino Feminino
Idade	Idade do paciente	< 65 anos ≥ 65 anos
Escolaridade	Anos de estudo	< 4 anos ≥ 4 anos
Renda	Renda familiar	< 2 salários mínimos ≥ 2 salários mínimos
Doença cardiovascular	Possui doença cardiovascular tardia	Sim Outras
Peritonites	Número de episódios de peritonites	0 a 4
Distância até o centro	Distância da residência até o centro de diálise	< 50 km ≥ 50 km
Diabetes	Apresentava diabetes	Não Sim
Doença arterial coronariana	Apresentava doença arterial coronariana	Não Sim
Hipertensão	Apresentava hipertensão	Não Sim
Hemodiálise anterior	Realizou hemodiálise anterior ao estudo	Não Sim
Modalidade	Modalidade da terapia de diálise peritoneal	APD CAPD

APD: Diálise peritoneal automatizada

CAPD: Diálise peritoneal ambulatorial contínua

Fonte: Baxter Healthcare Brasil (2017).

## 2.1.2 Recursos Computacionais

Foi utilizado o *software* livre R, versão 3.3.3 (R CORE TEAM, 2016) de computação estatística para as análises e tratamento dos dados com o auxílio dos pacotes *survival* (THERNEAU, 2016), *mstate* (PUTTER; WREEDE; FIOCCO, 2016) e *cmprsk* (GRAY, 2014). Este programa permite a utilização de códigos abertos, proporcionando um amplo espaço para análises estatísticas e gráficas e o gerenciamento de banco de dados.

## 2.2 Métodos

### 2.2.1 Funções de interesse

A análise de sobrevivência consiste em métodos utilizados para a análise de dados em que a variável dependente (ou variável resposta) é o tempo  $T$  até a ocorrência de um evento de interesse ou mudança de estado. Este tempo é denominado tempo de falha (COLOSIMO; GIOLO, 2006).

Uma característica importante de dados de sobrevivência é a presença de censura, uma observação parcial da resposta que pode aparecer de várias tipos. A censura à direita sabe-se que o tempo entre o início da observação e o evento é maior do que o tempo de fato observado ( $T > t$ ), em que  $t$  é o tempo observado e  $T$  o tempo de ocorrência do evento. A censura à esquerda acontece quando não se conhece o momento da ocorrência do desfecho, porém se sabe que aconteceu antes do tempo observado (CARVALHO et al., 2011). Quando não se sabe o tempo exato do acontecimento do evento, porém se sabe que este ocorreu dentro de um intervalo conhecido de tempo dá-se o nome de censura intervalar.

Considere  $T$  uma variável aleatória que denota o tempo até a ocorrência de um evento, não negativo e usualmente contínuo. Para descrever a distribuição de  $T$ , umas das funções mais usuais é a função de sobrevivência, denotada por  $S(t)$ , que é definida como a probabilidade de uma observação não falhar até certo tempo  $t$ ,  $S(t) = P(T \geq t)$ .

Outra função utilizada é a de taxa instantânea de falha. Esta é bastante útil para descrever a distribuição do tempo de vida de pacientes. Assim, define-se a função de taxa de falha de  $T$ , conforme mostra a Equação (2.1)

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \quad (2.1)$$

A taxa de falha é definida como a probabilidade de que a falha ocorra no intervalo  $[t, t + \Delta t)$ , condicionado à sobrevivência até o tempo  $t$ , sendo dividida pelo comprimento do intervalo.

## 2.2.2 Estimador de Kaplan-Meier

O estimador de Kaplan-Meier (KAPLAN; MEIER, 1958) – também conhecido como estimador do produto-limite (PL) – é usado para estimar a função de sobrevivência  $S(t)$  quando os dados apresentam censuras. Para esse estimador, não é necessário assumir nenhuma distribuição de probabilidade para a variável aleatória  $T$ , sendo assim um estimador não paramétrico e certamente o mais utilizado. O estimador de Kaplan-Meier é definido por (Equação 2.2):

$$\hat{S}(t) = \prod_{j:t_j \leq t} \left( \frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j \leq t} \left( 1 - \frac{d_j}{n_j} \right), \quad (2.2)$$

em que  $t_1, \dots, t_k$  representam os  $k$  tempos distintos e ordenados de falha,  $d_j$  é o número de falhas em  $t_j$  e  $n_j$  é o número de indivíduos sob risco em  $t_j$ , para  $j = 1, \dots, k$ .

## 2.2.3 Modelo de Regressão de Cox e suas extensões

### 2.2.3.1 Modelo de Cox

Para a análise de dados de sobrevivência, o modelo mais frequentemente usado é o de regressão de Cox (COX, 1972). Este modelo permite a análise em que a resposta é o tempo de ocorrência de um evento de interesse sendo ajustado por covariáveis.

O modelo de Cox é definido pela Equação 2.3

$$\lambda(t|x) = \lambda_0(t) \exp(x'\beta), \quad (2.3)$$

em que a função de risco base  $\lambda_0(t)$  é multiplicada pela exponencial do vetor de covariáveis ( $x$ ), e pelo vetor de parâmetros das covariáveis ( $\beta$ ). Este modelo é composto pelo produto de dois componentes: não paramétrico ( $\lambda_0(t)$ ), que é uma função não negativa livre de distribuição e o paramétrico ( $\exp(x\beta)$ ). Essa particularidade faz com que este modelo seja mais flexível e, por este motivo, muito utilizado. Uma suposição básica deste modelo é que as taxas de falha são proporcionais (COLOSIMO; GIOLO, 2006). A razão das funções das taxas de falha para os indivíduos  $i$  e  $j$  é constante e dada por:

$$\frac{\lambda(t|x_i)}{\lambda(t|x_j)} = \frac{\lambda_0(t) \exp\{x'_i\beta\}}{\lambda_0(t) \exp\{x'_j\beta\}} = \exp\{(x'_i - x'_j)\beta\}. \quad (2.4)$$

A função de verossimilhança é dada pela Equação 2.5

$$L(\beta) = \prod_{i=1}^n \left( \frac{\exp(x'_i\beta)}{\sum_{j \in R(t_i)} \exp(x'_j\beta)} \right)^{\delta_i}, \quad (2.5)$$

em que  $R(t_i)$  é o conjunto dos índices das observações sob risco no tempo  $t$  e  $\delta_i$  é o indicador de falha.

### 2.2.3.2 Modelo de Cox com variáveis tempo dependente

Em situações, como quando os valores das covariáveis observadas durante o estudo cujos valores mudam nesse período, ou quando é violada a suposição de riscos proporcionais, não é adequado assumir o Modelo de Cox. Com isso, existem extensões do modelo, como o modelo com covariáveis dependentes do tempo, ou seja, valores das covariáveis podem mudar ao longo do tempo. Uma variável dependente do tempo é definida como qualquer variável cujo valor se altera ao longo do tempo ( $t$ ) (KLEINBAUM; KLEIN, 2006). O modelo é definido pela Equação 2.6

$$\lambda(t|x(t)) = \lambda_0(t) \exp(x'(t)\beta), \quad (2.6)$$

em que a função de risco base ( $\lambda_0(t)$ ) é multiplicada pela exponencial do vetor de covariáveis indexadas no tempo ( $x(t)$ ) e pelo vetor de parâmetros das covariáveis ( $\beta$ ).

Sendo assim, o modelo (2.6) não possui mais taxas proporcionais ao longo de todo o tempo  $t$ , pois a razão das funções das taxas de falhas no tempo  $t$  para dois indivíduos  $i$  e  $j$  é dado por:

$$\frac{\lambda(t|x_i(t))}{\lambda(t|x_j(t))} = \exp\{(\mathbf{x}'_i(t) - \mathbf{x}'_j(t))\beta\}, \quad (2.7)$$

que é dependente do tempo. Apenas nos intervalos de tempo onde não há mudança de covariáveis para nenhum dos dois indivíduos haverá proporcionalidade de taxas de falha. Sendo assim diz-se que há riscos proporcionalidade de riscos tempo dependentes.

As estimativas dos parâmetros do modelo de regressão de Cox com covariáveis tempo dependentes são obtidas estendendo-se o logaritmo da função de verossimilhança parcial dado por:

$$L(\beta) = \sum_{i=1}^n \delta_i \left[ x_i(t_i) - \frac{\sum_{j \in R(t_i)} x_j(t_i) \exp(x'_j(t_i)\hat{\beta})}{\sum_{j \in R(t_i)} \exp(x'_j(t_i)\hat{\beta})} \right] = 0, \quad (2.8)$$

em que pode variar no tempo  $t$  e poderiam ser mais denotadas por  $x(t)$ .

### 2.2.3.3 Modelo de Cox estratificado

Na análise estratificada, o banco de dados é dividido em  $m$  estratos e o modelo é expresso por (Equação 2.9)

$$\lambda(t|x_{ij}) = \lambda_{0j}(t) \exp(x'_{ij}\beta), \quad (2.9)$$

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

Constrói-se uma função de verossimilhança para cada estrato e a estimação é baseada na soma dos logaritmos das verossimilhanças parciais, ou seja, maximiza-se a função

$$\ell(\beta) = \sum_{j=1}^m \log(L_j(\beta)), \quad (2.10)$$

onde  $L_j(\beta)$  é calculado utilizando somente os dados dos indivíduos no  $j$ -ésimo estrato.

Nesse trabalho foi utilizado o modelo de Cox estratificado com covariável tempo dependente

#### 2.2.3.4 Modelos Marginais

Modelos marginais são utilizados com fins exploratórios para verificar o comportamento de cada covariável no modelo na ausência das demais covariáveis. Um modelo com cada covariável é ajustado e a qualidade de ajuste, violação de suposições, efeitos de covariáveis e demais medidas são avaliadas e comparadas com os demais modelos bem como com o modelo que contém todas as covariáveis. Com isso também pode-se optar um conjunto de modelos marginais ou um modelo conjunto.

#### 2.2.3.5 Método de seleção de covariáveis

##### 2.2.3.5.1 Método de Collet

Em razão do número de covariáveis disponíveis no banco de dados, foi utilizado o método de seleção de covariáveis proposta por Collett (2003). Segundo esse método, o passo inicial consiste em ajustar um modelo para cada uma das covariáveis, sendo que apenas as que forem significativas ao nível de 0,10 continuam para a segunda etapa. No segundo passo, as covariáveis que foram selecionadas no primeiro passo são ajustadas conjuntamente visto que algumas podem deixar de ser significativas na presença de outras devido à relação ou não independência entre elas. Se isso de fato ocorrer, são ajustados modelos reduzidos no qual retira-se uma covariável de cada vez. Cada etapa também é avaliada por meio do teste da razão de verossimilhanças, para medir a contribuição de cada covariável para a função de verossimilhança. No modelo, serão mantidas somente as covariáveis que alcançarem a significância estabelecida. Por fim, ajusta-se um novo modelo com as covariáveis contidas no segundo passo.

##### 2.2.3.5.2 Teste de Wald

Quando existe a necessidade de testar hipóteses referentes a um único parâmetro, o teste de Wald geralmente é usado. Este teste é obtido pela comparação da estimativa de máxima verossimilhança do parâmetro  $\widehat{\beta}_i$  e a estimativa de seu erro padrão. A razão resultante, sob a hipótese nula  $H_0 : \beta_i = \beta_{0i}$  tem distribuição normal padrão.

Portanto, a estatística do teste Wald é dada por

$$W = \frac{(\hat{\beta}_q - \beta_{0q})^2}{\widehat{\text{Var}}(\hat{\beta}_q)}.$$

,  $q = 1, 2, \dots, p$ , em que  $p$  é o número de covariáveis.

Os  $p$ -valores de testes paramétricos como esse são sensíveis ao tamanho da amostra. Assim sendo, em análises cujo banco de dados tem número elevado de observações, não se deve dar muito poder de evidência aos baixos  $p$ -valores.

### 2.2.3.6 Adequação do Modelo de Cox

#### 2.2.3.6.1 Resíduos de Schoenfeld

Apesar do modelo de Cox ser bastante flexível, é necessário verificar a adequação do modelo. Uma forma utilizada para avaliar a suposição de taxas de falha proporcionais são os resíduos de Schoenfeld (SCHOENFELD, 1982).

Considerando que o  $i$ -ésimo indivíduo com o vetor de covariáveis  $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$  irá falhar, então, para esse indivíduo o vetor de resíduos de Schoenfeld  $r_i = (r_{i1}, r_{i2}, \dots, r_{ip})$  em que cada componente  $r_{qi}$ , para  $q = 1, \dots, p$ , será definido por:

$$r_{qi} = x_{qi} - \frac{\sum_{j \in R(t_i)} x_{qj} \exp(x'_j \hat{\beta})}{\sum_{j \in R(t_i)} \exp(x'_j \hat{\beta})}. \quad (2.11)$$

Portanto, os resíduos são definidos para cada falha e não são definidos para as censuras. O resíduo de Schoenfeld é a diferença entre os valores observados das covariáveis de um indivíduo com tempo de ocorrência do evento  $t_i$  e os valores esperados em  $t_i$  dado o grupo de risco  $R(t_i)$ .

Foi utilizada a forma padronizada dos resíduos de Schoenfeld que é definida por:

$$s_i^* = [\vartheta(\hat{\beta})]^{-1} \times r_i,$$

sendo que  $\vartheta(\hat{\beta})$  a matriz de informação observada (THERNEAU; GRAMBSCH, 2000). Valores de coeficiente de correlação de Pearson entre os resíduos de Schoenfeld e o tempo, para cada variável, próximos de zero mostram não haver evidência contra a suposição de taxas de falha proporcionais.

Um teste de hipótese global de proporcionalidade das taxas de falha sobre todas as covariáveis no modelo, assumindo  $g_q(t) = g(t)$ , utiliza a estatística de teste apresentada na Equação 2.12

$$T = \frac{(g - \bar{g})' S^* \vartheta S'^* (g - \bar{g})}{d \Sigma_l (g_l - \bar{g})^2}, \quad (2.12)$$

em que  $\vartheta$  é a matriz de informação observada,  $d$  o número de falhas,  $S = dR\vartheta^{-1}$ ,  $R$  a matriz  $d \times p$  dos resíduos de Schoenfeld não padronizados e  $p$  o número de parâmetros.

Sob a hipótese nula de proporcionalidade das taxas de falha,  $T$  tem aproximadamente distribuição qui-quadrado com  $p$  graus de liberdade. Valores de  $T > \chi_{p,1-\alpha}^2$  mostram evidências contra a suposição de taxas de falha proporcionais.

Adicionalmente pode-se testar a hipótese de taxas de falha proporcionais para a  $q$ -ésima covariável,  $q = 1, \dots, p$ , com a estatística:

$$T_q = \frac{d(\Sigma_l(g_l - \bar{g})s_{ql}^*)^2}{\vartheta_q^{-1}\Sigma_l(g_l - \bar{g})^2}, \quad (2.13)$$

em que  $\vartheta_q^{-1}$  é o  $q$ -ésimo elemento da diagonal do inverso da matriz de informação observada. Sob a hipótese nula de proporcionalidade de taxas de falha,  $T_q$  tem aproximadamente distribuição qui-quadrado com um grau de liberdade. Valores de  $T_q > \chi_{1,1-\alpha}^2$  mostram evidências contra a suposição de proporcionalidade das taxas de falha para a  $q$ -ésima covariável.

