



UNIVERSIDADE FEDERAL DA PARAÍBA (UFPB)
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS (CCSA)
DEPARTAMENTO DE FINANÇAS E CONTABILIDADE (DFC)
CURSO DE BACHARELADO EM CIÊNCIAS ATUARIAIS (CCA)

JAMILLYS DOS SANTOS SILVA

**ANÁLISE BIOMÉTRICA DE GERADORES DE CUSTO ASSISTENCIAL: UM
ESTUDO DE CASO**

JOÃO PESSOA, PB

2020

JAMILLYS DOS SANTOS SILVA

**ANÁLISE BIOMÉTRICA DE GERADORES DE CUSTO ASSISTENCIAL: UM
ESTUDO DE CASO**

Trabalho de Conclusão de Curso (TCC) para o curso de Ciências Atuariais na UFPB, como requisito parcial à obtenção do título de bacharel em Ciências Atuariais.

Área de concentração: Saúde Suplementar.

Orientador: Prof. Dr. Luiz Carlos Santos Júnior.

Coorientador: Prof. Me. Marcelo Coelho de Sá.

JOÃO PESSOA, PB

2020

S237a Santos, Jamillys.

ANÁLISE BIOMÉTRICA DE GERADORES DE CUSTO ASSISTENCIAL:
UM ESTUDO DE CASO / Jamillys Santos. – João Pessoa, 2020.
57f.: il.

Orientador(a): Prof^o Msc. Luiz Carlos Santos Junior.
Trabalho de Conclusão de Curso (Ciências Atuariais) – UFPB/CCSA.

1. Saúde Suplementar. 2. Geração de primeiro custo assistencial. 3.
Análise de sobrevivência. 4. Funções de sobrevivência e de risco. 5. Riscos
competitivos. I. Título.

UFPB/CCSA/BS

CDU:347.764(043.2)

Gerada pelo Catalogar - Sistema de Geração Automática de Ficha Catalográfica do
CCSA/UFPB, com os dados fornecidos pelo autor(a)

JAMILLYS DOS SANTOS SILVA

**ANÁLISE BIOMÉTRICA DE GERADORES DE CUSTO ASSISTENCIAL: UM
ESTUDO DE CASO**

Trabalho de Conclusão de Curso para o curso
de Ciências Atuariais na UFPB, como
requisito parcial à obtenção do título de
bacharel em Ciências Atuariais.

BANCA EXAMINADORA

Prof. Dr. Luiz Carlos Santos Junior
Orientador
Universidade Federal da Paraíba (UFPB)

Prof. Me. Filipe Coelho de Lima Duarte
Membro avaliador
Universidade Federal da Paraíba (UFPB)

Prof. Me. Werton José Cabral Rodrigues Filho
Membro avaliador
Instituto de Previdência do Município de João Pessoa (IPMJP)

AGRADECIMENTOS

Agradeço primeiramente a Deus, pois sem ele nada seria possível.

Agradeço ao meu amigo e orientador Luiz Carlos, por toda paciência e disposição a ajudar.

Agradeço ao meu noivo e aos familiares que sempre me apoiam, me dão força e nunca me deixam desistir, não conseguiria sem vocês.

Agradeço a minha melhor amiga Liliane e Ionnara.

Agradeço ao meu Coorientador e a todos os professores que me ensinaram.

RESUMO

Objetivou-se analisar o comportamento biométrico de um grupo de segurados referente à geração de custos assistenciais, a partir de uma amostra de 160.572 segurados (referentes ao período de 2013 a 2015) de uma operadora de saúde que optou pelo anonimato, mediante o uso de modelos de regressão, sob a perspectiva da análise de sobrevivência, na ausência e presença de riscos competitivos. Assim, na ausência de riscos competitivos, foram estimadas funções de sobrevivência de modo não paramétrico (KAPLAN-MEIER, 1958), além de funções de risco e de sobrevivência de modo semiparamétrico (COX, 1972); na presença, funções de risco e de sobrevivência de modo semiparamétrico (ALLIGNOL; BEYERSMANN; SCHUMACHER, 2008). Assim o tempo até a ocorrência do primeiro uso, por parte do segurado, do plano de saúde, bem como o risco de se gerar esse tipo de custo, é investigado considerando-se o efeito de algumas covariáveis (sexo, idade, custo e procedimento). Dentre os principais resultados destaca-se que na ausência de risco competitivo, o homem apresenta menos risco que a mulher de gerar o primeiro custo assistencial; o risco de gerar o referido custo tende a diminuir com o aumento da idade; em relação à procedência, todos os níveis apresentam maior risco de falha em relação às diárias e taxas; quanto aos custos, o risco de o primeiro custo assistencial ser “baixo” é maior quando comparado à categoria “alto”. Na presença de riscos competitivos, considerando cinco diferentes desfechos (procedimentos geradores do primeiro custo assistencial) o homem, em relação à mulher, apresenta maiores riscos de gerar custos assistenciais 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 (Serviços de Apoio ao Diagnóstico e Terapias); em relação ao custo, o risco de gerar o custo assistencial pela categoria moderado é maior quando estes são oriundos de materiais e medicamentos e honorários; já a categoria baixo custo é maior quando estes são oriundos de honorários.

