

O EFEITO DAS CARACTERÍSTICAS DOS SEGURADOS SOBRE O TEMPO DE GERAÇÃO DO PRIMEIRO CUSTO ASSISTENCIAL: UM ESTUDO DE CASO EM UMA OPERADORA DE PLANO DE SAÚDE

THE EFFECT OF INSURED CHARACTERISTICS ON THE GENERATION TIME OF THE FIRST ASSISTANCE COST: A CASE STUDY IN A HEALTH CARE OPERATOR

EL EFECTO DE LAS CARACTERÍSTICAS DE LOS ASEGURADOS SOBRE EL TIEMPO DE GENERACIÓN DEL PRIMER COSTO DE ASISTENCIA: UN ESTUDIO DE CASO EN UN OPERADOR DE SALUD

**Jamillys dos Santos Silva**

Universidade Federal da Paraíba  
jamillys.atuaria@gmail.com

**Luiz Santos**

Universidade Federal da Paraíba  
luiz.atuario@gmail.com

**Marcelo Coelho de Sá**

Universidade Federal de Sergipe  
mcs.atuarial@gmail.com



Este é um artigo de acesso aberto distribuído sob os termos da Creative Commons Attribution License  
This is an open-access article distributed under the terms of the Creative Commons Attribution License  
Este es un artículo de acceso abierto distribuido bajo los términos de la Creative Commons Attribution License

## RESUMO

**Objetivo:** Analisar o efeito das características dos segurados sobre a geração do primeiro custo assistencial num plano de saúde. **Método:** O estudo de caso possui caráter quantitativo e descritivo e utiliza dados longitudinais e secundários disponibilizados por uma operadora de saúde que optou pelo anonimato. Aplica-se o Modelo de Cox às informações mensais (entre 2013 e 2015) relativas a 160.572 beneficiários para estimar riscos e probabilidades em função de algumas covariáveis. **Fundamentação teórica:** o tempo até a geração do referido custo é uma variável aleatória que pode apresentar informações incompletas (censuras), indicando-se o uso da análise de sobrevivência para o seu tratamento. **Resultados:** Na ausência de riscos competitivos, o sexo masculino e o procedimento diárias e taxas apresentam menores riscos de gerar custos; na presença, o sexo masculino apresenta maior risco de gerar custos oriundos de diárias e taxas, materiais e medicamentos e outros. **Conclusões:** o método utilizado é considerado adequado para estimativa dos riscos relativos aos custos assistenciais, dada a sua aderência aos dados.

**Palavras-chave:** Saúde Suplementar. Análise de sobrevivência. Riscos competitivos.

## ABSTRACT

**Objective:** To analyze the effect of insured characteristics on the generation of the first care cost in a health plan. **Method:** The case study has a quantitative and descriptive character and uses longitudinal and secondary data provided by a health care provider who opted for anonymity. The Cox Model is applied to monthly information (between 2013 and 2015) relating to 160,572 beneficiaries to estimate risks and probabilities based on some covariates. **Theoretical basis:** the time until the cost is generated is a random variable that may present incomplete information (right-censoring), indicating the use of survival analysis for its treatment. **Results:** In the absence of competitive risks, males and daily procedures and fees have lower risks of incurring costs; in their presence, males are at greater risk of incurring costs from daily rates and fees, materials and medications, and others. **Conclusions:** the method used is considered adequate to estimate the risks related to care costs, given its adherence to the data.

**Keywords:** Supplementary Health. Survival analysis. Competitive risks.

## RESUMEN

**Objetivo:** Analizar el efecto de las características aseguradas en la generación del primer costo de atención en un plan de salud. **Método:** El estudio de caso tiene un carácter cuantitativo y descriptivo y utiliza datos longitudinales y secundarios proporcionados por un proveedor de salud que optó por el anonimato. El Modelo de Cox se aplica a la información mensual (entre 2013 y 2015) de 160.572 beneficiarios para estimar riesgos y probabilidades en base a algunas covariables. **Fundamento teórico:** el tiempo hasta que se genera el costo es una variable aleatoria que puede presentar información incompleta (censura), indicando el uso de análisis de supervivencia para su tratamiento. **Resultados:** En ausencia de riesgos competitivos, los hombres y los procedimientos y honorarios diarios tienen menores riesgos de incurrir en costos; en su presencia, los hombres corren un mayor riesgo de incurrir en costos de tarifas y tarifas diarias, materiales y medicamentos, y otros. **Conclusiones:** el método utilizado se considera adecuado para estimar los riesgos relacionados con los costos de atención, dada su adherencia a los datos.

**Palabras-clave:** Salud suplementaria. Análisis de supervivencia. Riesgos competitivos.

## INTRODUÇÃO

O sistema de saúde no Brasil, direito oferecido a todos os cidadãos, é composto por duas modalidades: pela saúde pública e pela saúde privada – saúde suplementar e serviços autônomos (CONASS, 2015). No âmbito da saúde suplementar, especifica-se nosso objeto de estudo: as operadoras de planos de saúde.

A formalização da saúde suplementar no Brasil se deu por meio da Constituição Federal de 1988 e permitiu que as empresas passassem a oferecer planos de saúde com assistência médica paralelamente ao Sistema Único de Saúde (SUS). Assim, ela se caracteriza por ofertar de forma privada planos e seguros de saúde que disponibilizam assistência médica aos seus beneficiários (CONASS, 2011).

A expansão desse segmento gerou a necessidade de regulação e, nesta perspectiva, teve seu marco regulatório estabelecido pela Lei nº 9.656/98 (BAHIA; VIANA, 2002). A sua normatização, inicialmente prestada pelo Conselho de Saúde Suplementar (CONSU) e posteriormente pela Agência Nacional de Saúde Suplementar (ANS), objetiva nortear todas as partes envolvidas (ofertante e demandante), abarcando quatro medidas: as condições de acesso, a cobertura assistencial, os aspectos econômico-financeiros e a fiscalização, não cabendo a ANS tarifar os planos de saúde (FENASAÚDE, 2014).

O setor tem apresentado ao longo dos anos uma ampliação em sua cobertura. Por exemplo, houve um aumento de 2.143.467 beneficiários entre dezembro de 2010 e dezembro de 2019 (ANS, 2020). Além disso, trata-se de um setor representativo economicamente e, em especial, para a arrecadação do Governo, pois representou 13,49% dentre os setores de maiores gastos tributários no Brasil em 2019 (LARA, 2020). As operadoras e os seguros de saúde associados à Federação Nacional de Saúde Suplementar (FENASAÚDE), ainda, custearam, entre setembro de 2018 e setembro de 2019, R\$ 68,5 bilhões em eventos na assistência médica de seus beneficiários e arrecadaram contraprestações de R\$ 83,1 bilhões (FENASAÚDE, 2019). Ademais, empregaram 3,6 (70,4%) das 5,1 milhões de pessoas que trabalhavam no setor de saúde em dezembro de 2019 (IESS, 2020).

Postos os expressivos valores, trata-se, a partir daqui, do desafio referente à definição da mensalidade (prêmio) a ser cobrada pela operadora, pois a “subprecificação” pode levar à insolvência e a “sobreprecificação” pode levar à perda de competitividade e de segurados (GOOVAERTS et al., 1986). Tal precificação pode ser pré ou pós estabelecida e tem como base a observação de seus custos assistenciais (MACIEL JR., 2016b). Em síntese, conforme os autores, a base que dá origem ao prêmio neste segmento deve ser estruturada de forma equilibrada e deve levar em consideração todos os fatores que impactam os custos assistenciais, além das restrições impostas pela Agência Nacional de Saúde Suplementar (ANS).

Os custos assistenciais são os custos das operadoras de saúde oriundos da utilização médico-hospitalar por parte do segurado, ou seja, correspondem às despesas das operadoras com consultas médicas, exames, terapias, atendimentos ambulatoriais, internações e demais despesas assistenciais. Os estudos sobre tais custos normalmente se reportam a sua composição junto aos preços cobrados nos planos (MACIEL JR., 2016b; LEAL E MATOS, 2009; RAMALHO, 2002), aos reajustes de preços e Variação de Custos Médico-Hospitalar – VCMH (MACIEL JR., 2016a; IPEA, 2006; IESS, 2019), além da relação entre o perfil do usuário e os gastos dos planos (SÁ, 2012; KANAMURA E VIANA, 2007; IESS, sd).

Em alinhamento com Sá (2012) e Kanamura e Viana (2007) – dado que os prêmios dependem dos custos assistenciais e que estes são afetados pelas características dos segurados – pergunta-se: qual o efeito que as características dos segurados exerce sobre os custos assistenciais de uma operadora de saúde?

Desta forma, objetiva-se analisar o efeito das características dos segurados sobre a geração de custos assistenciais de um plano de saúde, a partir de uma amostra de 160.572 segurados vinculados a uma operadora de saúde e observados de 2013 a 2015. Para isso, utilizou-se do Modelo de Cox, na ausência e presença de riscos competitivos, para estimar o risco de “o segurado gerar o Primeiro Custo

Assistencial (PCA) ao plano”.

O Modelo de Cox é um modelo de sobrevivência que pode ser aplicado tanto aos casos que apresentam único desfecho, ou seja, eventos que só acontecem uma vez durante o período de acompanhamento, tal qual a morte, quanto aos casos denominados eventos múltiplos, relativos a eventos que podem acontecer mais de uma vez para um mesmo indivíduo, como é o caso das gestações, internações, infartos, etc. (CARVALHO et al., 2011).