#### 2.2.3.6.2 Resíduos Martingal e Deviance

Para verificar a presença de observações atípicas e também a forma funcional das covariáveis (linear, quadrática) são construídos gráficos dos resíduos *martingal* e *deviance* versus os tempos. Se o modelo for apropriado, os resíduos devem apresentar um comportamento aleatório em torno de zero. Os resíduos *martingal* e *deviance* são expressos para  $i = 1, \dots, n$ , respectivamente, por:

$$\hat{m}_i = \delta_i - \hat{e}_i \quad (2.14)$$

e

$$\hat{d}_i = \text{sin}(\hat{m}_i)[-2(\hat{m}_i + \delta_i \log(\delta_i - \hat{m}_i))]^{1/2}, \quad (2.15)$$

em que  $\delta_i$  é a variável indicadora de falha e  $\hat{e}_i$  o resíduo de Cox-Snell.

#### 2.2.3.6.3 Resíduos de Cox e Snell

Outra maneira de avaliar a qualidade do ajuste, é através da análise dos resíduos de Cox e Snell (COX; SNELL, 1968). Esses resíduos são definidos por:

$$\hat{e}_i = \hat{\Lambda}(t_i | \mathbf{x}_i) = \hat{\Lambda}_0(t_i) \exp \left\{ \sum_{k=1}^p x_{ik} \hat{\beta}_k \right\} \quad (2.16)$$

em que  $\hat{\Lambda}(\cdot)$  é a função taxa de falha acumulada obtida a partir do modelo ajustado. Para o modelo ser adequado, os resíduos  $\hat{e}_i$  devem seguir uma distribuição exponencial padrão (LAWLESS, 2011).



O uso de técnicas gráficas é comum para a verificação desse fato. Uma proposta é construir o gráfico das probabilidades de sobrevivência dos resíduos  $\hat{e}_i$ , obtidas a partir do estimador de Kaplan-Meier, versus as probabilidades de sobrevivência destes resíduos obtidas pelo modelo exponencial padrão. Este deve ser, aproximadamente, uma reta com inclinação 1 para que o modelo seja considerado adequado.

Estes resíduos devem ser usados com cautela pois a distribuição exponencial mantém-se somente quando os verdadeiros valores dos parâmetros são usados em 2.16 (KLEIN; MOESCHBERGER, 1997). Quando as estimativas dessas quantidades são usadas para o cálculo dos resíduos, as falhas quanto à distribuição exponencial podem ocorrer devido, parcialmente, à incerteza no processo de estimação dos parâmetros  $\beta$ . Essa incerteza é maior na cauda direita da distribuição 5 (onde ocorrem as censuras) e para amostras pequenas.

#### 2.2.3.6.4 Contrastes de Estimativas Kaplan-Meier vs Cox

O estimador de Kaplan-Meier não é paramétrico e tem como única suposição a aleatoriedade da censura. Se as estimativas oriundas de um método alternativo ficarem próximas às estimativas de Kaplan-Meier isso significa que o método de Kaplan-Meier está sendo bem representado pelo método alternativo. Uma forma gráfica de verificar a proximidade desses dois métodos é através da sobreposição das curvas de sobrevivências  $S_{KM}(t)$  (estimativas oriundas do método de Kaplan-Meier) e  $S_a(t)$  (estimativas oriundas do método alternativo) nos tempos  $t$ . Se as curvas ficarem próximas pode-se utilizar o método alternativo sem grandes perdas. Para diminuir a subjetividade dessa avaliação é possível calcular o coeficiente de correlação de Pearson ( $\rho$ ) entre as estimativas. Para esse caso, coeficiente  $\rho = 1$  atesta equivalência perfeita, então quanto mais próximo de 1 o valor de  $\rho$  mais bem representadas estão as estimativas de Kaplan-Meier pelas estimativas do método alternativo. Para facilitar essa verificação sugere-se traçar um gráfico dos pontos em que no eixo  $x$  estão as estimativas  $S_{KM}(t_i)$  e no eixo  $y$  as estimativas  $S_a(t_i)$ , em que  $i = 1, \dots, n_t$  e  $n_t$  é o número de tempos de falha distintos. A esse gráfico acrescenta-se a reta  $y = x$ . Quanto mais próximos os pontos  $(S_{KM}(t_i), S_a(t_i))$  estiverem da reta  $y = x$ , mais se evidencia a boa qualidade de ajuste do modelo alternativo.

## 2.2.4 Modelos Multiestado

Os modelos multiestado podem ser considerados uma generalização do modelo de sobrevivência clássica. Enquanto o modelo de sobrevivência clássico considera somente o tempo até um determinado evento, os modelos multiestado lidam com sistemas de eventos relacionados. Tais sistemas de eventos são descritos como um conjunto de estados e de transições entre os estados.

Neste contexto, os modelos multiestado são frequentemente usados para descrever dados longitudinais e definidos a partir de um processo estocástico em tempo contínuo permitindo que os indivíduos se movam através de um número finito de estados. A complexidade de um modelo multiestado depende do número de estados definidos e do número de transições permitidas entre eles (FERNANDES, 2010).

Os estados podem ser considerados como transientes ou absorventes (FERNANDES, 2010). Um estado absorvente não permite transições e quando um indivíduo entra nesse estado, permanece nele sem mudar para outro estado. A morte, seria um exemplo desse tipo de estado. Por outro lado, um estado transiente é quando é possível ocorrer transições a partir dele. Como exemplo, em diversos estudos de câncer, são observados estados intermediários tais como: "vivo com recorrência", "vivo com metástases", entre outros.

Os modelos multiestado facilitam a compreensão do progresso do fenômeno de interesse. Proporcionam uma utilização mais eficiente da informação incompleta, quando o histórico da doença de um indivíduo é conhecida. Esse possibilita modelar estados intermediários, que podem ou não preceder os estados absorventes (desfechos finais). Outra flexibilidade do modelo, é a possibilidade de introduzir variáveis tempo-dependentes como estados intermediários. Tais resultados mostram a flexibilidade e a conveniência do método tanto para predições quanto para interpretações dos efeitos das covariáveis.

Note que não é adequado utilizar o método de Kaplan-Meier quando o evento de interesse está na presença de outros eventos que, caso ocorram, impossibilitam a ocorrência do evento de interesse. Por exemplo: no estudo de insuficiência renal no qual o óbito é o evento de interesse, uma abordagem usual é considerar transplante renal como censura. No entanto, o conceito de censura afirma que, para o indivíduo em questão, o evento de interesse de fato ocorrerá. Se o paciente sofreu um transplante, o evento não vai ocorrer e o paciente não pode ser representado pelos que ainda continuam no estudo. Sendo assim, é importante considerar mais de uma causa de falha sempre que possível.

Seja  $Z(t)$  o estado do indivíduo no tempo  $t$ . O movimento em um espaço discreto de estados  $1, \dots, W$  é definido pelas intensidades de transição  $\lambda_{rs}(t, X(t)) : r, s = 1, \dots, W$  que são função do tempo  $t$  e do vetor de covariáveis que podem ser tempo dependentes,  $X(t)$ .

De acordo com Jackson et al. (2011), a intensidade representa o risco instantâneo de mover do estado  $r$  para o estado  $s$ ,  $r \neq s$

$$\lambda_{rs}(t, X(t)) = \lim_{\Delta t \rightarrow 0} \frac{P[Z(t + \Delta t) = s | Z(t) = r, X(t)]}{\Delta t}. \quad (2.17)$$

Dentre todos os possíveis métodos de estimação das intensidades de transição, neste trabalho foi utilizado modelo de Cox estratificado, em que cada possível transição foi tratada como um estrato. Para o ajuste do modelo multiestado, é preciso preparar o banco de dados de maneira conveniente. Sendo assim, é necessário inserir linhas que representem, não só a transição que ocorreu com cada indivíduo, mas também as transições que poderiam ter ocorrido devidamente identificadas como censura. Dessa forma a transição é usada como variável de estratificação.

No modelo multiestados em que se tem  $m$  transições possíveis, o risco da  $j$ -ésima transição para o  $i$ -ésimo indivíduo condicionada ao vetor de covariáveis é expresso por:

$$\lambda_j(t|x_i) = \lambda_{0j}(t) \exp(x'_i \beta) \quad (2.18)$$

onde  $j = 1, \dots, m$  e  $i = 1, \dots, n_j$  em que  $n_j$  é o número de observações na  $j$ -ésima transição.

### 2.2.5 Modelos de Riscos Competitivos

À situação em que se tem mais de um desfecho possível e a ocorrência de um impede a ocorrência do outro, dá-se o nome de Riscos Competitivos (KALBFLEISCH; PRENTICE, 2002). Os tempos de falha na presença de riscos competitivos são modelados através da Função de Incidência Acumulada (FIA) que é a soma dos riscos de causa específica acumulados. A estimação da FIA (Equação 2.19) para a  $k$ -ésima causa, no tempo  $t$  condicionada ao vetor de covariáveis tempodependentes  $x(t)$  é dada por:

$$FIA_k(t|x(t)) = \sum_{\forall j, t_j|x(t) \leq t} \lambda_k(t_j|x(t)) S(t_j|x(t)), \quad (2.19)$$

em que  $\lambda_k(t|x(t))$  são os riscos da  $k$ -ésima causa específica,  $k = 1, 2, \dots, K$ , e  $S(t|x(t))$  a sobrevivência global (PUTTER; FIOCCO; GESKUS, 2007).

O risco (Equação 2.20) da  $k$ -ésima causa, também chamado de do função de risco de subdistribuição é dada por:

$$\lambda_k(t|x(t)) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P[t \geq T \geq t + \Delta t, \varepsilon = k | T \geq t, x(t)]. \quad (2.20)$$

Fine e Gray (1999) propuseram um método de estimação dos riscos de causa específica sem o pressuposto de independência entre os tempos de falha e suas causas. O ponto chave é que a Função de Incidência Acumulada (FIA) é função de todos os riscos

de causa específica. Já a função do risco de subdistribuição tem relação biunívoca (um-para-um) com os riscos de causa específica, tornando possível analisar separadamente as causas de falha e seus riscos. Essa análise, aliada à estimação dos efeitos das covariáveis, mostra-se bem útil não só na área médica (onde é mais usada), mas também na industrial.

Pelo modelo de Fine e Gray (FINE; GRAY, 1999), o risco de subdistribuição da  $k$ -ésima causa,  $k = 1, 2, \dots, K$ , é dado por

$$\lambda_k(t|x(t)) = \lambda_{0k}(t) \exp(x(t)\beta_k), \quad (2.21)$$

em que  $\beta_k$  é o vetor de coeficientes das covariáveis para a  $k$ -ésima causa de falha e  $\lambda_{0k}(\cdot)$  é o risco de base da  $k$ -ésima causa de falha e é estimado separadamente para cada causa. Os coeficientes  $\beta_k$  são estimados maximizando-se a função de verossimilhança parcial e a técnica de ponderamento pela probabilidade inversa das censuras (ROBINS; ROTNITZKY, 1992), dada por

$$L(\beta_k) = \prod_{i=1}^n \left( \frac{\exp(\beta_k x_i(t_i))}{\sum_{j=1}^n Y_j(t_i) w_j(t_i) \exp(\beta_k x_j(t_i))} \right)^{I(\delta_i \epsilon_i = k)}, \quad (2.22)$$

em que

$Y(t)$ : indicadora de indivíduo sob risco de falha,

$w(\cdot)$ : peso da observação no tempo,

$k = 1, \dots, K$  é a causa da falha.

$K$ : número de causas de falha.  $\delta_i$ : indicadora de falha.  $\epsilon$ : causa de falha

Sendo  $T$  uma variável aleatória de forma que  $T = \min(D, C)$ , em que  $D$  é o tempo latente (tempo "oculto", não manifesto) de falha e  $C$  o tempo latente de censura. Desse modo  $\delta_i = 1$ , se  $C > T$  e  $\delta_i = 0$ , caso contrário. Para o  $i$ -ésimo indivíduo,  $i = 1, \dots, n$ , sejam  $t_i$ ,  $\delta_i$ ,  $\epsilon_i$ , e  $X_i(t)$  o tempo observado, a variável indicadora de evento, a causa de falha e o vetor de covariáveis no tempo de acompanhamento  $t$ , respectivamente. Assuma que existam  $K$  causas de falha observáveis ( $\epsilon_i \in (1, \dots, K)$ ). Considere que a causa de falha 1 seja a de interesse, e as falhas por outras causas como eventos competitivos.

Sejam

$$N_i(t) = I(T_i \leq t, \epsilon_i = 1), \quad (2.23)$$

$$Y_i(t) = 1 - N_i(t-). \quad (2.24)$$

Note que se  $\epsilon_i = 1$ , então  $N_i(t) = I(T_i \leq t)$  e  $Y_i(t) = I(T_i \geq t)$ ; se  $\epsilon_i \neq 1$ , então  $N_i(t) = 0$  e  $Y_i(t) = 1$ . Seja

$$r_i(t) = I[C_i \geq \min(T_i, t)], \quad (2.25)$$

$$w_i(t) = r_i(t) \frac{G(t)}{G[\min(T_i, t)]}, \quad (2.26)$$

em que  $G(t)$  é a estimativa Kaplan-Meier da função de sobrevivência da variável de censura que é calculada usando  $\{T_i, 1 - \delta_i, i = 1, 2, \dots, n\}$ . Se  $\delta_i = 0$ , então  $r_i(t) = 1$  quando  $t \leq T_i$  e 0 caso contrário. Se  $\delta_i = 1$ , então  $r_i(t) = 1$ . A Tabela 2 mostra os pesos em função do tempo.

Tabela 2 – Peso para o  $i$ -ésimo indivíduo

$t, T_i$	Estado	$r_i(t)$	$Y_i(t)$	$w_i(t)$
$t \leq T_i$	$\delta_i = 0$	1	1	1
	$\delta_i \epsilon_i = 1$	1	1	1
	$\delta_i \epsilon_i \neq 1$	1	1	1
$t > T_i$	$\delta_i = 0$	0	1	0
	$\delta_i \epsilon_i = 1$	1	0	$G(t)/G(T_i)$
	$\delta_i \epsilon_i \neq 1$	1	1	$G(t)/G(T_i)$

Haverá um conjunto de coeficientes  $\beta$  para cada causa de falha separadamente, de forma que as covariáveis podem incidir de maneira diferente em cada uma delas. Os efeitos desses coeficientes sobre as funções de risco específico são proporcionais. Em um modelo onde não há riscos competitivos, isso leva à proporcionalidade das funções de incidência acumulada impedindo o cruzamento entre suas respectivas curvas. Porém, na presença de riscos competitivos a função de incidência acumulada depende não somente do risco da  $k$ -ésima causa, mas também dos riscos das outras causas. Isso dificulta a previsibilidade do efeito dos coeficientes no comportamento das curvas de incidência acumulada.