Palavras-chave: Saúde Suplementar. Geração de primeiro custo assistencial. Análise de sobrevivência. Funções de sobrevivência e de risco. Riscos competitivos.

ABSTRACT

The objective was to analyze the biometric behavior of a group of policyholders regarding the generation of assistance costs, based on a sample of 160,572 policyholders (referring to the period from 2013 to 2015) of a health operator that opted for anonymity, using regression models, from the perspective of survival analysis, in the absence and presence of competitive risks. Thus, in the absence of competitive risks, non-parametric survival functions were estimated (KAPLAN-MEIER, 1958), in addition to semi-parametric risk and survival functions (COX, 1972); in the presence, risk and survival functions in a semi-parametric way (ALLIGNOL; BEYERSMANN; SCHUMACHER, 2008). Thus, the time until the insured's first use of the health plan, as well as the risk of generating this type of cost, is investigated considering the effect of some covariables (sex, age, cost and procedure). Among the main results, it is highlighted that in the absence of competitive risk, men are less at risk than women of generating the first care cost; the risk of generating this cost tends to decrease with increasing age; in relation to origin, all levels have a higher risk of failure in relation to daily rates and rates; as for costs, the risk of generating assistance costs by the "low" category is greater when compared to the "high" category, and by the "moderate" category it is equal to the risk of the "high" category. In the presence of competitive risk, men, in relation to women, present greater risks of generating assistance costs when these come from daily rates and fees and others; the increase in age tends to increase the risks of generating assistance costs when they come from materials and medicines and SADT, in relation to the cost, the risk of generating the assistance cost by the moderate category is higher when these come from materials and medicines and fees; the low category is higher when they come from fees and others.

Keywords: Supplementary Health. First assistance cost generation. Survival analysis. Survival and risk functions.

LISTA DE ILUSTRAÇÕES

Figura 1: Beneficiários de planos privados de saúde por cobertura assistencial do plano	16
Figura 3: Sobrevivência estimada por Kaplan-Meier para o tempo até o segurado gerar o primeiro custo assistencial.....	29
Figura 4: Sobrevivência estimada por Kaplan-Meier, por níveis da covariável sexo, para o tempo até o segurado gerar o custo assistencial	30
Figura 5: Sobrevivência estimada por Kaplan-Meier, por níveis da covariável sexo, para o tempo que leva o segurado a gerar o custo	30
Figura 6: Sobrevivência estimada por Kaplan-Meier, por níveis da covariável custo, para o tempo que leva o segurado a gerar o custo por nível.....	31
Figura 7: Curvas de sobrevivência comparadas ao modelo de Cox	34