Os segurados aportarão recursos ao plano de saúde para saldarem compromissos futuros, que, a rigor, estarão sujeitos a fatores decrementais: envelhecimento, estágios de morbidez, morte, etc. Isto gera um ambiente em que o segurado estará permanentemente submetido a probabilidades conjuntas que atuam sobre seu status e reduzem suas chances de permanecer “saúdável e não gerador de custo assistencial” (RODRIGUES, 2008).

Neste sentido, é fundamental que a operadora mensure uma matriz de transição que seja capaz de informar os riscos e as probabilidades de seus segurados abandonarem o status inicial de “saúdáveis e não geradores de custos assistenciais” e migrarem para o status de “geradores de custos”. Mais que isso, é fundamental separar essas informações por cada uma das causas que concorrem para retirá-lo de seu status inicial – as inúmeras morbidades e os tipos de procedimento requeridos –, resultados obtidos pelo Modelo de Cox em sua abordagem de Riscos Competitivos.

Assim, este trabalho adiciona à literatura a aplicação dos modelos de sobrevivência, especificamente do Modelo de Cox em ambiente de riscos competitivos, para obtenção de riscos e probabilidades relativos à primeira mudança de status do segurado num plano de saúde. O tempo até a ocorrência do PCA pode expressar, isoladamente, a maturidade do plano; seus riscos e probabilidade, em conjunto com as severidades, são imprescindíveis na formação de provisões e preços.

O presente artigo é composto por esta introdução (contexto, problema, objetivo e justificativa), pelo referencial teórico (precificação dos planos e estudos anteriores), pelos aspectos metodológicos (abordagem, finalidade, procedimentos), pela análise e discussão dos resultados (descritiva, sobrevivência na ausência e presença de riscos competitivos relativos ao tempo até a geração do PCA) e pelas conclusões (limites, sugestões e fechamento).

## REFERENCIAL TEÓRICO

Esta seção apresenta o arcabouço teórico relacionado a presente pesquisa, sendo constituída por dois subtópicos: questões de precificação dos planos e resultados de trabalhos que trataram dos custos assistenciais em planos de saúde.

### **Precificação em saúde suplementar**

A precificação dos planos de saúde – que pode ser pré ou pós estabelecida – tem como base a observação de seus custos, ou seja, das utilizações dos beneficiários dos serviços médico-hospitalares. O preço pós-estabelecido corresponde ao preço calculado após a realização dos custos com as coberturas contratadas; não exige, portanto, cálculo atuarial. Já o preço pré-estabelecido corresponde ao preço calculado antes da utilização das coberturas contratadas, exigindo, assim, cálculo atuarial (MACIEL JR., 2016a).

Em relação aos pré-estabelecidos, tem-se que os preços comerciais (aqueles efetivamente cobrados aos segurados) são compostos por carregamentos administrativos, comerciais, lucro e prêmio puro. O prêmio puro, por sua vez, é formado pela margem de segurança estatística e pelo prêmio de risco. O prêmio de risco nada mais é do que a média anual das despesas assistenciais.

Os preços na saúde suplementar são segmentados em dez faixas etárias. Em decorrência da

existência de pacto intergeracional, o preço da décima faixa não pode ser superior a seis vezes o preço da primeira faixa etária. Essa e outras limitações estão dispostas na Resolução Normativa ANS nº 63/03 (ANS, 2003).

Simplificadamente, o processo de precificação do plano de saúde é constituído por quatro etapas (MACIEL JR., 2016b):

- 1º Passo - Cálculo do Prêmio de Risco (PR): consiste em calcular a frequência de utilização e a severidade média para cada uma das dez faixas etárias e para cada uma das despesas assistenciais agregadas (consultas, exames, terapias, internações, outros atendimentos ambulatoriais e demais despesas assistenciais).
- 2º Passo - Cálculo do Prêmio Puro (PP): por ser atuante num mercado de risco elevado, o plano é obrigado pela ANS a calcular e a aplicar a Margem de Segurança Estatística (MSE). O PP é formado pela soma entre a MSE e o PR.
- 3º Passo - Calcular o Prêmio Comercial (PC): após o cálculo de todos os PP, é necessário incluir as parcelas de carregamentos comerciais, de carregamentos administrativos e da margem de Lucro, pelo menos.
- 4º Passo - Alinhamento do PC: valor que ajusta o prêmio comercial e serve para atender aos interesses gerais (comerciais, jurídicos, etc.) da operadora.

Evidencia-se que a presente pesquisa contribui junto ao 1º passo do processo de precificação descrito, que usualmente precifica o produto no ano corrente com base na frequência observada. A construção da matriz de transição incrementa ao permitir realizar projeções para as frequências. Neste sentido, o presente trabalho não trata da severidade.

Dada a exposição do processo de precificação, e do destaque dado à projeção da frequência, elucida-se o papel basilar dos custos assistenciais na formação do preço num plano de saúde. Tendo em mente que os diferentes perfis dos segurados podem exercer efeitos significativos sobre os referidos custos, apresentam-se, a seguir, os resultados de alguns trabalhos que analisaram os custos em saúde.

### Estudos anteriores

Aldridge e Kelley (2015), a partir de dados de 2011 do MEPS – um conjunto de pesquisas em larga escala de famílias e indivíduos nos EUA realizada pela *National Health Expenditure Accounts e National Center for Health Statistics* –, estimaram as despesas geradas pela população estudada. Assim, identificaram que pessoas com perfil de alto custo assistencial podem ser encaixadas em três grupos: no primeiro, estão as pessoas que passaram por uma situação que gerou alto custo e voltaram a gerar custos mais baixos (como, por exemplo, um indivíduo que sofreu ataque isquêmico vascular); no segundo, estão as pessoas que possuem doenças crônicas e geram custos anuais elevados; e o terceiro, constituído por pessoas que geram custos elevados por estarem no fim de suas vidas.

De forma semelhante, o IESS (sd) objetivou caracterizar os beneficiários brasileiros de maior custo assistencial, com suporte em dados de 76,1 mil beneficiários no ano de 2015. Dentre seus principais resultados, observa-se que o custo assistencial pode estar atrelado ao perfil do beneficiário, onde o grupo de interesse representa os 5% (3.808 pessoas) da população que gera os custos mais altos. Destes, 56,8% são mulheres responsáveis por 66,5% (R\$ 354,9 milhões) da despesa assistencial total do grupo investigado; a idade avançada mostra uma relação direta com os valores elevados de despesa, logo, quanto maior a idade do beneficiário, maior o custo assistencial; a doença crônica corresponde a outro fator que eleva consideravelmente o custo assistencial, tendo em vista que indivíduos que possuem essas doenças estão mais propensos a serem internados (como a internação nesse grupo em 2015 foi responsável por 92% da despesa, essa característica se torna mais preocupante que a idade).

Ainda pensando nos custos e em como eles se propagam ao longo do tempo, Leal e Matos (2009) fizeram um estudo sobre a evolução dos custos assistenciais no período de 2002 a 2006, a partir de dados

do Sistema de Informações de Produtos (SIP) e de um índice de valor que compara a multiplicação dos preços e quantidades no período  $t$  com os respectivos valores no período 0. Desta forma, foi feita uma análise empírica em duas etapas: na primeira, comparou-se a evolução de custos médios com as despesas assistenciais por beneficiário, enquanto na segunda foi feita uma avaliação de custo médio (custos por expostos). Assim, concluiu-se que a elevação dos custos assistenciais é explicada predominantemente pelo aumento do custo médio dos eventos (associados à inflação por insumos e ao avanço tecnológico, em conformidade com Sá (2012)) e não pelo aumento da frequência (associada a fatores sociais, demográficos e epidemiológicos).

Destaca-se que faz sentido pensar, no curto prazo, no aumento do custo médio dos eventos como causa do aumento de custos assistenciais. No médio e longo prazo, entretanto, o envelhecimento certamente será a principal causa, justificando a necessidade de elaboração de ferramentas que prevejam o agravamento das probabilidades de transição dos status elencados no presente trabalho.

Além disso, o IESS criou a variável índice Variação de Custos Médico-Hospitalares (VCMH) para estudar os custos assistenciais. Ele capta a variação de preços e a frequência de utilização dos beneficiários de planos de saúde (LARA; REIS; MINAMI, 2018).

A variação nos custos também depende do tipo de serviço utilizado pelo beneficiário (IESS, 2019). Por exemplo, em dezembro de 2011, os gastos do setor foram compostos por consultas (9,2%), exames (6,0%), terapias (10,7%), Outros Serviços Ambulatoriais – OSA (11,8%), internação (13,5%) e outros (85,8%), gerando uma VCMH de 12,9%; em dezembro de 2018, os gastos em percentuais foram de 12,3% em consultas, 9,9% em exames, 31,3% em terapias, 19,7% em OSA, 16,5% em internação e 16,9% em outros, com VCMH igual a 17,3%.

Exposto o esforço central de grande parte dos trabalhos em analisar a composição dos custos médico-hospitalares por meio da observação de sua variação monetária, reitera-se a contribuição da presente pesquisa, que analisa a composição dos referidos custos por meio da observação de suas frequências. A literatura não apresenta, de forma exclusiva, tal tratamento. Mais que isso, não se utiliza da metodologia aqui proposta, isto é, da análise de sobrevivência. Isso fica claro quando se buscam, sem sucesso, trabalhos que combinem tal objeto e método.