Para ilustrar esse fenômeno, suponha uma situação onde se tem dois grupos de indivíduos (A e B) e duas causas de morte (1 e 2) sendo 1 a causa de interesse. O grupo A possui taxa de falha constante de 30 para ambas as causas de falha. O grupo B possui taxas de falha de 0,25 e 0,10 para as causas 1 e 2 respectivamente. Obviamente o grupo B possui taxas de falha proporcionais e menores para ambas as causas, porém as  $FIA_1$  não são proporcionais e após o tempo 2 a incidência acumulada do grupo B é maior que a do grupo A. Isso ocorre pois, para o grupo A, o risco da causa 2 (0.30) é muito maior que o do grupo B (0.10). Com isso, há mais indivíduos que sobreviveram a ambas as causas de morte no grupo B do que no grupo A. Como pode ser visto na Tabela 3, a quantidade de sobreviventes no Tempo 1 para o grupo B (6500) é maior que para o grupo A (4000).

Tabela 3 – Eventos Instantâneos

Grupo	Causa	Risco	Tempo 0	Tempo 1	Tempo 2	Tempo 3	Tempo 4	Tempo 5
A	1	0.3	0	3000	1200	480	192	77
	2	0.3	0	3000	1200	480	192	77
	Sobreviventes		10000	4000	1600	640	256	102
Grupo	Causa	Risco	Tempo 0	Tempo 1	Tempo 2	Tempo 3	Tempo 4	Tempo 5
B	1	0.25	0	2500	1625	1056	687	446
	2	0.1	0	1000	650	423	275	179
	Sobreviventes		10000	6500	4225	2746	1785	1160

Os indivíduos do grupo A morrem em grande número pela causa 2, o que impede que eles morram pela causa 1. Enquanto para os indivíduos do grupo B, eles morrem pouco pela causa 2, o que faz com que eles fiquem mais expostos aos risco 1 e, conseqüentemente, morram mais por esse risco.

A Tabela 5 mostra a proporção de indivíduos de cada grupo que morreu da causa 1. As proporções de falha, no Tempo 1, são de 30% e 25% para os grupos A e B respectivamente. No Tempo 2, as proporções são bem próximas (42% e 41%). Já no Tempo 3, a proporção de falhas pela causa 1 para o grupo B supera a do grupo A. Isso mostra que, apesar de a taxa de falha instantânea para a causa 1 do grupo A ser maior que para o grupo B, a probabilidade acumulada de falha pela causa 1 no grupo B pode superar a do grupo A. Isso ocorre pela diferença na taxa de falha pela causa 2 entre os grupos. Esta possibilidade pode ser vista na Figura 1.

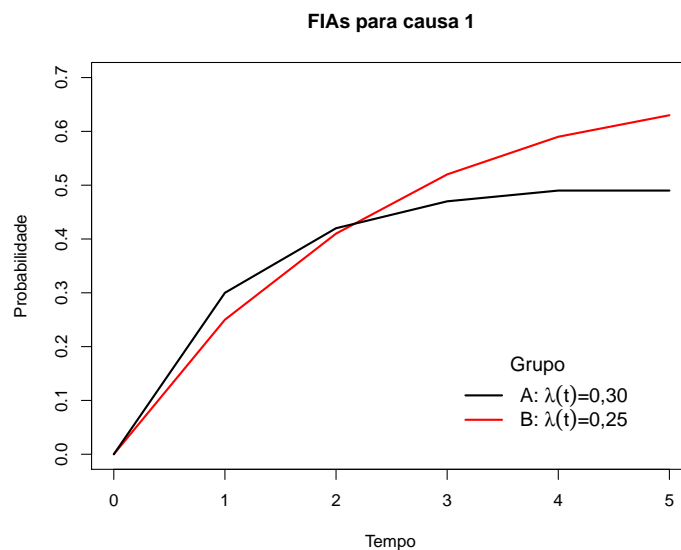
Tabela 4 – Eventos Acumulados (Causa 1)

Grupo	Tempo 0	Tempo 1	Tempo 2	Tempo 3	Tempo 4	Tempo 5
A	0	3000	4200	4680	4872	4949
B	0	2500	4125	5181	5868	6314

Tabela 5 – Incidência acumulada (Causa 1)

Grupo	Tempo 0	Tempo 1	Tempo 2	Tempo 3	Tempo 4	Tempo 5
A	0	0.30	0.42	0.47	0.49	0.49
B	0	0.25	0.41	0.52	0.59	0.63

Figura 1 – Gráfico de função de incidência acumulada



Fonte: Elaborado pelos autores (2017).

## 3 Resultados e Discussão

### 3.1 Análise Descritiva

Primeiramente, foi realizada uma análise descritiva dos dados. Na Tabela 6, temos as variáveis com as frequências absolutas e respectivos percentuais de pacientes, de falhas e de censuras.

Portanto, observa-se na Tabela 6, que para a variável sexo, temos que 52,6% são do sexo feminino. Para a idade, nota-se que 61,1% dos pacientes possuem menos de 65 anos. Quanto à escolaridade, temos que a maioria (65,1%) possui menos de 4 anos de estudo. Em relação à renda dos pacientes acompanhados durante esse estudo, 66,2% tem renda superior a dois salários mínimos.

Tendo em conta o número de episódios de peritonites, temos que 73,6% tiveram nenhum episódio, enquanto apenas 1,5% sofreram 4 peritonites. Para a maioria destes pacientes, a distância da residência até o centro de hemodiálise é menor que 50 km (57,0%). Com relação aos pacientes que têm diabetes, doença arterial coronariana e hipertensão, percebe-se que são 43,6%, 78,6% e 27,3% respectivamente. 63,8% não fizeram uma hemodiálise anterior. Dos indivíduos acompanhados, Acerca da modalidade de terapia, 48,0% fizeram a diálise peritoneal automatizada (APD), enquanto 52,0% realizaram a diálise peritoneal ambulatorial contínua (CAPD).

Em consideração ao percentual de falhas, nota-se que para as covariáveis sexo, idade, diabetes, distância até o centro, hemodiálise anterior e o tipo da modalidade da terapia, os percentuais estão bem distribuídos entre as categorias. No entanto, para as covariáveis escolaridade e renda, houve maior concentração nas categorias menos de 4 anos e menos de 2 salários mínimos, respectivamente.

Considerando o número de episódios de peritonites, há um percentual maior de falhas para nenhuma peritonite (13,1%), porém esta é a categoria com maior número de pacientes (4210). Para as covariáveis doença arterial coronariana e hipertensão, o percentual de falha é maior na categoria em que os pacientes que tem (12,9%) e na categoria não (13,7%), respectivamente.

Tabela 6 – Frequências absolutas e percentuais dos pacientes

Covariável	Categoria	N	Censura	Falha
Sexo	Feminino	2972 (52.6%)	2196 (38.8%)	483 (9.0%)
	Masculino	2679 (47.4%)	2419 (42.8%)	553 (10.0%)
Idade	< 65 anos	3452 (61.1%)	3011 (53.2%)	440 (7.7%)
	≥ 65 anos	2202 (38.9%)	1607 (28.4%)	595 (10.5%)
Escolaridade	< 4 anos	3686 (65.1%)	2932 (51.9%)	754 (13.3%)
	≥ 4 anos	1968 (34.9%)	1686 (30.0%)	282 (5%)
Renda	< 2 salários	1913 (33.8%)	1527 (27.0%)	386 (4.5%)
	≥ 2 salários	3741 (66.2%)	3091 (55.0%)	650 (11.5%)
Episódios de peritonites	Nenhuma	4210 (73.6%)	3464 (61.3%)	746 (13.1%)
	Uma	945 (16.9%)	730 (13.1%)	215 (3.8%)
	Duas	335 (6.0%)	285 (5.0%)	50 (0.8%)
	Três	116 (2.1%)	98 (1.7%)	18 (0.2%)
	Quatro	85 (1.5%)	41 (0.7%)	7 (0.1%)
Distância até o centro	< 50 km	3224 (57.0%)	2629 (46.6%)	595 (10.5%)
	≥ 50 km	2427 (43.0%)	1986 (35.2%)	441 (7.8%)
Diabetes	Não	3191 (56.4%)	2738 (48.4%)	453 (7.9%)
	Sim	2463 (43.6%)	1880 (33.2%)	583 (10.5%)
Doença arterial coronariana	Não	1205 (21.4%)	901 (16.0%)	304 (5.3%)
	Sim	4449 (78.6%)	3717 (65.7%)	696 (12.9%)
Hipertensão	Não	4109 (72.7%)	3329 (58.8%)	780 (13.7%)
	Sim	1545 (27.3%)	1289 (23.0%)	256 (4.5%)
Hemodiálise anterior	Não	2044 (36.8%)	1603 (28.3%)	441 (7.8%)
	Sim	3610 (63.2%)	3015 (53.3%)	595 (10.5%)
Modalidade da terapia	APD	2546 (48.0%)	1993 (37.6%)	553 (10.4%)
	CAPD	2751 (52.0%)	2299 (43.4%)	452 (8.5%)

APD: Diálise peritoneal automatizada

CAPD: Diálise peritoneal ambulatorial contínua

Fonte: Baxter Healthcare Brasil (2017).



## 3.2 Modelo de Cox

Para o primeiro ajuste, foi considerado o modelo de regressão de Cox, utilizando apenas como covariável o número total de episódios de peritonites sofridos por cada paciente. Esta variável foi considerada como fator.

Na Tabela 7, temos a estimativa do parâmetro, o erro padrão e o teste de Wald com o respectivo  $p$ -valor. Com essa estimativa, temos que quanto mais peritonites, menor será o risco de falha, ou que o acréscimo de uma peritonite, diminuirá o risco. Porém, isso é contra-intuitivo, do ponto de vista clínico. Pelo teste, temos evidências de que o parâmetro difere significativamente de zero.

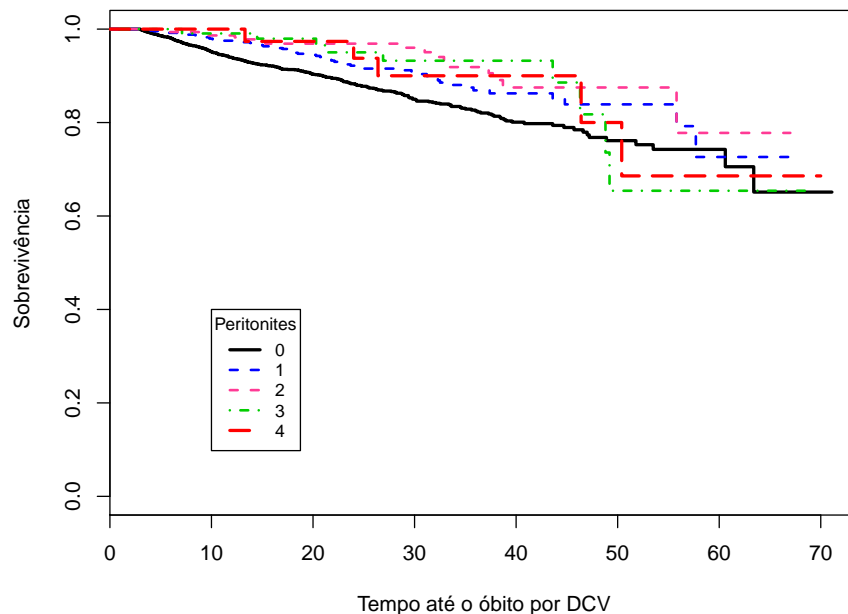
Tabela 7 – Ajuste do modelo de Cox para os dados de peritonites

Parâmetro	Coef	Exp(coef)	Erro padrão	Wald	valor $p$
Peritonites 1	-0.50	0.60	0.14	-3.62	0.00
Peritonites 2	-0.90	0.41	0.25	-3.60	0.00
Peritonites 3	-0.63	0.54	0.34	-1.85	0.06
Peritonites 4	-0.44	0.64	0.45	-0.99	0.32

Fonte: Elaborado pelos autores (2017).

Ao observar as curvas de sobrevivência pela Figura 2, nota-se que, quanto maior o número de peritonites, maior será a sobrevivência.

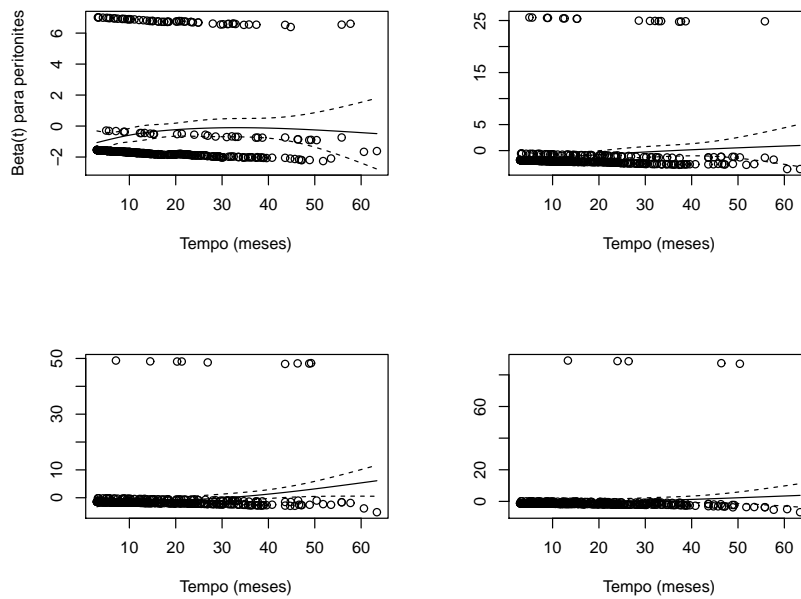
Figura 2 – Gráfico das curvas de sobrevivências considerando a covariável de episódios de peritonites



Fonte: Elaborado pelos autores (2017).

Na Figura 3, temos os resíduos de Schoenfeld. Visualmente, nota-se que existe uma tendência ao longo do tempo. Ao observar os valores obtidos para  $\rho$  (0.0855, 0.0982, 0.1331, 0.0923), que não são próximos de zero, leva-se então a rejeição da hipótese nula  $H_0 : \beta(t) = \beta$  ( $p$ -valor  $< 0.0001$ ). Logo, existe evidências de que a suposição de taxas proporcionais não é satisfeita para esse modelo.

Figura 3 – Resíduos de Schoenfeld para o modelo de Cox considerando a covariável número de episódios de peritonites



Fonte: Elaborado pelos autores (2017).

Dessa forma, conforme os resultados apresentados, o modelo de Cox com a covariável número de peritonites não se adequou bem os dados.

### 3.3 Modelo de Cox Estendido

Conforme visto, a falta de ajuste no modelo de Cox, apresentado na Seção 3.2, resultou no ajuste do modelo de Cox Estendido, levando em consideração o número de episódios de peritonites como um fator mudando ao longo do tempo e, como grupo basal, os pacientes que não tiveram nenhum episódio de peritonite.