LISTA DE TABELAS

Tabela 1: Receitas e despesas, por modalidade das operadoras, por tipo, segundo a modalidade da operadora.....	16
Tabela 2: Estatísticas descritivas referentes ao tempo que leva para o indivíduo gerar o primeiro custo assistencial e a idade do indivíduo na respectiva ocorrência entre 2013 e 2015	28
Tabela 3: Frequência da geração de primeiros custos assistenciais por tipo de Procedimento entre 2013 e 2015	28
Tabela 4: Resultados dos testes logrank utilizados para comparações das sobrevivências estimadas em cada nível (grupo), por covariáveis (sexo, procedimento e custo), considerados o tempo que leva para o segurado gerar o primeiro custo assistencial.....	31
Tabela 5: Resultados dos testes logrank utilizados para as comparações dos níveis do grupo causa considerados no tempo que leva para o segurado gerar o custo assistencial.....	32
Tabela 6: Resultados dos testes logrank utilizados para as comparações dos níveis do grupo custo considerados no tempo que leva para o segurado gerar o custo assistencial por nível...	32
Tabela 7: AIC dos setes modelos ajustados para explicar o risco de geração do primeiro custo assistencial por segurado do plano	33
Tabela 8: Desvio Quadrático Médio de cinco modelos ajustados para estimar a função de sobrevivência, ou seja, a probabilidade de não gerar o custo assistencial pela primeira vez...	34
Tabela 9: Resultados do modelo de Cox 8 ajustado para o tempo que leva o segurado a gerar o custo assistencial	35
Tabela 10: Teste de Schoenfeld para proporcionalidade de riscos do Modelo de Cox 8 ajustado para o tempo que leva o segurado a gerar o custo assistencial	36
Tabela 11: Resultados do modelo de Cox 1 ajustado para o tempo que leva o segurado a gerar o custo assistencial	36
Tabela 12: Teste de Schoenfeld para proporcionalidade de riscos do Modelo de Cox 1 ajustado para o tempo que leva o segurado a gerar o custo assistencial	37
Tabela 13: Resultados do efeito do sexo, idade e custo sobre as variáveis, para o risco do segurado a gerar o custo assistencial	38
Tabela 14: Funções de sobrevivência estimadas pelo estimador de Kaplan-Meier e ajustadas por cinco modelos de Cox	47

LISTA DE ABREVIATURAS E SIGLAS

ANS - Agência Nacional de Saúde
AIC- Akaike Information Criterion
CONSU- Conselho de Saúde Suplementar
EUA- Estados Unidos da América
EKM- Estimador de Kaplan-Meier
FIA-Função de Incidência Acumulada
HCPA- Hospital de Clínicas de Porto Alegre
IESS - Instituto de Estudos em Saúde Suplementar
IPCA - Índice de Preço do Consumidor Amplo
PIB- Produto Interno bruto
SIP- Sistema de Informações de Produtos
SUS- Sistema Único de Saúde
VCMH - Variação de Custo Médico-Hospitalar

SUMÁRIO

1 INTRODUÇÃO	11
1.1 Contexto e problema	11
1.2 Objetivos	12
1.2.1 Objetivo geral	12
1.2.2 Objetivos específicos.....	12
1.3 Justificativa	13
2 REFERENCIAL TEÓRICO	14
2.1 Saúde Suplementar	14
2.2 Revisão de Literatura	17
2.2.1 Custos Médico-Hospitalares	17
2.2.2 Análise de Sobrevivência em Custos	20
3 ASPECTOS METODOLÓGICOS	22
3.1 Tipos de pesquisa, universo, coleta de dados e variáveis	22
3.2 Análise de sobrevivência	23
4 ANÁLISE E DISCUSSÃO DE RESULTADOS	28
4.1 Análise descritiva	28
4.2 Análise de sobrevivência na ausência de riscos competitivos	29
4.3 Análise de sobrevivência na presença de riscos competitivos	37
5 CONCLUSÃO	40
REFERÊNCIAS	42
APÊNDICES	46
APÊNDICE A – FUNÇÕES DE SOBREVIVÊNCIAS ESTIMADAS	47
APÊNDICE B - SCRIPT DA ANÁLISE EM LINGUAGEM R	48

1 INTRODUÇÃO

A saúde suplementar é um serviço oferecido pelas operadoras de plano de saúde para que a população tenha acesso a atendimentos clínicos e hospitalares alternativos aos prestados pelo serviço público.

A fim de contribuir com a discussão que envolve alguns dos custos desse produto, o presente trabalho apresenta nesta introdução um breve contexto, além do problema de pesquisa, objetivos e justificativa.

1.1 Contexto e problema

A saúde privada se caracteriza por ofertar planos e seguros de saúde que oferecem assistência médica aos beneficiários (CONASS, 2011). Dentre seus subtipos, o ramo de saúde suplementar é, no Brasil, o mais predominante. Esse tipo de seguro age protegendo as pessoas de forma a garantir que quando ocorra um sinistro, os segurados não venham a desembolsar um valor monetário que as desestabilizem financeiramente (FEDERAÇÃO NACIONAL DE SAÚDE SUPLEMENTAR (FENASAÚDE), sd).