Análise de sobrevivência é um ramo da análise estatística que estuda a variável contínua “tempo de falha”, ou seja, o tempo até a ocorrência de um evento de interesse. Ela se difere das demais análises porque inclui em seus modelos a “censura”, que são informações incompletas que influenciam nos resultados apresentados (COLOSIMO; GIOLO, 2006).

A aplicação desse método em custos não é comum. Foi realizada uma busca extensiva nos sites (e fontes) “SciELO”, “Periódicos CAPES”, “Google Acadêmico”, “Libgen” e “Ciência & Saúde Coletiva”, pelos termos “sobrevivência em custos”, “sobrevivência em despesas”, “sobrevivência em custos”, “sobrevivência em despesas” e “costs survival analysis”. A seguir, comenta-se o resultado da busca.

Jackman et al. (2017), a partir de dados referentes a 160 pacientes *prepathway* (antes da intervenção) e 210 *postpathway* (depois da intervenção) e com o propósito de saber se a implementação de “*Clinical Pathways*” em pacientes (no estágio IV de câncer, do instituto Dana-Farber) impactaria no custo e na sobrevivência, i) utilizaram o estimador da função de sobrevivência de Kaplan-Meier para detectar os padrões de sobrevivência nos indivíduos e ii) estimaram a função de risco por meio de uma modificação da regressão de Cox, levando-se em conta o efeito de algumas covariáveis (sexo, raça, perfil genérico, distância entre residência e clínica, além da clínica utilizada). Na determinação do custo de implementação foi feita a comparação entre o custo total médio e o custo de cada categoria principal de serviço, por meio de testes e modelos de regressão múltipla ajustada às covariáveis demográficas e clínicas descritas. Dado isso e apesar das limitações do trabalho, notou-se por meio de análise de regressão que o custo foi de U\$\$ 69.122 antes da intervenção e U\$\$ 52.037 depois da intervenção, obtendo-se uma economia de U\$\$ 17,085 mil.

## ASPECTOS METODOLÓGICOS

Nesta seção são apresentadas as diretrizes gerais sobre os procedimentos necessários para a realização da pesquisa, ou seja, apresenta o tipo de pesquisa, o universo investigado, o tipo de coleta, as variáveis, o tipo de análise (análise de sobrevivência), o modelo (modelo de Cox) e os pacotes em R utilizados.

Esse estudo de caso, de caráter quantitativo, descritivo, bibliográfico e documental, faz uso de dados longitudinais e secundários disponibilizados por uma operadora de saúde que optou pelo anonimato. A escolha pela operadora, bem como pelo período analisado, se deu por conveniência, dado que informações sobre frequência e severidade não são comumente divulgadas. Aqui, tratam-se de informações mensais relativas a 160.572 beneficiários no período de 2013 a 2015, listadas no Quadro I.

**Quadro I - Variáveis disponibilizadas e suas especificações**

Variáveis	Tipo	Níveis
Data de nascimento	Quantitativa	-
Data de adesão ao plano	Quantitativa	-
Data de primeiro uso do plano	Quantitativa	-
Custo do primeiro uso do plano (em R\$)	Qualitativa	Baixo – valores menores ou iguais a R\$108,56; Moderado – valores entre R\$108,56 e R\$453,53; Alto – valores maiores ou iguais a R\$453,72.
Idade	Quantitativa	-
Tipo de procedimento	Qualitativa	Materiais e medicamentos, Honorário, Diárias e taxas, Outros, SADT.
Sexo	Qualitativa	Feminino e masculino.

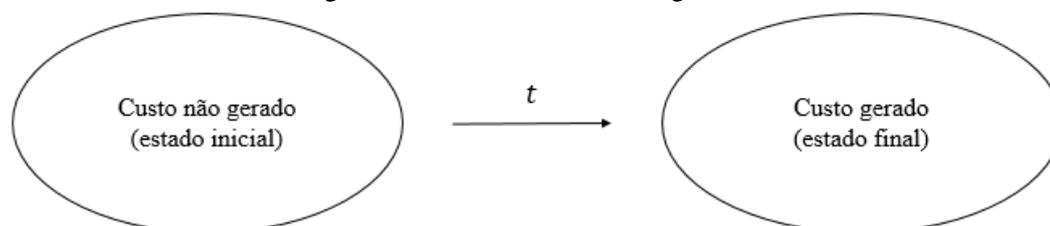
Fonte: Elaboração própria através de dados de uma operadora que optou pelo anonimato (2021).

Em análise de sobrevivência é necessário que sejam definidas a falha e a censura (quando houver). A falha corresponde à ocorrência do evento de interesse, enquanto a censura se refere a uma informação incompleta sobre a ocorrência. Deste modo, a falha no presente estudo corresponde à geração do PCA por parte do segurado, enquanto a censura se apresentaria caso o período delimitado de estudo se encerrasse e o segurado não tivesse gerado custo ao plano até aquela data, o que não ocorreu.

A variável resposta em análise de sobrevivência é o tempo de falha  $T$ , ou seja, o tempo que leva até que o indivíduo gere o PCA ao plano. Ela é definida pela subtração entre a data de uso e a data de adesão ao plano.

Optou-se pela realização de duas análises de sobrevivência sobre a variável aleatória  $T$ , uma chamada de análise padrão e outra chamada de análise de riscos competitivos. A análise padrão considera um contexto simplificado em que se observa, num  $T = t$ , se a falha ocorreu ou não, ou seja, se houve a transição de um estado inicial ( $J = 0$ ) para apenas um outro ( $J = 1$ ), conforme a Figura I. Isso significa que essa abordagem permite analisar se houve ou não a geração de custos assistenciais, sem a diferenciação de desfechos.

**Figura I - Matriz de transição da geração do PCA sob a abordagem da análise de sobrevivência clássica**

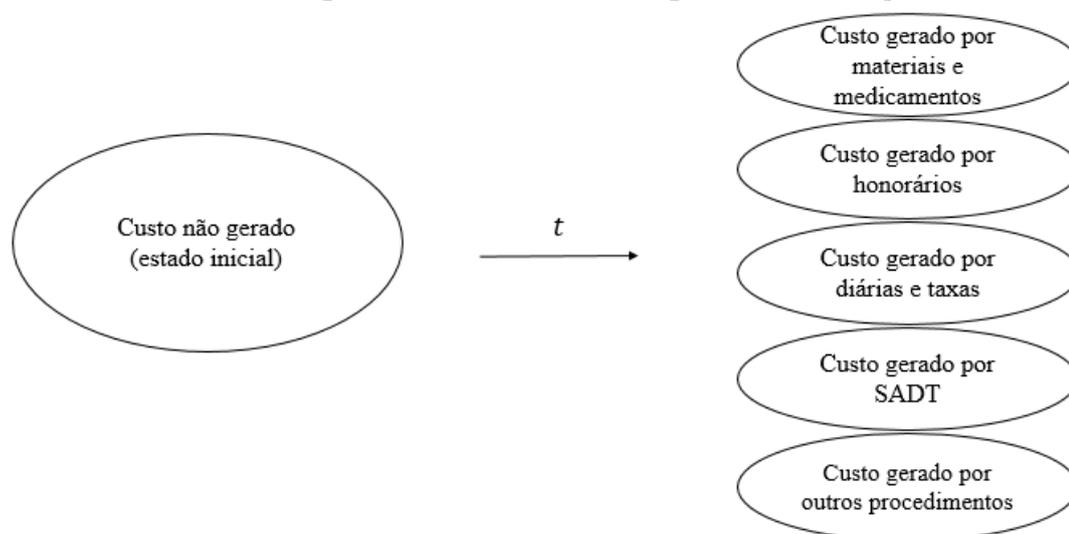


Fonte: elaboração própria (2021).

A análise de riscos competitivos torna o contexto um pouco mais complexo, visto que, transcorridos  $t$  períodos, permite que haja transição de um estado inicial ( $J = 0$ ) para outros estados, conforme a Figura 2. Isto significa que sua matriz de transição é maior que a da análise de sobrevivência padrão e permite considerar diferentes desfechos ou causas concorrentes que tendem a retirar a unidade observada do seu status inicial. Assim, essa abordagem permite analisar a geração de custos assistenciais por tipo de procedimento utilizado pelo segurado.

Quanto aos níveis da variável Tipo de Procedimento, “materiais e medicamentos” (matmed) englobam todos os materiais e medicamentos utilizados nos procedimentos por pessoa na sua primeira utilização do plano; “honorários” (hono) são relativos aos honorários médicos, honorários não médicos, procedimentos terapêuticos, hemoterapia e gasoterapia; “diárias e taxas” (dt) correspondem às diárias, taxas e internações em casa; serviços de apoio ao diagnóstico e terapias (sadt), que é um termo autoexplicativo; os demais procedimentos foram agrupados em “Outros”.

Figura 2 - Matriz de transição da geração do PCA sob a abordagem de riscos competitivos



Fonte: elaboração própria (2021).

Tipo de procedimento especifica o tipo de despesa gerada ao plano em caso de usufruto do beneficiário, ou seja, se o custo foi proveniente ao pagamento de matmed, hono, dt, sadt ou outros procedimentos. Observa-se que o tipo procedimento assume dois papéis distintos na análise: na ausência de riscos competitivos, é uma covariável que pode explicar o comportamento da variável resposta “tempo de falha”; na presença, corresponde a riscos específicos, ou seja, a cinco causas distintas e concorrentes a gerar o PCA.