As estimativas dos parâmetros, os erros padrão e o teste de Wald com o respectivo  $p$ -valores são apresentadas na Tabela 8. Ao verificar essas estimativas, temos que, à medida que o número de peritonites aumenta, maior será o risco de falha. O risco de óbito por DCV aumenta conforme o número de episódios (1.25, 1.29, 2.41 e 5.11). Considerando a razão dos riscos de óbito por DCV – isto é, a divisão entre o risco de óbito por DCV de pacientes que tiveram algum episódio de peritonite com aqueles que não tiveram nenhum

episódio de peritonite –, os pacientes com quatro episódios de peritonite têm um risco de mortalidade cardiovascular aproximadamente cinco vezes maior do que os pacientes que não sofreram nenhuma peritonite.

Tabela 8 – Ajuste do modelo de Cox estendido para os dados de peritonites

<b>Parâmetro</b>	<b>Coef</b>	<b>Exp(coef)</b>	<b>Erro padrão</b>	<b>Wald</b>	<b>Valor <math>p</math></b>	<b>IC (95%)</b>
1 peritonite	0.2222	1.2488	0.1408	1.578	0.114513	(0.9477;1.646)
2 peritonites	0.2607	1.2978	0.2520	1.034	0.300955	(0.7920;2.127)
3 peritonites	0.8832	2.4186	0.3427	2.577	0.009960	(1.2356;4.734)
4 peritonites	1.6326	5.1169	0.4642	3.517	0.000436	(2.0601;12.710)

Fonte: Elaborado pelos autores (2017).

Os valores encontrados para  $\rho$ , dessa vez, são próximos de zero. E a hipótese de não tendência no tempo, para cada parâmetro do modelo, não é rejeitada. Estes valores podem ser vistos na Tabela 9. É necessário certa atenção aos valores dos coeficientes de correlação de  $\rho$  pois, se amostra for muito grande, a significância estatística será apontada mesmo para valores muito próximos de zero (COLOSIMO; GIOLO, 2006). Analisando graficamente os resíduos de Schoenfeld, percebe-se que não há uma certa tendência para as variáveis. Apenas para a variável 1 peritonite, que aparenta ter uma tendência, mas vale ressaltar que é a categoria com mais indivíduos. Este gráfico encontra-se no Apêndice A, na Figura 10.

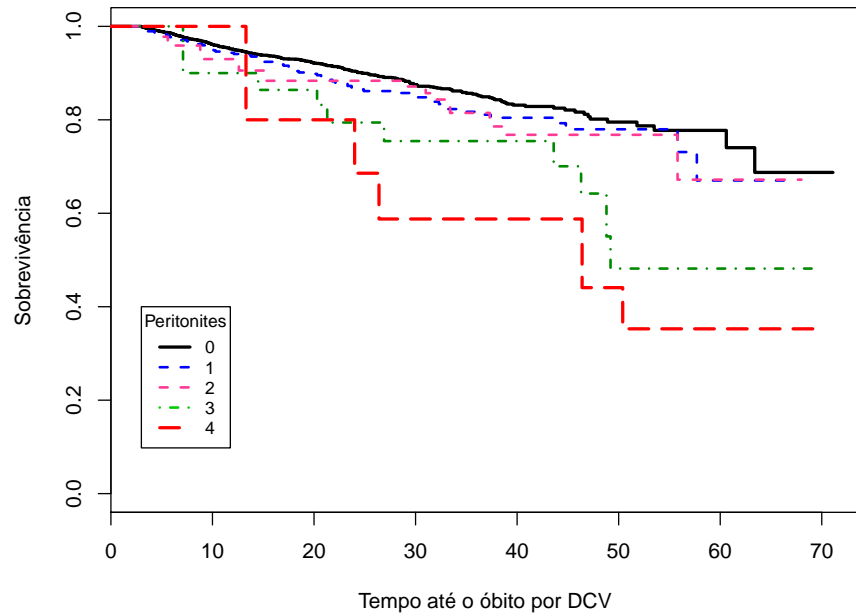
Tabela 9 – Testes da proporcionalidade de riscos tempo dependentes do modelo de Cox estendido

<b>Parâmetro</b>	<b><math>\rho</math></b>	<b><math>\chi^2</math></b>	<b>Valor <math>p</math></b>
1 peritonite	-0.0228	0.233	0.629
2 peritonites	-0.0160	0.114	0.736
3 peritonites	0.0242	0.257	0.612
4 peritonites	-0.0363	0.609	0.435
<b>Global</b>	-	1.214	0.876

Fonte: Elaborado pelos autores (2017).

Na Figura 4, observa-se que, quanto maior o número de peritonites, menor será a sobrevivência do indivíduo, conforme o esperado do ponto de vista clínico.

Figura 4 – Gráfico das curvas de sobrevivências considerando a covariável de episódios de peritonites mudando no tempo



Fonte: Elaborado pelos autores (2017).

Depois da verificação da inclusão da variável peritonites mudando ao longo tempo, foi feita a escolha das possíveis covariáveis para entrarem no modelo de regressão de Cox. Para isto, foi utilizado o método proposto por Collett (COLLETT, 2003). As etapas desse método estão na Tabela 10.

O modelo final ficou com as seguintes covariáveis: sexo, idade, escolaridade, renda, diabetes, doença arterial coronariana, hemodiálise anterior e o número de episódios de peritonites mudando no tempo. A escolha deste modelo, foi feita também pelo Critério de Informação de Akaike (AIC), por apresentar o menor valor. Com isso, foi ajustado o modelo de Cox, já apresentado na Equação (2.3).

Tabela 10 – Seleção das covariáveis pelo método proposto por Collett

Passo	Modelo	-2 log L( $\theta$ )	TRV	Valor $p$	AIC
Passo 1	Nulo	6902.657	-	-	
	Sexo	6789.064	20.74	0.0009	
	Idade	6881.660	113.6	0.0000	
	Escolaridade	6880.700	21.96	0.0005	
	Renda	6887.372	15.29	0.0092	
	Distância até o centro	6887.236	15.17	0.0096	
	Diabetes	6842.026	60.63	0.0000	
	Doença arterial coronariana	6834.543	68.11	0.0000	
	Hipertensão	6885.891	16.77	0.0049	
	Hemodiálise anterior	6881.694	20.96	0.0008	
	Modalidade da terapia	6636.607	14.13	0.0148	
Passo 2	Exceto Modalidade	6700.117	-	0.0000	6724.117
	Exceto Modalidade e Sexo	6704.759	197.9	0.0000	6726.759
	Exceto Modalidade e Idade	6778.464	123.9	0.0000	6800.464
	Exceto Modalidade e Escolaridade	6702.017	200.4	0.0000	6724.017
	Exceto Modalidade e Renda	6701.795	200.6	0.0000	6723.795
	Exceto Modalidade e Distância	6702.234	202.3	0.0000	6765.485
	Exceto Modalidade e Diabetes	6732.485	169.9	0.0000	6745.229
	Exceto Modalidade e Doença arterial	6728.229	174.2	0.0000	6750.229
	Exceto Modalidade e Hipertensão	6712.572	189.8	0.0000	6734.572
Passo 3	Exceto Modalidade e Hemodiálise	6728.229	174.2	0.0000	6750.229
	Exceto Hipertensão	6460.832	189.6	0.0000	6489.524
	Exceto Hipertensão e Sexo	6465.524	184.9	0.0000	6563.539
	Exceto Hipertensão e Idade	6539.539	110.9	0.0000	6486.366
	Exceto Hipertensão e Escolaridade	6462.366	188.1	0.0000	6510.560
	Exceto Hipertensão e Renda	6486.56	163.9	0.0000	6512.314
	Exceto Hipertensão e Distância	6488.314	162.2	0.0000	6498.402
	Exceto Hipertensão e Diabetes	6486.56	163.9	0.0000	6597.506
	Exceto Hipertensão e Doença arterial	6488.314	162.2	0.0000	6485.402
Passo 4	Exceto Hipertensão e Hemodiálise	6699.966	202.4	0.0000	6723.966
	Exceto Distância e Sexo	6705.443	197.2	0.0000	6725.443
	Exceto Distância e Idade	6778.508	201.7	0.0000	6798.508
	Exceto Distância e Escolaridade	6702.7	199.7	0.0000	6722.700
	Exceto Distância e Renda	6700.718	201.8	0.0000	6722.718
	Exceto Distância e Diabetes	6734.216	168.2	0.0000	6754.216
	Exceto Distância e Doença arterial	6729.351	173.1	0.0000	6749.351
	Exceto Distância e Hemodiálise	6713.355	189.9	0.0000	6733.355
Passo 5	Exceto Modalidade, Hipertensão e Sexo	6705.443	200.2	0.0000	6725.443
	Exceto Modalidade, Hipertensão e Idade	6778.508	123.9	0.0000	6798.508
	Exceto Modalidade, Hipertensão e Escolaridade	6702.72	198.3	0.0000	6722.700
	Exceto Modalidade, Hipertensão e Renda	6702.371	200	0.0000	6749.351
	Exceto Modalidade, Hipertensão e Diabetes	6734.216	168.2	0.0000	6760.384
	Exceto Modalidade, Hipertensão e Doença arterial	6767.692	134.7	0.0000	6710.394
	Exceto Modalidade, Hipertensão e Hemodiálise	6742.384	202.2	0.0000	620.89
	Exceto Modalidade, Hipertensão e Distância	6702.371	200	0.0000	6702.371

Fonte: Elaborado pelos autores (2017).

Na Tabela 11, apresentam-se os resultados obtidos para o coeficiente de correlação. Pode-se notar que os valores  $\rho$  são próximos de zero, o que evidencia uma não violação da suposição das taxas proporcionais. Analisando graficamente os resíduos de Schoenfeld, percebe-se que não há uma certa tendência para as variáveis. Estes gráficos encontram-se no Apêndice A, Figura 11.

Tabela 11 – Testes da proporcionalidade das taxas de falha do modelo de Cox estendido com covariáveis

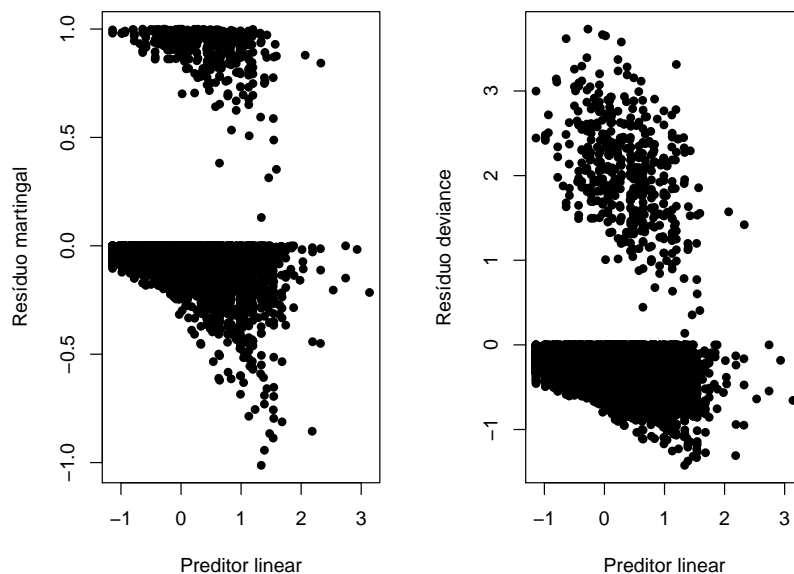
Parâmetro	$\rho$	$\chi^2$	Valor $p$
Sexo	0.01599	0.11119	0.739
Idade	0.00544	0.01386	0.906
Escolaridade	-0.05487	1.33991	0.247
Renda	0.04462	0.87800	0.349
Diabetes	0.01305	0.08001	0.777
Doença arterial coronariana	-0.02662	0.34329	0.558
Hemodiálise anterior	0.06078	1.59380	0.207
<b>Global</b>	-	5.19758	0.921

Fonte: Elaborado pelos autores (2017).

De modo geral, pode-se dizer, com base nos resultados, que o modelo de Cox não apresentou evidências de violação suas suposições. Após esta verificação, procedeu-se a fase de interpretação das estimativas.

Para avaliar a existência de observações atípicas foram obtidos os gráficos dos resíduos *martingal* e *deviance*. A partir da Figura 5 pode-se observar que os resíduos *deviance*, que são resíduos mais simétricos, não sugerem a existência de pontos que possam ser considerados atípicos (outliers), com uma exceção de alguns pontos mais próximos de 3.5 valores considerados aceitáveis para estes resíduos. Tem-se que que grande parte está distribuído em torno de zero.

Figura 5 – Resíduos *martingal* (Esquerda) e *deviance* (Direita) versus preditor linear associados ao modelo de Cox



Fonte: Elaborado pelos autores (2017).

As estimativas, os erros padrão, os valores de  $p$  para o teste de Wald, para o modelo selecionado, e intervalo de confiança de 95% são apresentadas na Tabela 12. Para todas covariáveis, foi evidente o efeito significativo ( $p < 0.05$ ), menos para escolaridade. Esta variável é considerada importante do ponto de vista de médico, por isso foi optado por deixá-la.

A partir das estimativas apresentadas na Tabela 12, percebe-se que, para o sexo masculino, o risco foi de 1.22 vezes a taxa estimada para o sexo feminino e este risco podendo variar entre 1.01 e 1.48 com 95% de confiança. Enquanto para os pacientes com idade superior a 65 anos, o risco de óbito por doença cardiovascular foi de 2.37 vezes o dos pacientes com idade inferior a 65 anos. Para escolaridade, obteve-se que o risco para os que estudaram menos de 4 anos foi de aproximadamente 1.17 vezes o de indivíduos com mais de 4 anos de estudo. Para renda, obteve-se que o risco para os que tinham renda inferior a 2 salários mínimos foi de aproximadamente 1.14 vezes o de indivíduos com mais de 2 salários mínimos.

Os pacientes que tinham diabetes apresentaram o risco de óbito por doença cardiovascular de 1.74 vezes o dos pacientes que não tem diabetes. Para os pacientes que tiveram doença arterial coronariana, o risco foi de 1.74 em comparação aos que não as tiveram. Para os que fizeram hemodiálise anterior, a taxa estimada foi de aproximadamente, 1.41 vezes a dos pacientes que não haviam feito uma hemodiálise anterior, podendo este risco variar entre 1.17 e 1.70, com 95% de confiança.

Em relação ao número de peritonites, observa-se que o risco de óbito por doença cardiovascular para pacientes que tiveram quatro episódios de peritonite é de 5.75 vezes o dos pacientes que não tiveram peritonite. Observa-se que, para um episódio peritonite, dois episódios peritonites este risco é de 1.22 e 1.11, podendo variar entre (0.67;1.81), e (0.97;3.73) respectivamente com 95% de confiança.