De acordo com o Art. 1º, § 1º, da Lei nº 9.656, de 3 de junho de 1998, o órgão fiscalizador dos planos assistenciais de saúde suplementar é a Agência Nacional de Saúde (ANS), que se responsabiliza pela regulamentação de contratos, produtos e serviços prestados, garantindo que a cobertura assistencial seja feita de forma adequada (BRASIL, 1998).

A saúde suplementar no Brasil é um setor de grande impacto econômico, responsável por assegurar os beneficiários dos planos de saúde e pela geração de empregos e arrecadação de impostos. Em 2015, o setor de saúde suplementar pagou R\$ 36,4 bilhões em tributos (VISÃO SAÚDE, 2016) e gerou aproximadamente 3,3 milhões de empregos (IESS, 2017).

Dada a sua importância social e econômica, muitas pesquisas foram e vem sendo realizadas dentro desta temática, principalmente com foco em questões legais e/ou composição de custos assistenciais.

Dentre estes, destacam-se as análises de tais custos, que explicam seus aumentos, geralmente, em função de fatores (SILVA; CARNEIRO, 2018; IESS, 2019): demográficos (como o envelhecimento); macroeconômicos (como a inflação); tecnológicos (como o avanço de técnicas medicinais); do uso inadequado dos serviços prestados pelas operadoras; e da contração econômica referente ao PIB.

Ainda em relação aos custos assistenciais, o Instituto de Estudo em Saúde

Suplementar (IESS) investiga anualmente a Variação de Custo Médico-Hospitalar (VCMH), além dos fatores que impactam na utilização desses recursos. Segundo esses estudos, a VCMH em 2018 foi de 17,3%, chegando a superar a variação do Índice de Preço do Consumidor Amplo (IPCA), que foi de 3,7%, avaliados no mesmo período.

A precificação (e atualização) desse serviço – comumente baseada na Teoria do Risco – deve considerar tais custos observando-se as restrições impostas pela ANS. Assim, é frequente o uso de alguns métodos de tarifação: sinistralidade, utilizada para atualizar a tarifa, é a razão entre os custos assistenciais (o sinistro total médio, em valores monetários) e as receitas assistenciais (o prêmio coletivo, em valores monetários); prêmio puro, por sua vez, carrega o prêmio de risco (o sinistro total médio, em valores monetários).

Complementarmente, levanta-se, aqui, a hipótese de que custos assistenciais são afetados por, além da frequência e da severidade (RODRIGUES, 2008), outro importante e pouco abordado fator: o biométrico, ou seja, fatores que impactam no modelo do ponto de vista da extensão e saúde da vida humana em várias dimensões, procurando refletir o comportamento biométrico de determinado grupo. Tendo isso em vista, o presente trabalho apresenta a seguinte problemática: qual o efeito que o fator biométrico exerce sobre os custos assistenciais de uma operadora de saúde?

1.2 Objetivos

1.2.1 Objetivo geral

Analisar o comportamento biométrico de um grupo de segurados referente à geração de custos assistenciais, a partir de uma amostra de 160.572 segurados (referentes ao período de 2013 a 2015) de uma operadora de saúde que optou pelo anonimato, mediante o uso de modelos de regressão sob a perspectiva de riscos competitivos em análise de sobrevivência.

1.2.2 Objetivos específicos

- Estimar, para o grupo de segurados analisado, as probabilidades relativas à geração de custo assistencial, por meio do estimador de Kaplan-Meier.
- Estimar o risco, para cada causa e considerando-se o efeito de covariáveis, de gerar custo assistencial, por meio da Regressão de Cox.
- Construir tábua biométrica sob a ótica dos riscos competitivos intitulada “geração

de custo assistencial”

1.3 Justificativa

De acordo com Goovaerts et al. (1986), definir o prêmio a ser cobrado pela seguradora é um grande desafio, pois se a mesma cobra um valor abaixo do necessário, logo se tornará insolvente; por outro lado, quando cobrado um valor excessivo, outros problemas surgem, como a perda de competitividade e de beneficiários. Assim, a base que dá origem ao prêmio deve ser estruturada de forma equilibrada, levando em consideração todos os riscos e fatores que impactam no custo assistencial.