Assim, na ausência de riscos competitivos, os níveis “sexo”, “idade”, “custo” e “tipo de procedimento” são covariáveis que estimam o risco de geração de PCA; na presença de riscos competitivos, apenas “sexo”, “idade” e “custo” são covariáveis que estimam o risco de geração do PCA, já que procedimento redefine as causas de falha.

Em relação ao tipo de análise, realizam-se as análises descritivas e de modelagem por meio da análise de sobrevivência. Assim, tem-se que a nossa variável de interesse é investigada por meio de duas funções: a função de sobrevivência e a função de risco. A função de sobrevivência, segundo Herrmann (2011), é a probabilidade de uma observação sobreviver ou não falhar até um certo tempo  $t$ . Aqui, representa a probabilidade de o segurado não gerar o PCA até um instante  $t$ . É representado pela equação (I):

$$S(t) = P(T > t), t \geq 0. \quad (1)$$

A função de risco, por sua vez, é a taxa instantânea de ocorrência do evento de interesse no tempo condicionada a sua não ocorrência anterior a  $t$  (CARVALHO et al., 2011), conforme expressão (2):

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \quad (2)$$

Na ausência de riscos competitivos, de acordo com Colosimo e Giolo (2006), a função de sobrevivência é comumente estimada pelo estimador de Kaplan-Meier (1958), que considera tanto o intervalo de tempo quanto o número de falhas e é escrito conforme (3):

$$\hat{S}(t) = \prod_{j:t_j < t} \left(1 - \frac{d_j}{n_j}\right). \quad (3)$$

$t_1 < t_2 \dots < t_k$ , os  $k$  são tempos distintos e ordenados de falha;  $d_j$  é o número de falhas em  $t_j$ ,  $j = 1, \dots, k$ ; e  $n_j$  é o número de indivíduos sob risco em  $t_j$ ; Desta forma, será estimado o  $\hat{q}_j$ , que é a probabilidade do indivíduo gerar o PCA no intervalo de tempo  $[t_{j-1}, t_j)$ .

O modelo semiparamétrico de Cox (1972), que exige proporcionalidade de riscos, estima a função de risco por meio da expressão (4):

$$\lambda(t | x) = \lambda_0(t) g(x' \beta). \quad (4)$$

$g(x' \beta)$  é o componente paramétrico, uma função não-negativa que deve ser especificada;  $\lambda_0(t)$  é o componente não paramétrico, uma função não negativa do tempo e não especificada (COLOSIMO; GIOLO, 2006).

Por fim, pretende-se construir uma tábua biométrica intitulada “geração do PCA” por meio de técnicas de eventos competitivos e levando-se em conta as diferentes formas de utilização do plano (decrementos) pelo usuário: materiais e medicamentos, honorário, diárias e taxas, diagnóstico e terapias e outros procedimentos.

Na presença de riscos competitivos, conforme Carvalho et al. (2011), é possível modelar a função de risco pelos métodos Risco Específico por Causa e Risco de Subdistribuição. Por sua simplicidade, utiliza-se, aqui, daquele.

O método Risco Específico por Causa assume que o risco por cada causa (tipo de procedimento) segue o modelo (5).

$$\lambda_{0j;i}(t; Z_i) = \lambda_{0j;0}(t) \times \exp(\beta_{0j} \times Z_i), \quad j = 1, 2, \dots, k \quad e \quad i = 1, 2, \dots, n \quad (5)$$

em que  $\beta_{0j}$  é um vetor  $1 \times p$  de coeficientes de regressão,  $Z_i$  é um vetor  $p \times 1$  de covariáveis para indivíduo  $i$  e  $\lambda_{0j;0}(t)$  é uma função de risco de linha basal não especificada e não negativa.

Em relação a testes e procedimentos relativos à seleção de modelos e verificação de pressupostos, utiliza-se das seguintes ferramentas: a) o teste de Logrank; b) a comparação gráfica entre as sobrevivências estimadas pelos modelos de Cox ajustados e os valores estimados empiricamente por Kaplan-Meier; c) o Critério de Informação de Akaike (AIC); d) o Desvio Quadrático Médio (DQM); e) Teste de Schoenfeld.

O teste de Logrank verifica se as sobrevivências estimadas pelo estimador de Kaplan-Meier por níveis de cada covariável são iguais. A estatística deste teste é a diferença entre o número observado de

falhas em cada grupo e uma quantidade correspondente ao número esperado de falhas sob a hipótese nula (MANTEL, 1966 apud COLOSIMO; GIOLO, 2006).

Para verificar se o modelo de Cox ajustado é adequado ou não, é possível realizar um teste subjetivo, isto é, comparar graficamente as curvas estimadas pelos modelos de Cox com àquela resultante do estimador de Kaplan-Meier, considerando-se adequado(s) o(s) modelo(s) de Cox parecido(s) com aqueles estimados por Kaplan-Meier.

O Critério de Informação de Akaike (AIC) é uma medida de distância entre o modelo ajustado e um teórico “modelo mais adequado”, representado por  $AIC_p = -2 \ln f(x|\hat{\theta}) + 2k$ . A verossimilhança é inversa e diretamente proporcional com os termos  $-2 \ln f(x|\hat{\theta})$  e  $2k$ , respectivamente. Dessa forma, AIC pondera a adequação aos dados e a complexidade do modelo (SOBRAL; BARRETO, 2016), de modo que quanto menor o AIC, melhor é o modelo.

O Desvio Quadrático Médio (DQM) permite avaliar a distância entre os dados observados ( $d_o$ ) e esperados ( $d_e$ ). Definida como  $DQM_t = [(d_o - d_e)/d_e]^2$ , quanto menor o DQM, mais aderente é o ajuste (PINHEIRO, 2007).

O Teste de Schoenfeld verifica a suposição de riscos proporcionais no modelo de Cox. Desta forma  $q_i$  é dado por (6)

$$r_{qi} = x_{qi} - \frac{\sum_{j \in R_{(ti)}} x_{qj} \exp\{x'_{j\hat{\beta}}\}}{\sum_{j \in R_{(ti)}} \exp\{x'_{j\hat{\beta}}\}}. \quad (6)$$

com  $k \leq n$ ;  $t_1 < t_2 < \dots < t_k$ ;  $x_i = (x_{1i}, x_{2i}, \dots, x_{pi})$  e  $r_i = (r_{1i}, r_{2i}, \dots, r_{pi})$  (SCHOENFELD, 1982, apud COLOSIMO; GIOLO, 2006).

A análise é realizada em ambiente R, viabilizada pelos pacotes survival (THERNEAU, 2015), mvna (ALLIGNOL; BEYERSMANN; SCHUMACHER, 2008) e etm (ALLIGNOL; SCHUMACHER; BEYERSMANN, 2011).

## ANÁLISE E DISCUSSÃO DE RESULTADOS

### Análise descritiva

A Tabela I mostra as estatísticas descritivas referentes ao tempo necessário para que o segurado gere o PCA, à idade do segurado e ao (primeiro) custo gerado pelo segurado. É possível observar, por exemplo, que o menor dos tempos é I, ou seja, algum dos usuários utilizou o plano já no primeiro mês de cobertura.

Tabela I - Estatísticas descritivas referentes ao tempo que leva para o indivíduo gerar o PCA e a idade do indivíduo na respectiva ocorrência - 2013 a 2015

Variável	Mínimo	Máximo	Média	Mediana	3rd Qu
Tempo (meses)	I	31,40	17,49	16,17	24,40
Idade (anos)	I	115,5784	41,1626	37,2868	60,0602
Custo (R\$)	0,02	53857,40	108,56	70,00	113,50

Fonte: elaboração própria (2021).

Destaca-se que levou em média 17 meses para que os segurados do plano analisado gerassem o PCA e cujos valores monetários médios representaram R\$ 108,56. A variável custo, para fins de regressão,

foi categorizada, sendo composta por 3.571 casos considerados como de custo alto, 119.774 como de custo baixo e 37.227 como de custo moderado.

Para fins descritivos, optou-se por categorizar a idade, conforme indica a Resolução Normativa ANS N° 63/03, com o objetivo de verificar se há ou não diferença dos tempos de falha entre as faixas etárias: Faixa 1 (de 0 a 18 anos); Faixa 2 (19 a 23 anos); Faixa 3 (de 24 a 28 anos); Faixa 4 (de 29 a 33 anos); Faixa 5 (de 34 a 38 anos); Faixa 6 (de 39 a 43 anos); Faixa 7 (de 44 a 48 anos); Faixa 8 (de 49 a 53 anos); Faixa 9 (de 54 a 58 anos); Faixa 10 (59 anos ou mais). A Tabela 2 apresenta a descritiva do tempo de falha por faixa etária.

**Tabela 2 - Estatísticas descritivas do tempo (em meses) de geração do PCA por faixa etária - 2013 a 2015**

Faixa etária	Média	Desvio padrão	Coefficiente de variação
0 a 18 anos	17,21	7,41	43,05%
19 a 23 anos	17,05	7,48	43,89%
24 a 28 anos	17,55	7,56	43,10%
29 a 33 anos	17,71	7,73	43,66%
34 a 38 anos	18,15	7,84	43,20%
39 a 43 anos	17,71	7,77	43,84%
44 a 48 anos	17,50	7,81	44,65%
49 a 53 anos	17,32	7,83	45,20%
54 a 58 anos	16,97	7,83	46,13%
59 anos ou mais	17,56	7,93	45,19%

Fonte: elaboração própria (2021).