Tabela 12 – Resultados do ajuste do modelo de Cox estendido para os dados de peritonite

Covariável	Coef	Exp(coef)	Erro padrão	Valor $p$	IC (95%)
Sexo (Feminino <sup>a</sup> )	0.2065	1.2294	0.0971	0.0334	(1.0163;1.487)
Idade (< 65 <sup>a</sup> )	0.8647	2.3743	0.0990	0.0000	(1.9554;2.883)
Escolaridade (< 4 <sup>a</sup> )	-0.1551	0.8563	0.1110	0.1621	(0.6889;1.064)
Renda (< 2 <sup>a</sup> )	-0.1382	0.8709	0.1069	0.1960	(0.7063;1.074)
Diabetes (Não <sup>a</sup> )	0.5565	1.7446	0.0971	0.0000	(1.4989;2.1991)
Doença arterial coronariana (Não <sup>a</sup> )	0.5539	1.7401	0.1007	0.0000	(1.4285;2.120)
Hemodiálise anterior (Não <sup>a</sup> )	0.3532	1.4236	0.0964	0.0002	(1.1706;1.709)
Peritonite 1	0.1998	1.2212	0.1408	0.1558	(0.9267;1.609)
Peritonite 2	0.1047	1.1104	0.2519	0.6776	(0.6889;1.8492)
Peritonite 3	0.6464	1.9086	0.3428	0.0593	(0.9749;3.737)
Peritonite 4	1.7500	5.7549	0.4669	0.0002	(2.3048;14.369)

<sup>a</sup>Categoria de referência

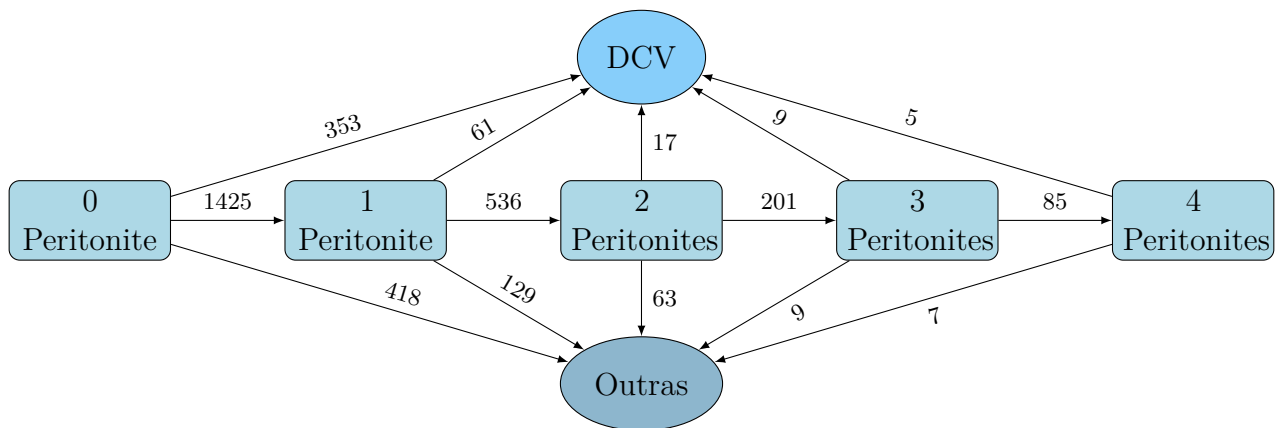
Fonte: Elaborado pelos autores (2017).

### 3.4 Modelo Multiestado

Outra abordagem para este problema é com os modelos multiestado. Nesta modelagem, todos os indivíduos iniciaram no estado de partida estado inicial, no qual todos os pacientes não tiveram nenhuma peritonite anteriormente, e terminam em um estado final (morte por doenças cardiovasculares e/ou outras causas). O caminho que cada indivíduo faz poderá ser diferente, assim havendo várias trajetórias possíveis.

A Figura 6, apresenta um esquema gráfico das possíveis transições entre os estados estudados. Por exemplo, um indivíduo começa no estado inicial que é nenhuma peritonite. Após algum tempo pode-se sofrer sua primeira peritonite, mudando assim para o estado 1 peritonite, morrer por doença cardiovascular (DCV) ou outras causas (Outras) ou então sofrer a segunda peritonite. Posto isso, os indivíduos não poderão transitar direto para a segunda peritonite sem ter sofrido a primeira e nem voltar para estágio anterior. Por este motivo, as setas apenas vão para os estados e não voltam. Os estados DCV e Outras, são absorventes, ou seja, o indivíduo entrando nele, permanecerá nele.

Figura 6 – Representação gráfica das transições entre estados



Fonte: Elaborado pelos autores (2017).

Um dos objetivos adicionais no modelo multiestado, é estudar as relações entre vários preditores e as intensidades das transições entre os estados. Estas frequências de mudança de estado encontram-se na Tabela 13. Na primeira linha, todos indivíduos estavam no estágio inicial, ou seja, nenhuma peritonite (P0). A partir de P0, estes indivíduos poderiam transitar apenas para 3 estados: primeira peritonite (P1), morte por doença cardiovascular (DCV) e morte por outras causas (Outras). Desse modo, 1456 tiveram uma peritonite, 353 morreram por morte cardiovascular e 418 por outras causas de mortes. Já, para censura, tem-se 3464 pacientes. Consequentemente totalizando 5691 eventos entre (P1), morte por doença cardiovascular, outras mortes e censuras.

Para a segunda linha, 536 indivíduos tiveram a segunda peritonite, 61 morreram por doença cardiovascular e 129 por outras causas de morte. Para censuras, tivemos 730



casos, totalizando 1456 eventos. E assim sucessivamente para as demais linhas.

Tabela 13 – Frequências das transições entre os estados

De	Para								
	P0	P1	P2	P3	P4	DCV	Outras	Censura	Total
P0	0	1456	0	0	0	353	418	3464	5691
P1	0	0	536	0	0	61	129	730	1456
P2	0	0	0	201	0	17	33	285	536
P3	0	0	0	0	85	9	9	98	201
P4	0	0	0	0	0	5	7	73	85

Fonte: Elaborado pelos autores (2017).

Na Tabela 14, tem-se as proporções para as transições. Logo, sabe-se que 0.2558 passaram de nenhuma peritonite para uma peritonite, 0.0620 para morte cardiovascular, 0.07344 para outras causas de morte e 0.6086 censuraram. As demais colunas são zero, pois são transições inexistentes. Por exemplo, um indivíduo não pode sair de nenhuma peritonite e ir para segunda, sem ter tido a sua primeira peritonite.

Tabela 14 – Proporções para as transições entre os estados

De	Para							
	P0	P1	P2	P3	P4	DCV	Outras	Censura
P0	0.0000	0.2558	0.0000	0.0000	0.0000	0.0620	0.07344	0.6086
P1	0.0000	0.0000	0.3681	0.0000	0.0000	0.0418	0.0885	0.5013
P2	0.0000	0.0000	0.0000	0.3750	0.0000	0.0317	0.0615	0.5317
P3	0.0000	0.0000	0.0000	0.0000	0.4228	0.0447	0.0447	0.4875
P4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0588	0.0823	0.8588

Fonte: Elaborado pelos autores (2017).

### 3.4.1 Ajustes Marginais

A fim de avaliar marginalmente o impacto de cada covariável no risco de morte por DCV, foi ajustado um modelo para cada covariável. A Tabela 15 mostra o impacto de cada covariável na transição para morte DCV. O valor na célula  $A_{ij}$  é o efeito (exponencial do coeficiente) da covariável  $i$  na transição do estado  $j$  para morte DCV. Nas células que contém um risco (-), não convergiram no modelo.

O sexo masculino mostrou-se como fator de proteção para pacientes sem nenhum episódio de peritonite. O risco de morte por DCV de indivíduos do sexo masculino é 78% do risco de indivíduos do sexo feminino a 5% de significância. O efeito da idade sobre o risco de morte por DCV permanece em torno de 2.5, enquanto o indivíduo tem zero ou um caso de peritonite. Ao atingir o 3º caso de peritonite, esse efeito é praticamente duplicado. Para mais episódios de peritonite, essa variável não se demonstrou significativa. Já para

Tabela 15 – Exponencial dos efeitos marginais das covariáveis para o risco de morte por DCV

Covariável	P0	P1	P2	P3	P4
Sexo (Feminino <sup>a</sup> )	0.78*	0.84	0.82	1.52	1.19
Idade (< 65 <sup>a</sup> )	2.69***	2.2**	2.19	5.17*	-
Escolaridade (< 4 <sup>a</sup> )	0.69**	1.45	0.72	0.15.	1.43
Renda ( $\geq 2^a$ )	1.07	0.87	0.78	6.66*	0.54
Distância até o centro (< 50 km <sup>a</sup> )	1.06	0.93	1.21	0.94	-
Diabetes (Não <sup>a</sup> )	2.07***	1.84*	0.66	1.6	0.96
Doença arterial coronariana (Não <sup>a</sup> )	2.17***	2.65***	0.64	1.42	1.06
Hipertensão (Não <sup>a</sup> )	1.26.	0.73	1.31	0.78	-
Hemodiálise anterior (Não <sup>a</sup> )	1.16	1.3	2.78.	5.1*	2.24
Modalidade da Terapia (APD <sup>a</sup> )	1.04	1.31	0.48	0.42	0.58
Cód. signif: 0 "****"0.001 "***"0.01 "*"0.05 "."0.1					

<sup>a</sup> Categoria de referência

Fonte: Elaborado pelos autores (2017).

o nível de escolaridade só teve efeito significativo para pacientes com nenhum ou três episódios de peritonite, sendo esses efeitos sempre de proteção. Os indivíduos com três episódios de peritonite e com renda baixa, têm risco de morte DCV de 6.6 vezes o risco de indivíduos com renda alta. A distância até o centro de hemodiálise não teve efeito significativo.

Para os pacientes que têm diabetes, aumenta em 107% e 84% o risco de morte por DCV para indivíduos com zero e um episódio de peritonite, respectivamente. E para doença arterial coronariana mais que duplica o risco de morte por DCV para pacientes com zero e um caso de peritonite. Já a hipertensão não se demonstrou impactante para o risco de morte por DCV. Ser hipertenso aumenta em 26% o risco de morte por DCV, mas apenas a níveis de 10% de significância. Pacientes que fizeram hemodiálise anterior e com três episódios de peritonite têm 5 vezes o risco de morte por DCV de pacientes com que não fizeram hemodiálise anterior. A modalidade de terapia, não apresentou-se significativa.

Para avaliar a qualidade dos ajustes marginais, foram utilizados os resíduos de Schoenfeld e o método de comparação entre as estimativas de sobrevivência do método de Kaplan-Meier e do modelo de Cox. O maior valor absoluto de coeficiente de correlação de Pearson entre os resíduos padronizados de Schoenfeld e o tempo foi 0,044 o que não sugere existência violação da suposição de taxas de falha proporcionais.

Para verificar a qualidade de ajuste foram comparadas as estimativas de Cox e de Kaplan-Meier. As estimativas de Cox ficaram satisfatoriamente próximas das estimativas de Kaplan-Meier sendo os coeficientes de correlação de Pearson entre as estimativas superiores a 0.98 para todas as covariáveis. A representação visual, bem como os coeficientes de correlação, encontram-se no Apêndice B, nas Figuras 12, 13 e 14.

### 3.4.2 Ajuste Conjunto

Sabe-se que não somente é possível, mas também recomendado que os modelos sejam ajustados com todas as covariáveis, pois assim é possível quantificar o efeito de cada covariável na presença das demais e verificar possíveis colinearidades. Foi ajustado o modelo com todas as covariáveis conjuntamente. Uma vez isso, foi utilizado o método de Collett de seleção de covariáveis (COLLETT, 2003) e obteve-se o modelo reduzido.

Na Tabela 16 contém o impacto de cada covariável, expresso pelo exponencial do coeficiente, no desfecho morte por DCV. Células sem valor, identificadas com um -, significam que o efeito da  $i$ -ésima covariável na transição do estado  $j$  para morte DCV, não incluído no modelo. Os efeitos das variáveis idade, renda e hemodiálise anterior mudaram bastante em comparação com o ajuste mostrado na Tabela 15. As outras covariáveis não sofreram grandes mudanças em seus efeitos. O tamanho da amostra teve impacto nos  $p$ -valores dos parâmetros dos estratos com maior número de indivíduos fazendo com que os efeitos fossem altamente significativos. Já o percentual elevado de censura trouxe problemas de convergência para os estratos com poucos indivíduos.

Tabela 16 – Modelo reduzido exponencializado para os dados de peritonite

Covariável	P0	P1	P2	P3	P4
Sexo (Feminino <sup>a</sup> )	0.79*	-	-	-	-
Idade (< 65 <sup>a</sup> )	2.42***	2.15**	2.54.	7.35*	-
Escolaridade (< 4 <sup>a</sup> )	0.73*	1.86*	-	-	-
Renda ( $\geq 2^a$ )	-	-	-	6.22*	-
Distância até o centro (< 50 km <sup>a</sup> )	-	-	-	0.24.	-
Diabetes (Não <sup>a</sup> )	1.7***	1.73*	-	-	-
Doença arterial coronariana (Não <sup>a</sup> )	1.79***	2.16**	-	-	-
Hemodiálise anterior (Não <sup>a</sup> )	1.27*	-	3.17*	8.73*	-
Cód. signif.: 0 "***"0.001 "***"0.01 "*"0.05 "."0.1					

<sup>a</sup> Categoria de referência

Fonte: Elaborado pelos autores (2017).