A revisão mostra que grande parte dos trabalhos que relatam a variação dos custos assistenciais em planos de saúde aponta como principais fatores, o envelhecimento, a taxa de inflação, avanço da tecnologia e até mesmo o uso inadequado dos serviços prestados. (LEAL E MATOS, 2009), (SÁ, 2012), (ALDRIDGE E KELLEY, 2015) e (REIS, 2018). No entanto, nenhum deles relaciona o referido custo com o risco de um segurado sinistrar; sendo assim, este trabalho propõe uma forma complementar de precificação, agregando a função biométrica.

Assim o presente trabalho, simplificadamente, pretende contribuir com a discussão por meio da inclusão da componente biométrica na composição dos custos assistenciais. Desta forma, a probabilidade de geração do custo assistencial, bem como sua função de risco, será estimada e analisada por meio de modelos semiparamétricos e não paramétricos de sobrevivência, na presença e ausência de riscos competitivos; Que são de acordo com Carvalho et al. (2011), múltiplos decrementos são situações no qual o indivíduo está exposto ao risco de ocorrer k eventos, ou seja, k forças atuam sobre os indivíduos para tirá-los de seu estado inicial; nesse caso, trataremos como decrementos as diferentes formas de utilização do plano pelo usuário, ou seja, as diferentes formas de geração de custo assistencial (Materiais e medicamentos, Honorário, Diárias e taxas, Outros, SADT).

2 REFERENCIAL TEÓRICO

2.1 Saúde Suplementar

O sistema de saúde no Brasil é um direito oferecido a todos os cidadãos. Esse, por sua vez, se divide em duas modalidades: o público, onde o governo atua dando assistência e proteção por meio do Sistema Único de Saúde (SUS); e o privado, que é composto pelo ramo de saúde suplementar e pelos serviços autônomos (CONASS, 2015).

Nos anos 60, o acesso à saúde era algo bastante limitado, mas com o passar dos anos, com o crescimento populacional e com a incapacidade de oferta do serviço público com qualidade, a necessidade de assistência à saúde aumentou, abrindo portas para o mercado de saúde suplementar, e consequentemente impulsando o mercado tecnológico (FENASAÚDE, sd).

A criação do SUS, prevista na constituição de 1988, faz referência à assistência de saúde pública para a população, de forma a garantir que esta tenha seus direitos referidos aos serviços de saúde pública reservados, sejam suas necessidades de caráter permanente ou provisório, isoladamente ou em conjunto, prestados por pessoas físicas ou jurídicas de direito público e privado. A partir dessa linha de pensamento foram designados alguns artigos que fazem referência à assistência de saúde privada, que dispõe das características de atuação, bem como as diretrizes de funcionamento; no entanto, não trata da regulação das operadoras de prestação de serviço privado de saúde (BRASIL, 1990).

Paralelamente a criação do SUS, as empresas passaram a oferecer aos seus funcionários assistência médico-hospitalares como forma de mediação entre médicos e pacientes. Assim, surgiram os planos de saúde oferecidos por empresas especializadas nesse ramo, e o mercado de saúde suplementar se expandiu, gerando a necessidade de regulação (BAHIA; VIANA, 2002).

As regularizações estabelecem regras a fim de nortear e assegurar que os acordos feitos entre as partes sejam cumpridos. Assim, a regulação da saúde suplementar no Brasil é feita através da intervenção do governo, com o intuito de nortear os serviços prestados pelas operadoras (CONASS, 2007).

A elaboração da Lei nº 9.656, de 3 de junho de 1998 veio organizar e garantir de forma legal que os serviços oferecidos fossem efetuados de acordo com a cláusulas pré-estabelecidas (BRASIL,1998). No entanto alguns pilares estruturais deixaram a desejar, causando aumento de preços em planos com tabelas que exploravam o consumidor, fazendo-o

optar por planos mais baratos. Em decorrência disso, médicos foram mal remunerados, tornando-os avessos a trabalhar sob essas circunstâncias; assim, no quinto ponto de debate sobre regulações, surgiu a ANS, uma agência reguladora de planos de saúde privados (BAHIA; VIANA, 2002).