De qualquer modo, a fim de verificar tal “igualdade”, optou-se pela comparação de médias (teste t) e de variância (teste F) do tempo até a geração do PCA entre as diferentes faixas etárias. Supondo-se normalidade dos dados devido ao tamanho amostral, realizaram-se os testes duas a duas (faixas etárias) e se considerou um nível de significância de 5%. Assim, tem-se que em relação à média houve diferença estatística em 34 das 45 combinações possíveis, ou seja, 34 comparações de faixas etárias apresentaram diferenças significativas quanto ao tempo médio necessário para a ocorrência do primeiro custo. Quanto à variância, 24 das 45 combinações apontaram diferença estatística.

Complementa-se que a despeito dos resultados dos testes estatísticos, em termos práticos e de gestão, tem-se que as diferenças do tempo até a ocorrência do PCA entre as faixas etárias não são expressivas e que o processo de envelhecimento não implica na antecipação do PCA. Em outros termos, o estudo revela que em média não há expressiva diferença do tempo para o PCA entre jovens e idosos.

Contabilizam-se, ainda, o número de primeiros custos assistenciais gerados ao plano por tipo de procedimento. De um total de 160.572, 97.229 casos se referem ao pagamento de honorários, 33.842 casos de SADT, 13.178 de outros procedimentos, 9.323 de materiais e medicamentos e 7.000 diárias e taxas. Ademais, o banco de dados é formado por 93.207 mulheres e 67.365 homens.

### **Análise de sobrevivência na ausência de riscos competitivos**

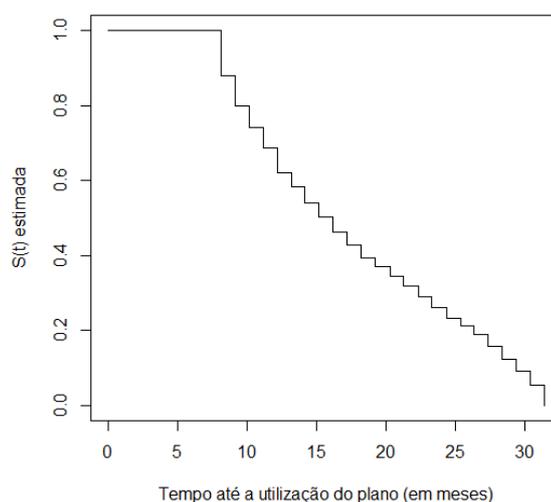
Na análise não paramétrica, estimam-se as funções de sobrevivência relativas ao grupo de segurados a) sem segmentação do grupo e b) segmentando-se o grupo por níveis de variáveis categóricas (sexo, tipo de procedimento e custo categorizado). Em seguida, averigua-se a igualdade de curvas entre os níveis de covariáveis comparados via teste de logrank.

Na análise semiparamétrica ajustam-se (via regressão de Cox) e selecionam-se (via AIC e DQM) modelos para estimar o risco de gerar o PCA. Assim, interpreta-se o efeito das covariáveis (sexo, idade,

tipo e custo de procedimento) sobre o referido risco. Além disso, verifica-se o pressuposto de proporcionalidade dos riscos por meio do teste de Schoenfeld.

Neste sentido, a Figura 3 apresenta as funções de sobrevivência estimadas pelo estimador não paramétrico de Kaplan-Meier para o conjunto de segurados. Para o grupo investigado, desconsiderando-se a existência de um período de carência, as pessoas utilizaram o plano pela primeira vez em até 31 meses depois do início do acompanhamento (1/1/2013).

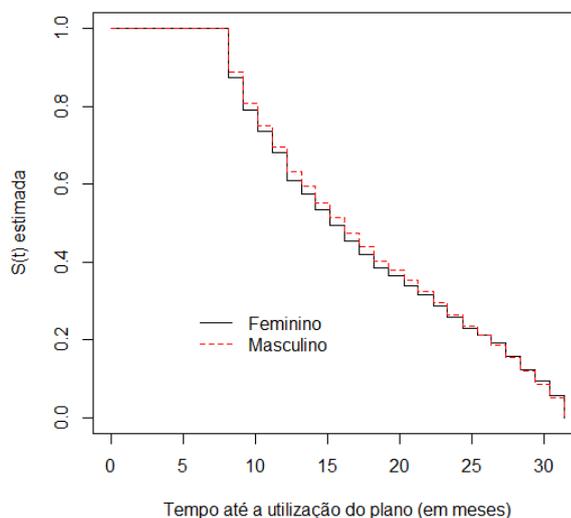
Figura 3 - Sobrevivência estimada por Kaplan-Meier para o tempo até o segurado gerar o PCA ao plano



Fonte: elaboração própria (2021).

A Figura 4 apresenta as funções de sobrevivência estimadas pelo estimador não paramétrico de Kaplan-Meier para cada categoria de sexo. Nota-se, de um modo geral, que as curvas de sobrevivência estimadas para homens e mulheres parecem não diferir entre si, apesar do indício, no recorte dos primeiros 15 meses, de que a mulher gera o PCA antes que o homem.

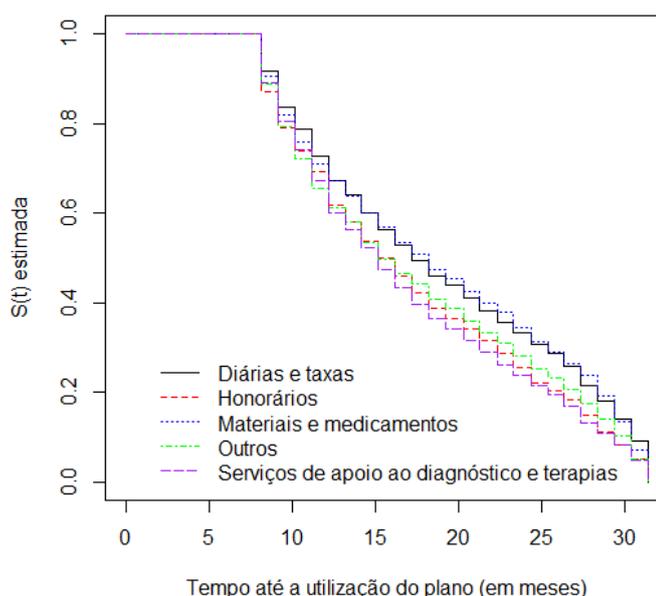
Figura 4 - Sobrevivência estimada por Kaplan-Meier, por níveis de Sexo, para o tempo até PCA



Fonte: elaboração própria (2021).

A Figura 5 apresenta as funções de sobrevivência estimadas pelo estimador de Kaplan-Meier para cada tipo de procedimento. Nota-se que algumas das curvas de sobrevivência estimadas por tipo de procedimento parecem diferir entre si. Por exemplo, Diárias e taxas e Materiais e medicamentos parecem ter sobrevivências similares entre si e divergentes daquelas apresentadas por Honorários, SADT e Outros. Isso significa que é mais provável que o PCA do segurado seja oriundo de algum procedimento pertencente ao segundo grupo.

**Figura 5 - Sobrevivência estimada por Kaplan-Meier, por níveis de Procedimento, para o tempo até o PCA**



Fonte: elaboração própria (2021).

A Figura 6 apresenta as funções de sobrevivência estimadas pelo estimador não paramétrico de Kaplan-Meier para cada categoria de custo. Assim, parece não haver diferença significativa entre as sobrevivências das diferentes categorias. Apesar disso, é possível observar que a partir do vigésimo mês de cobertura a probabilidade de o PCA gerado ser baixo é maior que as probabilidades de ser moderado ou alto.

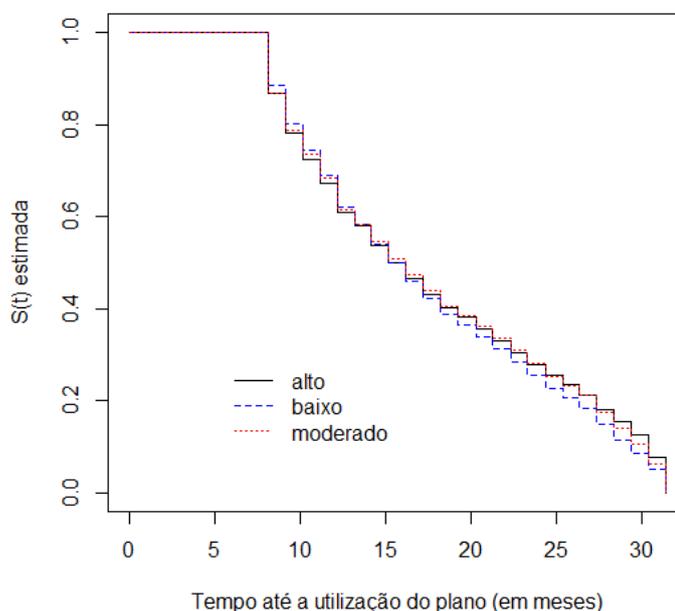
Para complementar a análise das Figuras 4, 5 e 6, realiza-se o teste de logrank, assumindo-se um nível de significância de 5%. Como os valores p foram menores ou iguais a 5%, rejeitam-se as hipóteses de que homens e mulheres possuem sobrevivências estimadas iguais, de que os tipos de procedimentos possuem sobrevivências estimadas iguais e que os tipos de custos possuem sobrevivências iguais. Como para Tipo de procedimento e Custos, que apresentam mais de dois níveis, rejeitam-se as hipóteses de igualdade de sobrevivência, foram realizados testes de logrank para cada combinação possível (dois a dois) de seus níveis. Conforme a Tabela 3, a única dupla que não diferiu foi “diárias e taxas” e “materiais e medicamentos”.