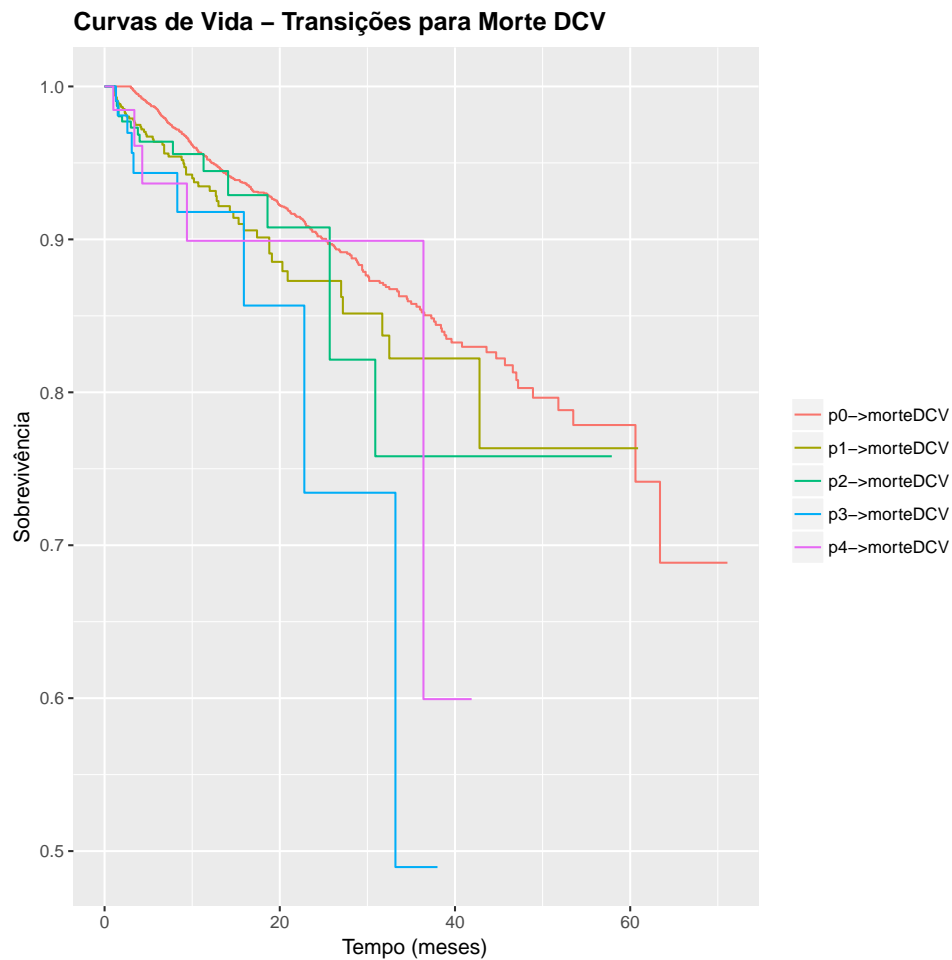
Sendo assim, os indivíduos que eram do sexo masculino mostraram-se como fator de proteção, e este risco de morte por DCV é 79% do risco do sexo feminino. O efeito de idade na transição de P3 para morte DCV é triplo, comparado com as transições partindo de P0, P1 e P2. Escolaridade não apresentou-se significativa. Ser de renda baixa aumenta em 522% o risco de transição de P3 para morte DCV. O aumento no efeito de diabetes com a com a ocorrência do primeiro episódio de peritonite foi de apenas 3%. O aumento do risco de morte por DCV de pacientes com doença arterial coronariana e nenhum episódio de peritonite salta de 75% para 116%. O efeito de hemodiálise anterior que era de 1.27, para indivíduos sem peritonites, saltou para 8.73 com a ocorrência do terceiro episódio de peritonite.

O coeficientes de correlação de Pearson entre os resíduos padronizados de Schoenfeld e o tempo, mostraram não haver evidências contra a suposição de taxas de falhas proporcionais. O maior valor absoluto de  $\rho$  foi 0.032. Testes paramétricos da hipótese de proporcionalidade das taxas de falha não foram utilizados em função do tamanho do banco de dados.

Como pode ser visto, o aumento na quantidade de episódios de peritonite intensifica os efeitos das covariáveis e deixa os indivíduos ainda mais propensos à morte por DCV.

Em um contexto exploratório com o intuito de verificar as probabilidades das transições em função do tempo, foram traçadas curvas de sobrevivência utilizando o método de Kaplan-Meier (do pacote *survival*) introduzindo a transição como covariável. O gráfico (Figura 7) mostra que a sobrevivência para o desfecho morte por DCV é maior para os indivíduos cujo estado inicial é P3 (3º episódio de peritonite) do que para os indivíduos cujo estado inicial é P0 (nenhum episódio de peritonite). Isso dá indícios de que quanto maior o número de episódios de peritonite maior é a probabilidade de morte por DCV.

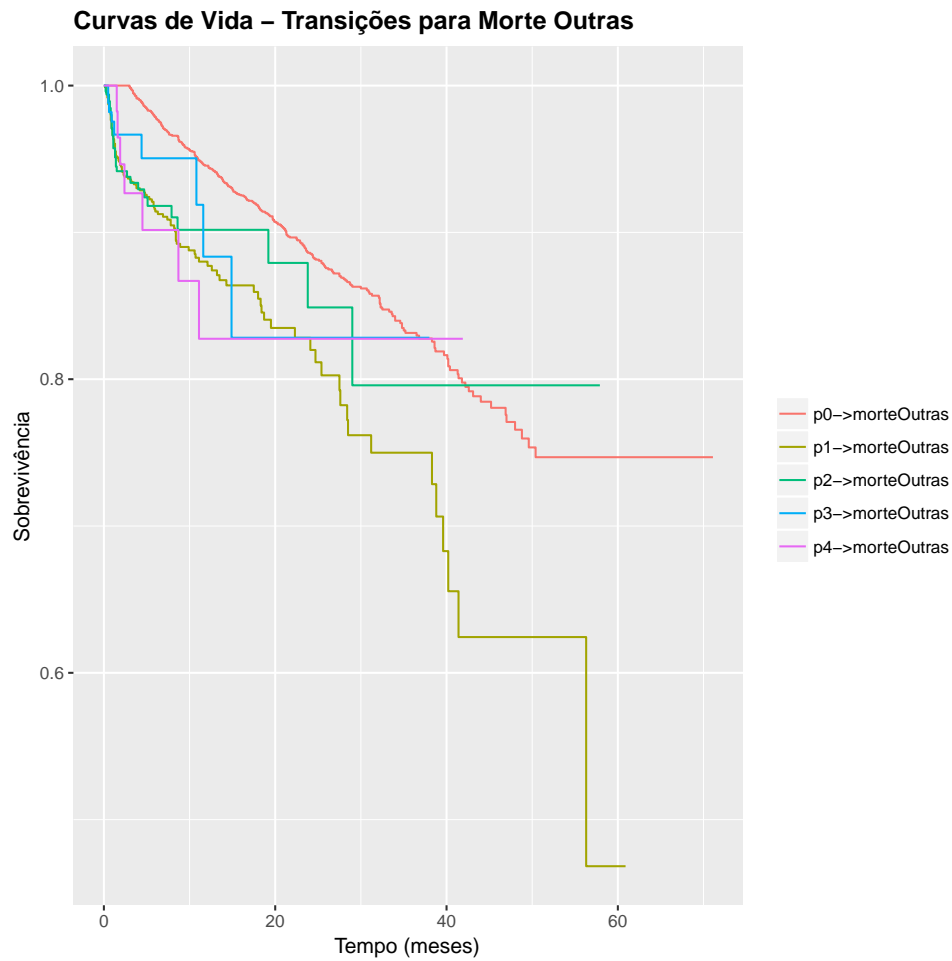
Figura 7 – Gráfico de sobrevivência para as transições para morte por DCV



Fonte: Elaborado pelos autores (2017).

O número de peritonites também diminui a sobrevivência quando o desfecho é morte por outras causas. Na Figura 8, é possível ver que a curva de sobrevivência dos indivíduos que partiram de P0 (nenhum episódio de peritonite) é maior que a dos que partiram de P1 (1º episódio de peritonite). Outra conveniência da utilização de modelos multiestados é a possibilidade de estimar as probabilidades de ocorrências de efeitos intermediários.

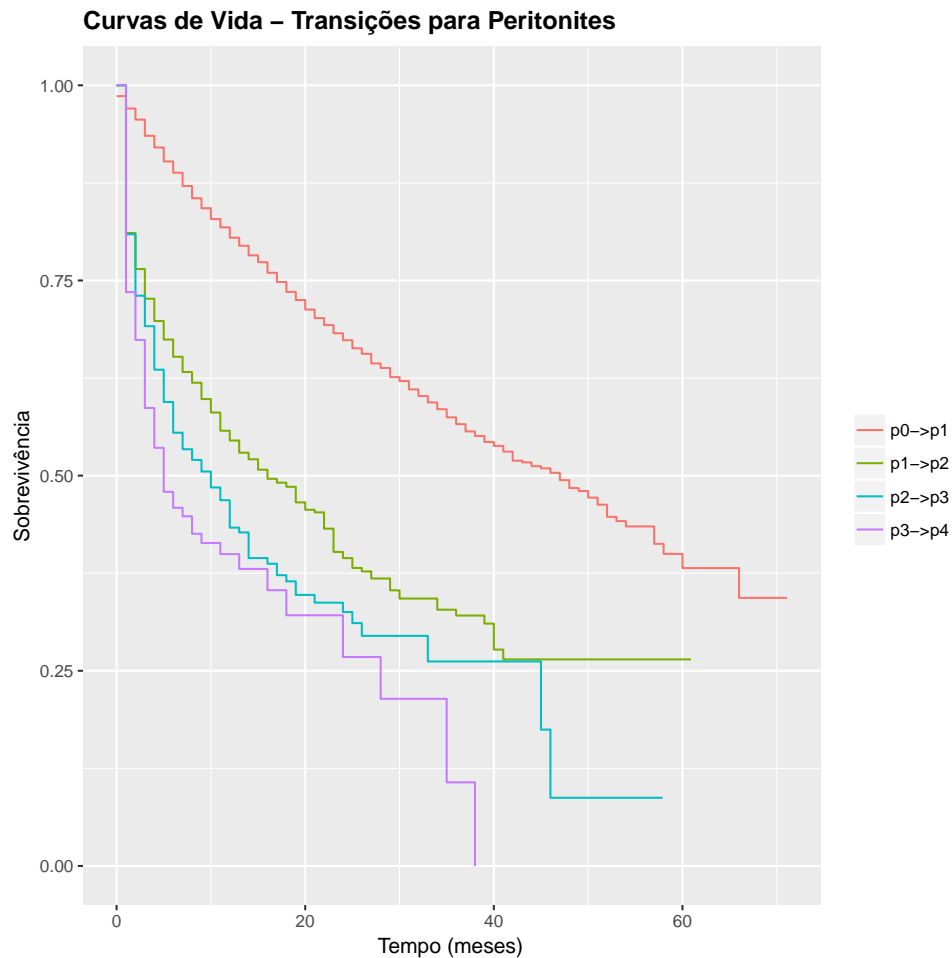
Figura 8 – Gráfico de sobrevivência para as transições para morte por outras causas



Fonte: Elaborado pelos autores (2017).

Na Figura 9 tem-se o gráfico que contém as curvas de sobrevivência para os episódios de peritonite. Como era de se esperar, quanto mais episódios de peritonite o indivíduo teve maior a probabilidade da ocorrência do próximo episódio de peritonite.

Figura 9 – Gráfico de curvas de vida para as transições para peritonites



Fonte: Elaborado pelos autores (2017).

Para mais detalhes sobre os ajustes, encontram-se disponível no link <https://goo.gl/hDolGt>

### 3.5 Modelo de Riscos Competitivos

Foi ajustado o modelo de riscos competitivos no qual o número de episódios de peritonite é usado como variável tempo dependente. Estes resultados estão na Tabela 17.

Tabela 17 – Modelo exponencializado reduzido de riscos competitivos

Covariável	Exp(coef)	Valor <i>p</i>	Signif
Sexo (Feminino <sup>a</sup> )	0.80	0.02	*
Idade (< 65 <sup>a</sup> )	2.21	0.00	***
Escolaridade (< 4 <sup>a</sup> )	0.88	0.28	
Renda (≥ 2 <sup>a</sup> )	1.08	0.47	
Distância até o centro (< 50 km <sup>a</sup> )	1.05	0.61	
Número de peritonites	1.24	0.00	***
Diabetes (Sim <sup>a</sup> )	1.61	0.00	***
Doença arterial coronariana (Sim <sup>a</sup> )	1.73	0.00	***
Hipertensão (Sim <sup>a</sup> )	1.05	0.70	
Hemodiálise anterior (Sim <sup>a</sup> )	1.33	0.00	**
Modalidade (APD <sup>a</sup> )	0.96	0.65	
Cód. signif.: 0 "****"0.001 "***"0.01 "**"0.05 "."0.1			

<sup>a</sup> Categoria de referência

Fonte: Elaborado pelos autores (2017).

Da mesma forma que no modelo multiestado, o efeito de ser do sexo masculino apresentou-se como fator de proteção e este risco é 80% do risco do sexo feminino. Indivíduos com mais de 65 anos têm risco de morte por DCV 1,2 vezes maior do que dos indivíduos com menos de 65 anos. O risco de morte por DCV de pacientes com diabetes é 1,61 vezes o de pacientes sem diabetes. Para os pacientes que possuem doença arterial coronariana, o risco de morte por DCV aumenta em 73%. Pacientes que já foram submetidos ao tratamento de hemodiálise, têm aumento de 33% no risco de morte por DCV. Para cada novo episódio de peritonite que o paciente tem o risco de morte por DCV tem um aumento em 24%.

O modelo com as interações entre todas as covariáveis e o número de peritonites como fator foi ajustado e não teve nenhuma interação significativa.



## 4 Considerações Finais

A diálise é o tratamento mais utilizado para pacientes com insuficiência renal. Dentre os brasileiros, mais de 50 mil estavam em diálise no ano de 2014 dos quais aproximadamente 3400 faziam diálise peritoneal. Existe a preocupação do bem estar desses pacientes, sendo a peritonite uma infecção comum. A peritonite também está sendo associada a doenças cardiovasculares tardias. Neste contexto, as análises feitas neste trabalho tiveram como objetivo identificar a associação entre os episódios da peritonite e morte por DCV, e determinar fatores importantes para a sobrevida dos 5691 pacientes acompanhados que faziam diálise peritoneal em 122 centros de diálise no Brasil, no período de Dezembro de 2004 a Janeiro de 2011.

Primeiramente, foi feita uma análise descritiva para um melhor conhecimento do dados. Observou-se que a maioria eram do sexo feminino, maiores de 65 anos, com menos de 4 anos de estudo, com a renda superior a dois salários mínimos, residindo a menos de 50 km do centro de diálise, não tinham diabetes e hipertensão. A maioria, os pacientes apresentava doença arterial coronariana, fizeram hemodiálise anterior e a modalidade do tipo de terapia foi equilibrada. Após essa primeira análise, foi ajustado o modelo de regressão de Cox. Com esse modelo, pode-se notar que a interpretação era errônea do ponto de vista clínico, em que quanto menos episódios de peritonites, menor seria a sobrevivência.

Em seguida, ajustou-se o modelo de Cox estendido considerando o número de episódios de peritonites mudando no tempo. Verificou-se, com apenas a covariável peritonites, que o risco de óbito por DCV aumentou conforme o número de episódios. Depois, foi ajustado com todas as variáveis. Para seleção das variáveis que iriam permanecer no modelo, utilizou-se do método de seleção de Collett do qual ficou as variáveis: sexo, idade, escolaridade, renda, diabetes, doença arterial coronariana e hemodiálise anterior. Observou-se que o risco de morte por DVC foi maior para o sexo masculino e com idade superior a 65 anos. Também foi visto que o risco foi maior para os que estudaram menos de 4 anos. Os pacientes que possuíam diabetes, doença cardiovascular coronariana desfrutaram de um risco maior dos que não tinham e o mesmo acontece para os fizeram hemodiálise anteriormente.