Em relação ao custo, conforme a Tabela 4, a única dupla que não diferiu quanto às sobrevivências estimadas foi “Moderado” e “Alto”.

Na análise semiparamétrica foi utilizado o modelo de Cox para estimar a função de risco relativa à geração do PCA por parte do segurado, considerando quatro covariáveis (sexo, idade, tipo de procedência e custo). Realizaram-se ajustes para todas as combinações possíveis dessas covariáveis, ou seja, um total de 16 ajustes, que vai do modelo composto somente pelo intercepto ao modelo completo (com

as quatro covariáveis). Para a seleção de modelos, opta-se pelo uso de critérios objetivo e subjetivo, dados pela Tabela 5 e Figura 7, respectivamente.

**Figura 6 - Sobrevivência estimada por Kaplan-Meier, por níveis de Custo, para o tempo até o PCA**



Fonte: elaboração própria (2021).

**Tabela 3 - Resultados dos testes logrank utilizados para as comparações das sobrevivências estimadas, relativas ao tempo até o segurado gerar o PCA, por níveis de Procedimento**

Níveis comparados	Valor p
Diárias e taxas x Honorários	<0,001
Diárias e taxas x Materiais e medicamentos	0,3
Diárias e taxas x outros	<0,001
Diárias e taxas x SADT	<0,001
Honorários x Materiais e medicamentos	<0,001
Honorários x Outros	<0,001
Honorários x SADT	<0,001
Materiais e medicamentos x Outros	<0,001
Materiais e medicamentos x SADT	<0,001
Outros x SADT	<0,001

Fonte: elaboração própria (2021).

**Tabela 4 - Resultados dos testes logrank utilizados para as comparações das sobrevivências estimadas, relativas ao tempo até o segurado gerar o PCA, por níveis de Custo**

Níveis de custos comparados	Valor p
Moderado x Alto	0,2
Baixo x Alto	<0,001
Moderado x Baixo	<0,001

Fonte: elaboração própria (2021).

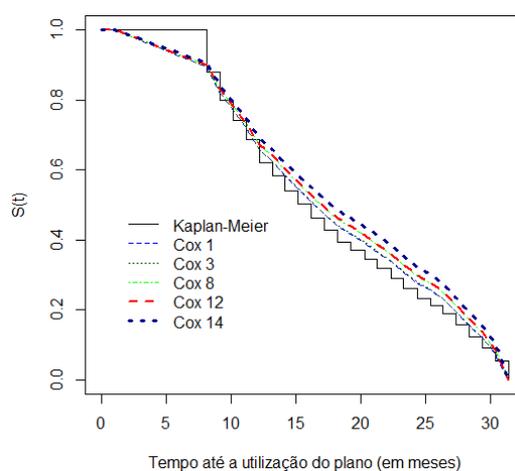
Tabela 5 - Resultados de AIC e DQM dos cinco melhores modelos ajustados para explicar o risco de geração do PCA por segurado do plano

Modelo ajustado	AIC	DQM
Cox I	3527458	1,89%
Cox 3	3527460	2,02%
Cox 8	3527333	3,39%
Cox I2	3527336	3,57%
Cox I4	3527459	5,59%

Fonte: elaboração própria (2021).

Desta forma, a Tabela 5 apresenta os valores de AIC e DQM dos cinco melhores modelos, ou seja, aqueles com menores valores das referidas métricas: Cox I, Cox 3, Cox 8, Cox I2 e Cox I4.

Figura 7 - Sobrevivências estimadas por modelos de Cox para o tempo até o PCA



Fonte: Elaboração própria (2021).

A Figura 7 compara as sobrevivências estimadas pelos modelos de Cox àquela estimada por Kaplan-Meier: quanto mais próxima a curva de Cox em relação a de Kaplan-Meier, melhor o modelo ajustado. Por meio de sua observação, nenhum dos modelos pré-selecionados destoam da sobrevivência estimada pelo estimador de Kaplan-Meier. No entanto, dentre eles, Cox I e Cox 3, sobrepostos neste caso, parecem estimar valores mais próximos daquele a partir do 15º mês de cobertura. Assim, opta-se pela apresentação dos resultados referentes a dois modelos: Cox I (menor DQM) e Cox 8 (menor AIC), indicando-se o uso de ambos os modelos para  $t \leq 15$  meses e apenas Cox I para  $t > 15$  meses. Seus resultados são exibidos na Tabela 6.

Tabela 6 - Resultados dos modelos de Cox I e Cox 8 ajustados para o tempo até a ocorrência de PCA

Variável – Nível	$e^{\beta}$	
	Cox I	Cox 8
Sexo – Masculino	0,989434*	0,9881*
Idade	0,998674*	0,9988*
Tipo de procedência – Honorários	1,223950*	1,221*
Tipo de procedência - Materiais e medicamentos	1,020620	1,016
Tipo de procedência - Outros	1,168610*	1,184*
Tipo de procedência - SADT	1,274831*	1,288*
Custo baixo	-	1,069*
Custo moderado	-	1,000

Fonte: Elaboração própria (2021).

Em relação ao modelo Cox I, verifica-se que as covariáveis sexo, idade e tipo de procedência exercem efeitos significativos sobre a função de risco estimada, dados que os valores p são inferiores ao nível de significância de 5%. Seus efeitos sobre o risco são constatados na coluna  $e^{\hat{\beta}}$ : o homem apresenta menos risco que a mulher de gerar o PCA (em  $1 - 0,989434 = 0,010566 = 1,05\%$ ); como  $e^{\hat{\beta}} < 1$ , o risco de gerar o referido custo tende a diminuir com o aumento da idade; em relação ao tipo procedência, como  $e^{\hat{\beta}} > 1$  para todos os níveis explicitados, todos eles apresentam maior risco de gerar o PCA que Diárias e taxas.

Em relação a Cox 8, verifica-se que as covariáveis sexo, idade, tipo de procedência e custo (categorizado) exercem efeitos significativos sobre a função de risco estimada, dados que os valores p são inferiores ao nível de significância de 5%. Seus efeitos sobre o risco são constatados na coluna  $e^{\hat{\beta}}$ : o homem apresenta menos risco que a mulher de gerar o PCA (em  $0,9881 - 1 = |-0,0119| = 1,19\%$ ); como  $e^{\hat{\beta}} < 1$ , o risco de gerar o referido custo tende a diminuir com o aumento da idade; em relação ao tipo de procedência, como  $e^{\hat{\beta}} > 1$  para todos os níveis explicitados (com exceção de materiais e medicamentos, cujo efeito não é significativo), todos eles apresentam maior risco de falha em relação ao nível implícito, qual seja, Diárias e taxas; tem-se ainda que o risco de o baixo custo gerar o PCA é maior que o de alto custo (em  $1,069 - 1 = 0,069 = 6,9\%$ ).

Observa-se, ainda, que os efeitos da idade sobre o risco de geração de PCA em ambos os modelos são muito próximos de 1, ou seja:  $e^{\beta} \cong 1 \rightarrow \beta = \ln(1) = 0$ . Isto significa que em termos práticos o efeito da idade sobre o risco não é significativo.

Verificou-se, ainda, a suposição de proporcionalidade de riscos dos modelos Cox I e Cox 8 por meio do Teste de Schoenfeld. Seus resultados são apresentados na Tabela 7 e indicam rejeição de proporcionalidade dos riscos de ambos os modelos, já que o valor p global é inferior ao nível de significância de 5%. Analisando-se variável por variável, tem-se que sexo e idade rejeitam a proporcionalidade, enquanto que alguns níveis de Tipo de procedência (em Cox I e 8) e de Custo (em Cox 8) não rejeitam.

**Tabela 7 - Teste de Schoenfeld para proporcionalidade de riscos dos Modelos de Cox I e Cox 8 ajustados para o tempo que leva o segurado a gerar o PCA**

Variável – Nível	$\chi$	
	Cox I	Cox 8
Sexo – Masculino	103,86*	101,52*
Idade	204,63*	184,66*
Tipo de procedência – Honorários	1,29	1,96
Tipo de procedência - Materiais e medicamentos	6,61*	5,59*
Tipo de procedência – Outros	7,82*	5,18*
Tipo de procedência – SADT	2,75	2,08
Custo baixo	-	12,43*
Custo moderado	-	3,12*
GLOBAL	364,51*	395,67*

Fonte: Elaboração própria (2021).

A princípio, a não assunção da premissa indicaria a não adequação do modelo ao tipo de dado. No entanto, de acordo com Hjort (1992), um modelo de sobrevivência mal especificado é consistente, não para esse parâmetro, mas para o que ele chama de parâmetro menos falso. Com base nessa ideia, opta-se pelo uso do modelo especificado e pela interpretação de seus efeitos sobre o risco estimado de geração do PCA.