Em continuidade, foi ajustado o modelo multiestados. Com base neste modelo, foi possível observar o risco entre as transições de episódios de peritonite. Primeiro, foram ajustados os modelos marginais para cada covariável. Para seleção das covariáveis, também foi utilizado o método de Collett. O modelo final ficou com as seguintes variáveis: sexo, idade, escolaridade, renda, distância da residência até o centro, diabetes, doença arterial coronariana e hemodiálise anterior. Foi observado que o efeito de idade na transição terceira peritonite para morte DCV foi o triplo triplo quando comparado com as transições

partindo de nenhuma, uma e duas. Também que os indivíduos do sexo masculino tiveram um risco de morte por DCV de 78% maior que os indivíduos do sexo feminino a 5% de significância. Foi verificado que pacientes com doença arterial coronariana tiveram o risco de morte por DCV aumentado em 30% em relação aos que tiveram nenhum e um episódio de peritonite. Notou-se que o efeito de hemodiálise anterior aumentou significativamente para pacientes sem peritonites para a ocorrência do terceiro episódio de peritonite. A qualidade do ajuste do modelo foi verificada e atestou-se ajuste satisfatório.

Prosseguindo as análises, foi ajustado o modelo de riscos competitivos, considerando a variável episódios de peritonites. As covariáveis significativas foram: sexo, idade, número de episódios de peritonites, diabetes, doença arterial coronariana e hemodiálise anterior. Pode-se notar que o risco de morte por doenças cardiovasculares foi maior para pacientes que apresentaram idade superior a 65 anos, que tinham diabetes e doença arterial coronariana. Observou-se que os pacientes que foram submetidos a hemodiálise anterior a diálise peritoneal, também tiveram um risco maior do que os que não foram submetidos. E, conforme o número de episódios de peritonites aumentou, também aumentou o risco de morte por doenças cardiovasculares.

A abordagem de multiestados mostrou-se flexível o bastante para acomodar várias situações complexas, sobretudo nas áreas industrial e empresarial que, em geral, têm interesse em predições. Com as estimativas de probabilidade de cada transição é possível saber qual cenário é mais provável e assim traçar a melhor estratégia baseando-se nos riscos e custos envolvidos. Contudo, um modelo como esse envolve muitos parâmetros e o percentual elevado de censuras trouxe problemas de estimação desses parâmetros. A abordagem de riscos competitivos tem a vantagem de representar mais fielmente as probabilidades de ocorrência de cada evento (no sentido de que não desrespeita os axiomas de probabilidade) e de ter um número bem menor de parâmetros, o que diminui a complexidade e problemas de estimação. Em um contexto em que se deseja inferir sobre os parâmetros, o modelo de riscos competitivos é mais atraente por ser mais simples. Pesquisadores e acadêmicos, usualmente, pretendem identificar fatores de riscos e proteção. Também têm interesse em comparar tratamentos e grupos de indivíduos. Sendo assim, sugere-se um modelo de interpretação mais direta e objetiva, que é o caso do modelo de riscos competitivos.

Para a continuação do estudo, pretende-se implementar riscos competitivos nas transições entre os estados. Propõe-se também um estudo de simulação para verificar e possivelmente quantificar a influência do percentual de censura nas estimações. Além disso, pretende-se transformar o script R em um pacote implementando resíduos de Schoenfeld, teste t, gráficos específicos para o modelo de riscos competitivos.

# Referências

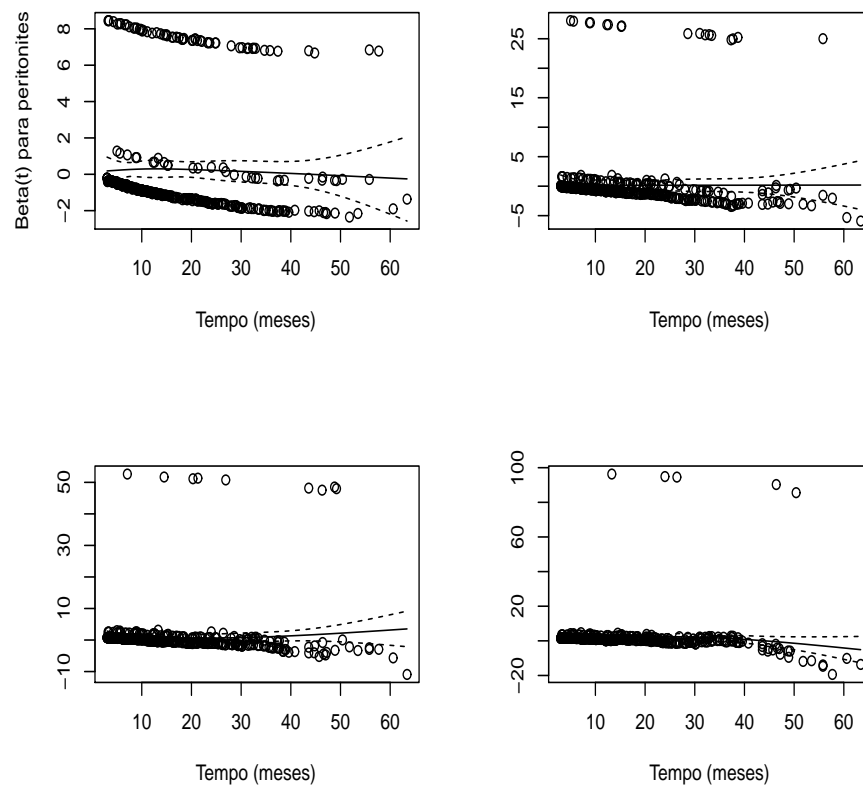
- BARBOSA, D. A. et al. Co-morbidade e mortalidade de pacientes em início de diálise. *Acta paul enferm*, SciELO Brasil, v. 19, n. 3, p. 304–9, 2006.
- BARBOSA, K. R. A. et al. Indicações de hemodiálise de emergência em uma unidade de terapia intensiva de um hospital particular da cidade de atibaia-sp. *Science in Health*, v. 3, n. 3, p. 131–8, 2012.
- BARRETI, P.; MONTELLI, A. C. et al. Atualização em diálise: tratamento atual das peritonites em diálise peritoneal ambulatorial contínua. *Jornal Brasileiro de Nefrologia*, São Paulo, v. 23, n. 2, p. 114–20, 2001.
- CARVALHO, M. S. et al. *Análise de Sobrevivência: teoria e aplicações em saúde*. [S.l.]: SciELO-Editora FIOCRUZ, 2011.
- CHO, Y.; JOHNSON, D. W. Peritoneal dialysis-related peritonitis: Towards improving evidence, practices, and outcomes. *American Journal of Kidney Diseases*, Elsevier, v. 64, n. 2, p. 278–289, 2014.
- COLLETT, D. Modelling survival data in medical research. CRC Press, 2003.
- COLOSIMO, E. A.; GIOLO, S. R. *Análise de sobrevivência aplicada*. São Paulo, Brasil: Edgard Blücher, 2006. 392 p. (ABE - Projeto Fisher). ISBN 9788521203841.
- COX, D. R. Regression models and life-tables. In: *Journal of the Royal Statistical Society. Series B (Methodological)*. [S.l.]: Springer, 1972. p. 187–220.
- COX, D. R.; SNELL, E. J. A general definition of residuals. *Journal of the Royal Statistical Society. Series B (Methodological)*, JSTOR, p. 248–275, 1968.
- FERNANDES, A. M. R. *Análise de dados em modelos multiestado*. Tese (Doutorado) — Universidade do Minho-Escola de Ciências, 2010.
- FINE, J. P.; GRAY, R. J. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American statistical association*, Taylor & Francis Group, v. 94, n. 446, p. 496–509, 1999.
- GRAY, B. *Subdistribution Analysis of Competing Risks*. [S.l.], 2014. Version 2.2-7. Disponível em: <<https://cran.r-project.org/web/packages/cmprsk>>.
- HONDA, H. et al. Serum albumin, c-reactive protein, interleukin 6, and fetuin a as predictors of malnutrition, cardiovascular disease, and mortality in patients with esrd. *American Journal of Kidney Diseases*, Elsevier, v. 47, n. 1, p. 139–148, 2006.
- JACKSON, C. H. et al. Multi-state models for panel data: the msm package for r. *Journal of Statistical Software*, v. 38, n. 8, p. 1–29, 2011.
- KALBFLEISCH, J. D.; PRENTICE, R. L. Relative risk (cox) regression models. *The Statistical Analysis of Failure Time Data, Second Edition*, Wiley Online Library, p. 95–147, 2002.

- KAPLAN, E. L.; MEIER, P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, Taylor & Francis, v. 53, n. 282, p. 457–481, 1958.
- KLEIN, J. P.; MOESCHBERGER, M. L. Statistics for biology and health. *Stat. Biol. Health, New York*, v. 27238, 1997.
- KLEINBAUM, D. G.; KLEIN, M. *Survival analysis: a self-learning text*. [S.l.]: Springer Science & Business Media, 2006.
- LAWLESS, J. F. *Statistical models and methods for lifetime data*. New York: John Wiley & Sons, 2011. v. 362.
- PECOITS-FILHO, R. et al. Peritonitis as a risk factor for long-term cardiovascular mortality in peritoneal dialysis patients: the case of a friendly fire? *Nephrology*, Wiley Online Library, 2016.
- PUTTER, H.; FIOCCO, M.; GESKUS, R. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in medicine*, Wiley Online Library, v. 26, n. 11, p. 2389–2430, 2007.
- PUTTER, H.; WREEDE, L. de; FIOCCO, M. *Data Preparation, Estimation and Prediction in Multi-State Models*. [S.l.], 2016. Version 0.2.9. Disponível em: <<https://CRAN.R-project.org/package=mstate>>.
- R CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2016. Disponível em: <<https://www.R-project.org/>>.
- ROBINS, J. M.; ROTNITZKY, A. Recovery of information and adjustment for dependent censoring using surrogate markers. In: *AIDS Epidemiology*. [S.l.]: Springer, 1992. p. 297–331.
- SCHOENFELD, D. Partial residuals for the proportional hazards regression model. *Biometrika*, JSTOR, p. 239–241, 1982.
- SESSO, R. C. et al. Inquérito brasileiro de diálise crônica 2014. *J. bras. nefrol*, v. 38, n. 1, p. 54–61, 2016.
- SILVA, G. L. D. F. d.; THOMÉ, E. G. d. R. Complicações do procedimento hemodialítico em pacientes com insuficiência renal aguda: intervenções de enfermagem. *Revista gaúcha de enfermagem. Porto Alegre. Vol. 30, n. 1 (mar. 2009), p. 33-39*, 2009.
- SILVA, H. G.; SILVA, M. J. da. Motivações do paciente renal para a escolha a diálise peritoneal ambulatorial contínua. *Revista Eletrônica de Enfermagem*, v. 5, n. 1, 2006.
- THERNEAU, T. M. *A Package for Survival Analysis in S*. [S.l.], 2016. Version 2.38-3. Disponível em: <<http://CRAN.R-project.org/package=survival>>.
- THERNEAU, T. M.; GRAMBSCH, P. M. *Modeling survival data: extending the Cox model*. [S.l.]: Springer Science & Business Media, 2000.
- YU, L. Diálise em uti: indicações e princípios. *Zilberstrin B. Manual de terapia intensiva: procedimentos práticos*. São Paulo: Robel, p. 385–400, 1995.

## Apêndices

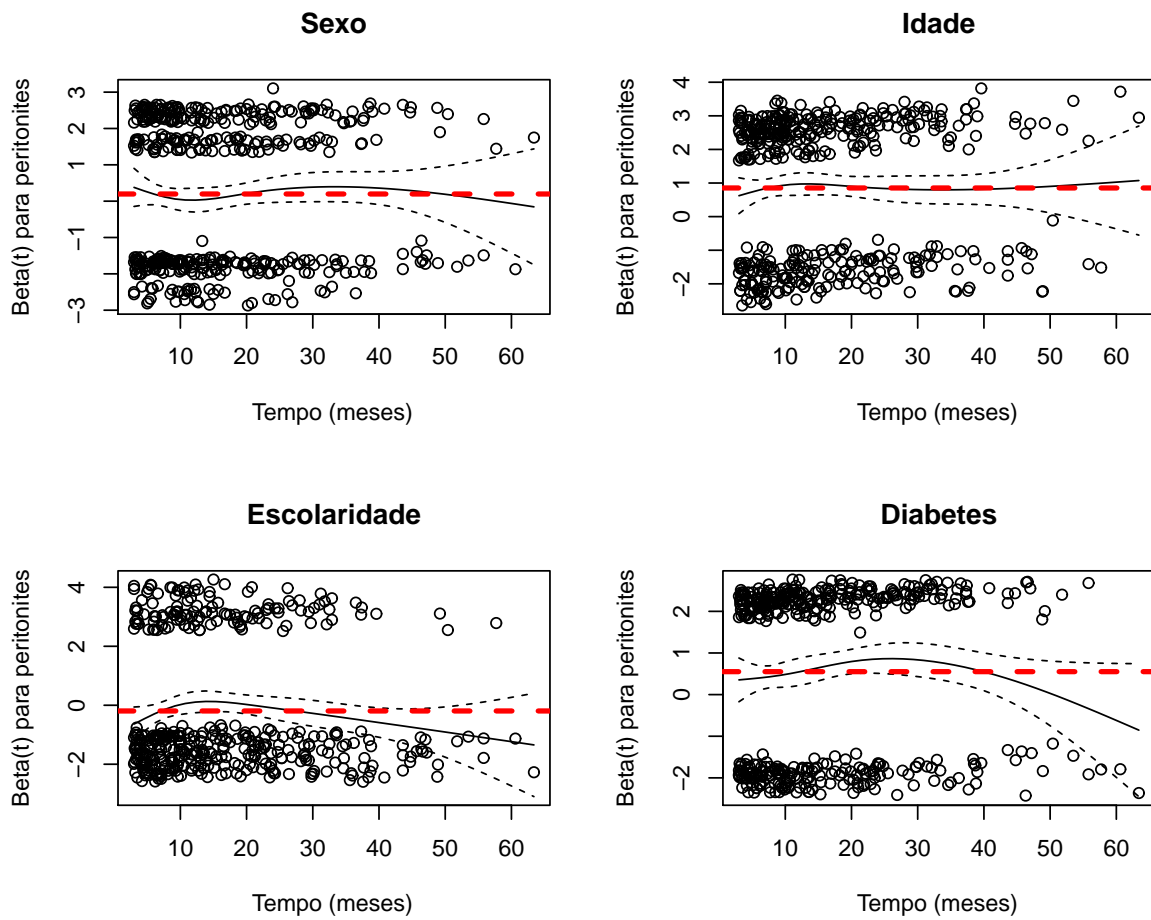
## APÊNDICE A - Gráficos de resíduos de Schoenfeld