### Análise de sobrevivência na presença de riscos competitivos

Na Análise de sobrevivência na presença de riscos competitivos se estimam semiparametricamente as funções de risco específico por causa, relativas à geração do PCA, levando-se em conta o efeito de duas covariáveis (sexo e idade). Aqui as cinco causas que concorrem para retirar o segurado de seu status inicial de “não gerador de PCA” são as gerações de custo oriundas de a) Materiais e medicamentos, b) Honorários, c) Diárias e taxas, d) Outros, e) SADT.

A Tabela 8 exibe os resultados de cinco modelos referentes aos efeitos de sexo, idade e custo sobre cada uma das cinco causas específicas.

Deste modo, em relação ao risco específico “Tipo de procedência – Diárias e taxas”, tem-se que sexo, idade e tipo de custo geram efeitos significativos (já que os valores p são inferiores a 5%) sobre o risco de gerar o PCA pelo uso de Diárias e taxas: o risco do homem é 16,55% maior que o da mulher; o aumento da idade tende a diminuir esse risco; e o custo moderado é menos arriscado que o custo alto em 46,77%.

Em relação ao risco específico “Tipo de procedência – Honorários”, tem-se que sexo, idade e tipo de custo geram efeitos significativos (já que os valores p são inferiores a 5%) sobre o risco de gerar o PCA pelo uso de Honorários: o risco do homem é 5,33% menor que o da mulher; o aumento da idade tende a diminuir o risco; o custo baixo é 332% mais arriscado e o custo moderado é 225% mais arriscado quando comparados ao custo alto.

Quanto ao risco específico “Tipo de procedência – Materiais e Medicamentos”, tem-se que sexo, idade e tipo de custo geram efeitos significativos sobre o risco de gerar o PCA pelo uso de Materiais e medicamentos: o risco do homem é 5,9% maior que o da mulher; o aumento da idade tende a aumentar o risco; o custo moderado é 54,01% menos arriscado quando comparado ao custo alto.

**Tabela 8 - Resultados do efeito do sexo, idade e custo sobre as variáveis, para o risco do segurado a gerar o PCA**

Risco específico	Variável – Nível	$e^{\hat{\beta}}$	Valor p
Tipo de procedência – Diárias e taxas	Sexo – Masculino	1,165585	<0,0001
	Idade	0,996555	<0,0001
	Custo baixo	0,934082	0,357
	Custo moderado	0,532349	<0,0001
Tipo de procedência – Honorários	Sexo – Masculino	0,9467500	<0,0001
	Idade	0,9975790	<0,0001
	Custo baixo	4,3271617	<0,0001
	Custo moderado	3,2565854	<0,0001
Tipo de procedência - Materiais e medicamentos	Sexo - Masculino	1,0597506	0,00558
	Idade	1,0044649	<0,0001
	Custo baixo	1,1047661	0,11960
	Custo moderado	0,4598646	<0,0001
Tipo de procedência – Outros	Sexo - Masculino	1,2421669	<0,0001
	Idade	0,986075	<0,0001
	Custo baixo	0,3895971	<0,0001
	Custo moderado	0,5853793	<0,0001
Tipo de procedência – SADT	Sexo - Masculino	0,9507245	<0,0001
	Idade	1,0060826	<0,0001
	Custo baixo	0,3456820	<0,0001
	Custo moderado	0,6106059	<0,0001

Fonte: Elaboração própria (2021).

Em relação ao risco específico “Tipo de procedência – outros”, tem-se que sexo, idade e tipo de custo geram efeitos significativos (já que os valores p são inferiores a 5%) sobre o risco de gerar o PCA pelo uso de Outros: o risco do homem é 24,21% maior que o da mulher; o aumento da idade tende a diminuir o risco; o custo baixo é 61,04% menos arriscado e o custo moderado é 41,46% menos arriscado quando comparados ao custo alto.

Quanto ao risco específico “Tipo de procedência – SADT”, tem-se que sexo, idade e tipo de custo geram efeitos significativos (já que os valores p são inferiores a 5%) sobre o risco de gerar o PCA pelo uso de SADT: o risco do homem é 4,93% menor que o da mulher; o aumento da idade tende a aumentar o risco; o custo baixo é 65,43% menos arriscado e o custo moderado é 38,94% menos arriscado quando comparados ao custo alto.

Observa-se novamente que os efeitos da idade sobre o risco de geração de PCA, para todas as causas específicas, são muito próximos de 1, significando em termos práticos que o efeito da idade sobre o risco não é significativo.

Assim, diferentemente de Jackman et al. (2017) – que utilizaram a análise de sobrevivência para estimar efeitos de uma intervenção clínica sobre a vida dos pacientes e os custos hospitalares – e de Sá (2012) – que investigou os beneficiários que geram altos e baixos custos num plano –, o presente trabalho se propôs a estimar o efeito de sexo, idade, serviço e custo demandado pelos segurados, sobre risco de gerar o PCA ao plano.

## CONCLUSÕES

Objetivou-se nesse trabalho, através de dados de uma operadora de saúde que optou pelo anonimato, analisar o tempo até a geração do PCA por segurado, a partir de uma amostra (referente às competências de 2013 a 2015) de tamanho 160.572. A base conta com 93.207 pessoas do sexo feminino e 67.365 do sexo masculino; 7.000 pessoas geraram o PCA pelo uso de “Diárias e Taxas”, 97.229 pessoas por “Honorários”, 9.323 pessoas por “Materiais Medicamentos”, 13.178 pessoas por “Outros” e 33.842 pessoas por “SADT”.

Constatou-se que o tempo médio que leva para que uma pessoa gere o PCA é de aproximadamente 17 meses, cujo valor mínimo é de 1 mês e o valor máximo, de 31 meses; em relação à idade, os segurados que fizeram o primeiro uso do plano têm, em média, 41 anos.

Na ausência de riscos competitivos, a análise não paramétrica, por meio do teste logrank, permitiu constatar que as sobrevivências (probabilidade de o segurado não gerar o PCA) dadas pelo estimador de Kaplan-Meier, por nível de covariáveis (sexo, custo e tipo de procedência), são diferentes; a análise semiparamétrica permitiu estimar, por meio do modelo de Cox selecionado (pelo menor AIC e DMQ), os riscos de geração de PCA, tendo como principal resultado o fato de as covariáveis sexo, idade, tipo de custo e tipo de procedência exercerem efeitos significativos sobre a função de risco estimada: o homem apresenta menos risco que a mulher de gerar o PCA; o risco de gerar o referido custo tende a diminuir com o aumento da idade; em relação ao tipo de procedência, todos os níveis apresentam maior risco de falha em relação às Diárias e taxas; quanto aos Custos, o risco de gerar um baixo PCA é maior quando comparado à categoria “alto”.

Na presença de riscos competitivos, realizou-se somente a análise semiparamétrica. O modelo de riscos específicos por causa estimou os riscos de geração de PCA para cada tipo de procedimento, tendo como principal resultado o fato de as covariáveis sexo, idade e custo exercerem efeitos (conjunto) significativos sobre as funções de risco estimadas. Assim, o homem, em relação à mulher, apresenta maiores riscos de gerar custos assistenciais quando estes são oriundos de Diárias e taxas, Materiais e medicamentos e Outros; o aumento da idade tende a aumentar os riscos de gerar custos assistenciais quando estes são oriundos de Materiais e medicamentos e SADT; o risco de gerar um baixo PCA é maior que um alto

primeiro custo para Honorários e Materiais e medicamentos.

Por fim, elencam-se as limitações deste trabalho e algumas sugestões para pesquisas futuras, a fim de idealizar a realização de pesquisas:

- Recomenda-se a adição de informações sobre os segurados, tais quais a escolaridade, a renda, a profissão, o tipo de plano, etc.
- Além disso, dentro da perspectiva de riscos competitivos, poderiam ser ajustadas tanto as regressões paramétricas, quanto as semiparamétricas de subdistribuições de riscos.
- A depender do tamanho da amostra, e conseqüentemente da viabilidade de convergência das estimativas, poder-se-ia extrapolar a variável resposta “tempo até a ocorrência do PCA por parte do segurado”, incluindo também o segundo, o terceiro, o n-ésimo custo, considerando-se a possibilidade de ampliação da matriz de transição de estados. O modelo de eventos múltiplos seria capaz de modelar essa situação.

Conclui-se que as utilizações dos modelos de sobrevivência podem ser implementadas nas operadoras de saúde e contribuem para que se obtenham informações – probabilidades e riscos – que melhor subsidiem planejamentos, ações e decisão baseadas em frequência.

## REFERÊNCIAS

ANS - AGÊNCIA NACIONAL DE SAÚDE SUPLEMENTAR - ANS. **Resolução Normativa nº 63, de 22 de dezembro de 2003**. Define os limites a serem observados para adoção de variação de preço por faixa etária nos planos privados de assistência à saúde contratados a partir de 1º de janeiro de 2004. Rio de Janeiro, RJ, 22 de dezembro 2003.

ANS - AGÊNCIA NACIONAL DE SAÚDE SUPLEMENTAR. **Beneficiários de planos privados de saúde, por cobertura assistencial**, 2020. Disponível em: <http://www.ans.gov.br/perfil-do-setor/dados-gerais>. Acesso em: 11 jul. 2020.

ALDRIDGE, Melissa D.; KELLEY, Amy S.. **Dying in America: improving quality and honoring individual preferences near the end of life**. 2015. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK285684>. Acesso em: 15 ago. 2019.

ALLIGNOL, A.; BEYERSMANN, J.; SCHUMACHER, M. mvna: an r package for the nelson-aalen estimator in multistate models. **R News**. n. 8, p. 48-50, 2008.

ALLIGNOL, A.; SCHUMACHER, M.; BEYERSMANN, J. Empirical transition matrix of multi-state models: the etm package. **Journal of Statistical Software**, n. 38, p. 1- 5, 2011.

BAHIA, L.; VIANA, A. A regulamentação da assistência médica suplementar. In: BRASIL. Ministério da Saúde (org.). **Regulação e saúde: estrutura, evolução e perspectivas da assistência médica suplementar**. Rio de Janeiro: ANS, 2002.

CARVALHO, Marília Sá et al. **Análise de sobrevivência: Teoria e aplicações em saúde**. 2. ed. Rio de Janeiro: Editora Fiocruz, 2011. 434 p.

COLOSIMO, Enrico Antônio; GIOLO, Suely Ruiz. **Análise de sobrevivência aplicada**. São Paulo: Edgard Blucher, 2006. 392 p.

CONASS – CONSELHO NACIONAL DE SECRETÁRIOS DE SAÚDE. **Saúde suplementar**. São Paulo: CONASS, 2011. Disponível em: [https://www.conass.org.br/bibliotecav3/pdfs/colecao2011/livro\\_12.pdf](https://www.conass.org.br/bibliotecav3/pdfs/colecao2011/livro_12.pdf). Acesso em: 28 jun. 2019.

CONASS – CONSELHO NACIONAL DE SECRETÁRIOS DE SAÚDE. **Saúde suplementar**. São Paulo: CONASS, 2015. Disponível em: [https://www.conass.org.br/biblioteca/pdf/atualizacao-2015/LI2\\_SAUDE-SUPLEMENTAR\\_jun2015.pdf](https://www.conass.org.br/biblioteca/pdf/atualizacao-2015/LI2_SAUDE-SUPLEMENTAR_jun2015.pdf). Acesso em: 01 jul. 2019.

COX, D. R. Regression models and life-tables (with discussion). **Journal of the Royal Statistical Society**, Series B, v. 34, p. 187-220, 1972.

FENASAÚDE - FEDERAÇÃO NACIONAL DE SAÚDE SUPLEMENTAR. **Saúde suplementar: Aspectos regulatórios**. Rio de Janeiro: Fenasaúde, 2014. 44 p. Disponível em: [http://midias.cnseg.org.br/data/files/53/35/05/75/5F81B310ED0901B33D4818A8/saude\\_suplementar\\_aspectos\\_regulatorios%2014.II.II.pdf](http://midias.cnseg.org.br/data/files/53/35/05/75/5F81B310ED0901B33D4818A8/saude_suplementar_aspectos_regulatorios%2014.II.II.pdf). Acesso em: 14 jul. 2019.

FENASAÚDE - FEDERAÇÃO NACIONAL DE SAÚDE SUPLEMENTAR. **Despesa assistencial e receita de contraprestações**. 2019. Disponível em: <http://fenasaude.org.br/estatisticas/economico-financeiro-fenasaude.html>. Acesso em: 11 jul. 2020.

GOOVAERTS, M.; VYLDER, F. de; HAEZENDONCK, J.. **Insurance and risk theory**. Netherlands: Springer, 1986.

HERRMANN, Letícia. **Estimação de Curvas de Sobrevivência para estudos de custo-efetividade**. 2011. 120 f. Monografia (Especialização) - Curso de Estatística, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011.

HJORT, Nils Lid. On Inference in Parametric Survival Data Models. **International Statistical Review**, v.60, n. 3, p. 355-387, 1992.

IESS – INSTITUTO DE ESTUDOS DE SAÚDE SUPLEMENTAR. **Análise Especial: Caracterização dos beneficiários de alto custo assistencial - Um estudo de caso**. sd. Disponível em: [https://www.iess.org.br/cms/rep/estudo\\_especial\\_gastos.pdf](https://www.iess.org.br/cms/rep/estudo_especial_gastos.pdf). Acesso em: 15 ago. 2019.

IESS – INSTITUTO DE ESTUDOS DE SAÚDE SUPLEMENTAR. **Variação de Custos Médico-Hospitalares**. 2019. Disponível em: <https://www.iess.org.br/cms/rep/VCMH.pdf>. Acesso em: 15 ago. 2019.

IPEA – INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **Uma descrição do comportamento dos preços dos planos de assistência à saúde: 2001 a 2005**. Texto para discussão n. 1232 *Boletim de políticas sociais* – acompanhamento e análise n. 10, Rio de Janeiro: Ipea, 2006. Disponível em: [https://www.ipea.gov.br/portal/index.php?option=com\\_content&view=article&id=4379](https://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=4379). Acesso em: 12 mai. 2020.

JACKMAN, David M. et al. Cost and survival analysis before and after implementation of dana-farber clinical pathways for patients with stage IV non-small-cell lung cancer. **Journal Of Oncology Practice**. Boston, p. 346-352, abr. 2017.

KANAMURA, Alberto Hideki; VIANA, Ana Luiza D'ávila. Gastos elevados em plano privado de saúde: com quem e em quê. **Revista de Saúde Pública**, [S.L.], v. 41, n. 5, p. 814-820, out. 2007. Disponível em: <https://www.scielo.br/pdf/rsp/v41n5/5809.pdf>. Acesso em: 24 jul. 2020.

KAPLAN, E. L.; MEIER, P. Nonparametric estimation from incomplete observations. **Journal of the American Statistical Association**, v. 53, p. 457-481, 1958.

LARA, Natalia; REIS, Amanda; MINAMI, Bruno. **A Variação de Custos Médicos Hospitalares (VCMH): um compêndio dos estudos do IESS e uma atualização do tema**. São Paulo: Instituto de Estudos em Saúde Suplementar, 2018. 22 p. Disponível em: [https://www.iess.org.br/cms/rep/TD71\\_CompendioVCMH.pdf](https://www.iess.org.br/cms/rep/TD71_CompendioVCMH.pdf). Acesso em: 15 jul. 2019

LARA, Natalia. Texto para Discussão nº 77 – 2020. **Gasto tributário como ferramenta para o desenvolvimento econômico e social do país**. São Paulo: Instituto de Estudos em Saúde Suplementar, 2020. 13p. Disponível em: <https://www.iess.org.br/cms/rep/td-77-Gastos-Tributarios.pdf>. Acesso em: 11 jul.2020.

LEAL, Rodrigo Mendes; MATOS, João Boaventura Branco de. Planos de saúde: uma análise dos custos assistenciais e seus componentes. **Revista de Administração de Empresas**, [S.L.], v. 49, n. 4, p. 447-458, dez. 2009. Disponível em: <https://www.scielo.br/pdf/rae/v49n4/v49n4a08.pdf>. Acesso em: 23 jul. 2020.

MACIEL JÚNIOR, José Nazareno. **Precificação dos planos de saúde**: apresentando alguns aspectos técnicos envolvidos. Fortaleza, 01 mai. 2016a. LinkedIn: José Maciel Nazareno Júnior. Disponível em: <https://pt.linkedin.com/pulse/precifica%C3%A7%C3%A3o-dos-planos-de-sa%C3%BAde-apresentando-alguns-maciel-junior>. Acesso em: 17 jul. 2020.

MACIEL JÚNIOR, José Nazareno. **Reajuste dos planos de saúde**: apresentando alguns aspectos técnicos envolvidos. Fortaleza, 28 fev. 2016b. LinkedIn: José Maciel Nazareno Júnior. Disponível em: <https://www.linkedin.com/pulse/reajuste-dos-planos-de-sa%C3%BAde-alguns-aspectos-maciel-junior?trk=mp-reader-card>. Acesso em: 17 jul. 2020.

PINHEIRO, Ricardo Pena. **A demografia dos fundos de pensão**. Brasília: Ministério da Previdência Social. Secretaria de Políticas de Previdência Social, 2007. 292 p. – (Coleção Previdência Social. Série estudos; v. 24).

RAMALHO, Ricardo Luiz. **Precificação de planos de saúde**. 2002. 141 f. Dissertação (Mestrado) - Curso de Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2002. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/82655/189741.pdf?sequence=1&isAllowed=y>. Acesso em: 4 jun. 2020.

RODRIGUES, José Angelo. **Gestão de risco atuarial**. São Paulo: Saraiva, 2008.

SÁ, Marcelo Coelho de. **Análise dos custos assistenciais de uma operadora de plano de saúde no Brasil**. 2012. 94 f. Dissertação (Mestrado) - Curso de Engenharia de Produção, Universidade Federal do Rio Grande do Norte, Natal, 2012.

SOBRAL, Thales Esteves Lima; BARRETO, Gilmar. Utilização dos critérios de informação na seleção de modelos de regressão linear. **Proceeding series of the brazilian society of applied and computational mathematics**, São Carlos, v. 4, n. 1, p.1-7, jan. 2016. Disponível em:

<https://proceedings.sbmec.org.br/sbmec/article/view/1144/1157>. Acesso em: 12 mar. 2020.

THERNEAU, T. **A package for survival analysis in S**. 2015. Disponível em: <https://CRAN.R-project.org/package=survival>. Acesso em: 15 mai. 2021.