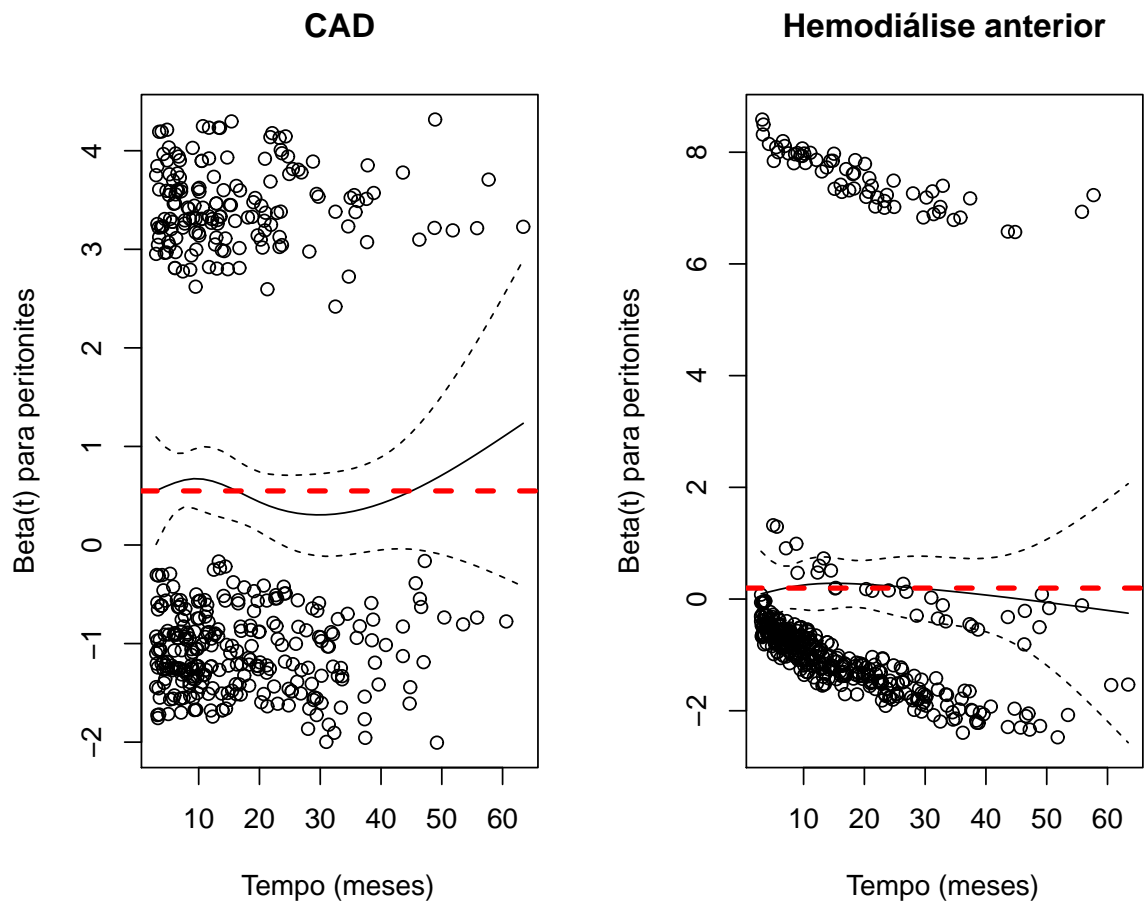
Figura 10 – Resíduos para o modelo de Cox estendido considerando variável peritonites mudando no tempo



Fonte: Elaborado pelos autores (2017).

Figura 11 – Verificação da suposição de riscos proporcionais para as covariáveis utilizando os resíduos de Schoenfeld





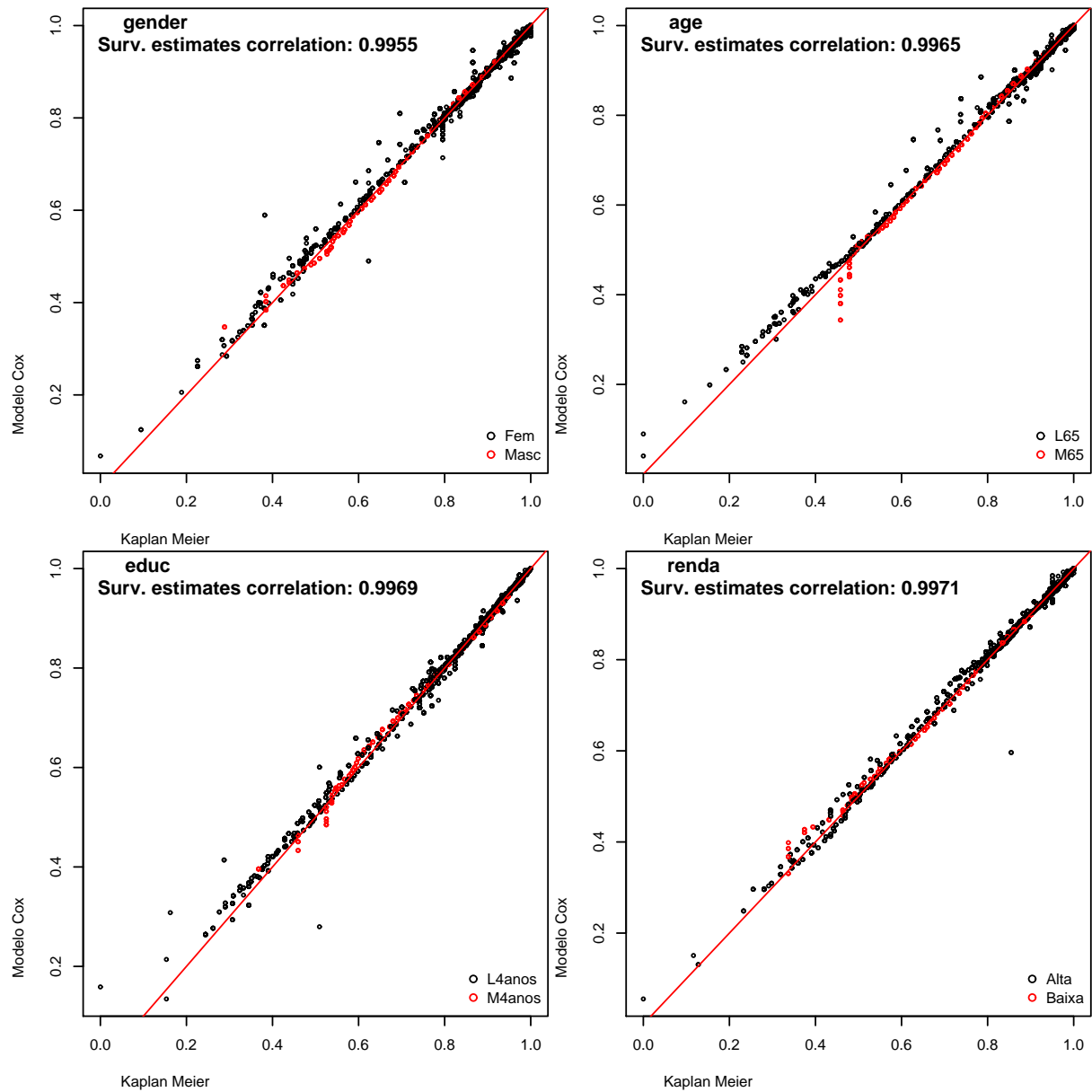
CAD: Doença arterial coronariana

Fonte: Elaborado pelos autores (2017).



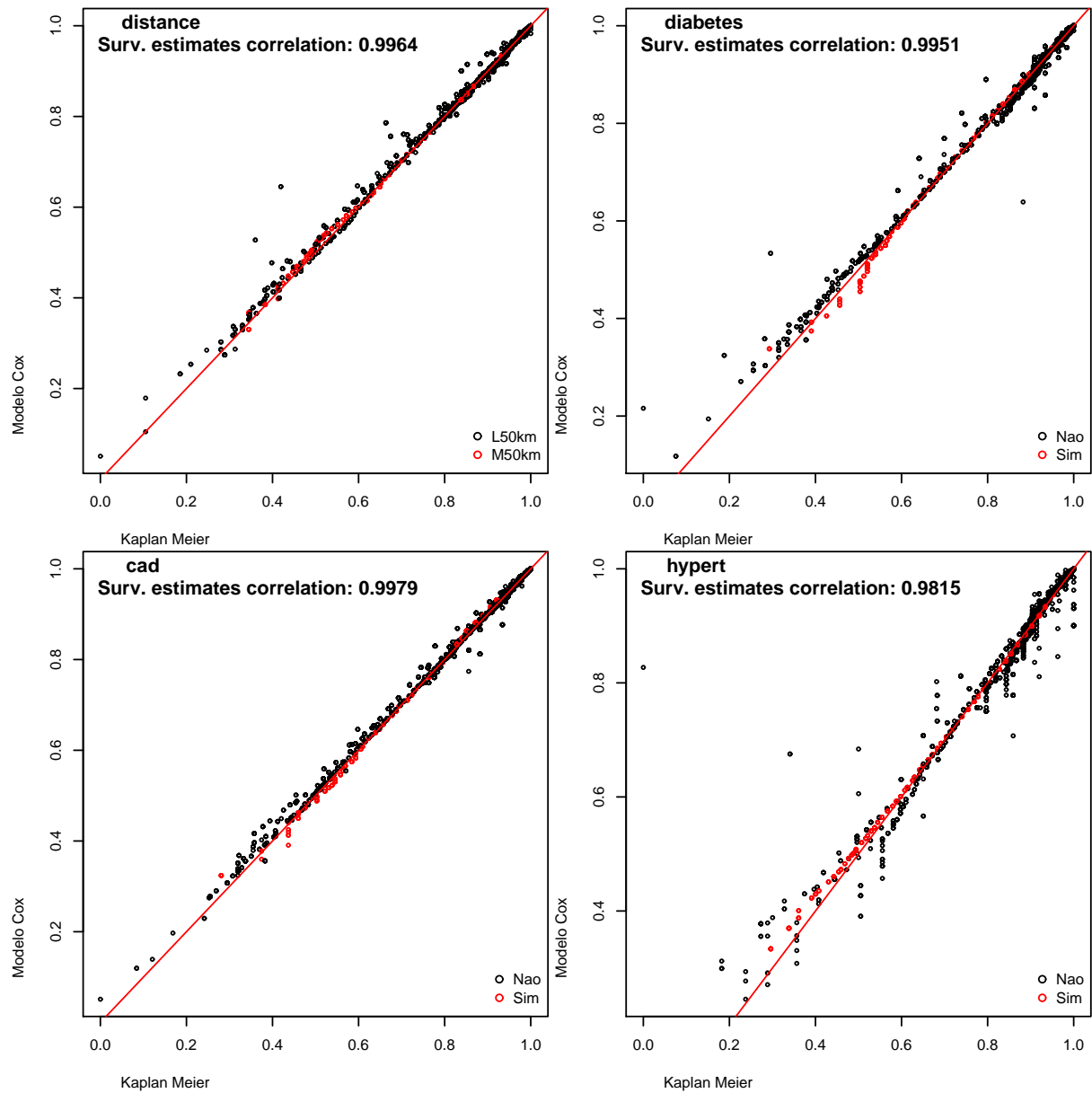
## APÊNDICE B - Gráficos de Kaplan-Meier versus Modelo de Cox

Figura 12 – Gráficos de correlações entre estimativas Kaplan-Meier e Cox para as variáveis sexo, idade, escolaridade e renda



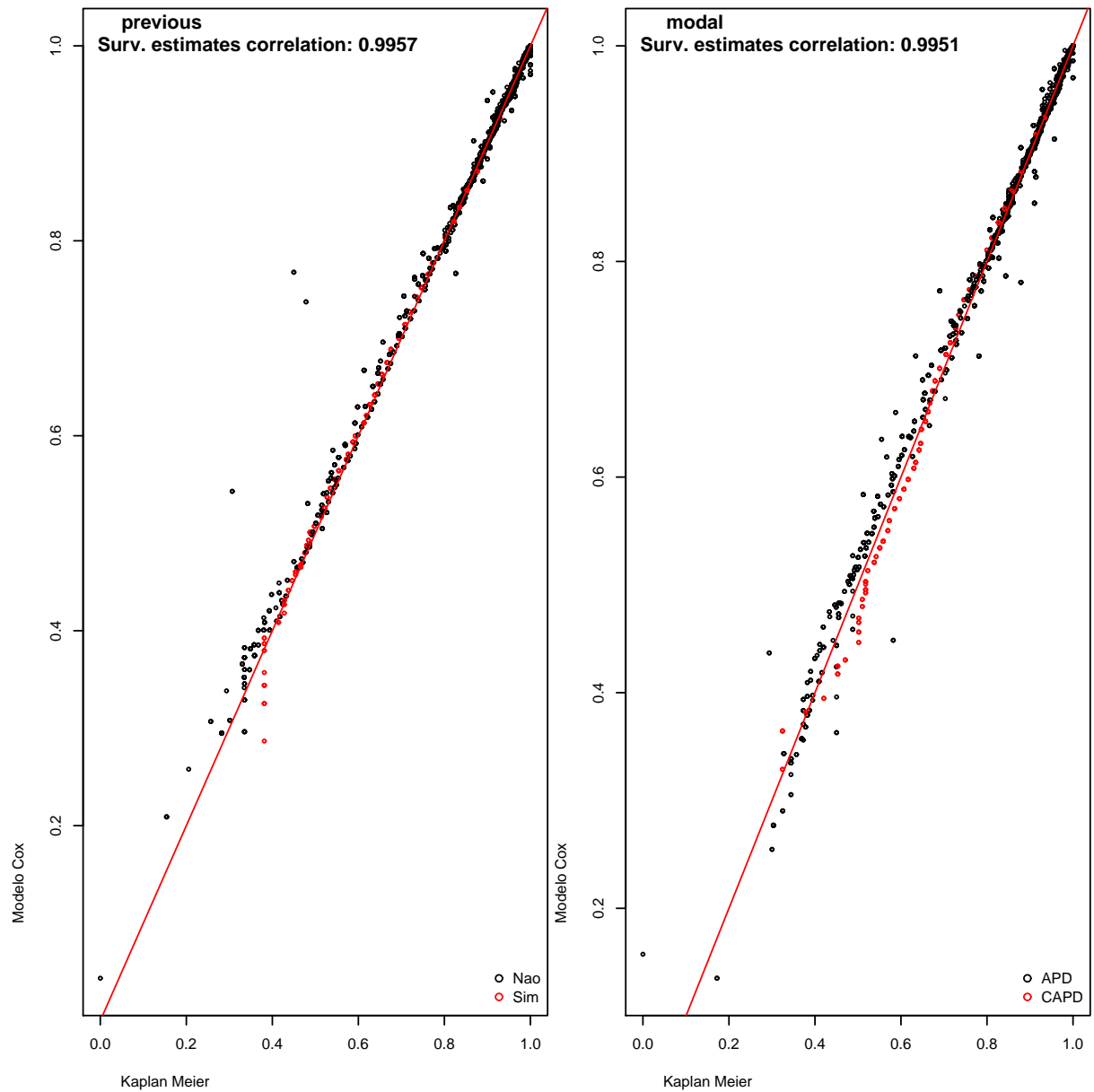
Fonte: Elaborado pelos autores (2017).

Figura 13 – Gráficos de correlações entre estimativas Kaplan-Meier e Cox para as variáveis distância até o centro, diabetes, doença arterial coronariana e hipertensão



Fonte: Elaborado pelos autores (2017).

Figura 14 – Gráficos de correlações entre estimativas Kaplan-Meier e Cox para as variáveis hemodiálise anterior e modalidade da terapia



Fonte: Elaborado pelos autores (2017).