A ANS é uma autarquia vinculada ao ministério da saúde, responsável por regular, normatizar, controlar e fiscalizar as atividades das operadoras de saúde suplementar, garantindo que os segurados tenham seus direitos reservados (BRASIL, 2000).

Segundo Montone (2000), dentre as características das agências, três são essenciais:

- O poder de regular de acordo com a legislação;
- A administração flexível;
- A autonomia financeira.

De acordo com Art. 4º, da Lei nº 9.961, de 28 de janeiro de 2000, compete a ANS monitorar os preços cobrados pelas operadoras, bem como os produtos e seus produtores, estabelecendo critérios e obrigações, a fim de normatizar esse departamento.

A regulação desse setor, inicialmente prestado pelo Conselho de Saúde Suplementar (CONSU) e posteriormente passado para a ANS, têm por objetivo regular o setor para ambas as partes envolvidas, abarcando quadro medidas: as condições de acesso, a cobertura assistencial, os aspectos econômico-financeiros e a fiscalização (FENASAÚDE, 2014).

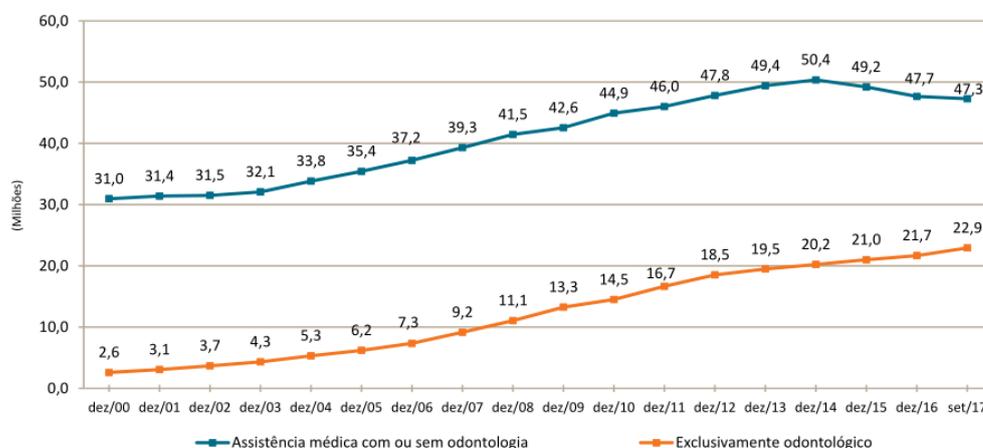
A ANS não é responsável pela tarifação cobrada pelos planos de saúde, no entanto, ela impõe limites e diretrizes a serem seguidas pelas operadoras, tendo em vista que como um órgão regulador, ela não pode deixar esse mercado se tornar inviável para os consumidores. Sendo assim, o art. 15 da Lei nº 9.656/98 permite as operadoras estabelecerem preços que variam de acordo com a idade, respeitando os aspectos regulatórios da ANS (FENASAÚDE, 2014).

Com esses avanços da legislação que regularam a relação operadora e consumidor, percebe-se que número de beneficiários vem aumentando a cada ano, fazendo com que o mercado de saúde suplementar se torne cada vez mais forte na economia. Uma análise feita entre o período de dezembro de 2012 a dezembro de 2017 mostrou que o número de beneficiários de planos médicos teve um aumento de 18,2% (FENASAÚDE, 2018).

Nota-se também que mercado de saúde suplementar no Brasil domina 42,6% do mercado segurador, chegando a arrecadar cerca de R\$183,4 bilhões, ou seja, 2,8% do PIB brasileiro no ano de 2017(FENASAÚDE, 2018).

A Figura 1 mostra a quantidade de beneficiários em planos de saúde por cobertura assistencial.

Figura 1: Beneficiários de planos privados de saúde por cobertura assistencial do plano (Brasil - dezembro/2000-setembro/2017).



Fontes: SIB/ANS/MS – 09/2017.

Caderno de Informações da Saúde Suplementar – dezembro/2017

Notas: 1) O termo “beneficiário” refere-se a vínculos aos planos de saúde, podendo incluir vários vínculos para um mesmo indivíduo; 2) Dados preliminares, sujeitos a revisão.

As operadoras constituem um papel fundamental nesse mercado, pois viabilizam o acesso das pessoas aos serviços de saúde privada, deixando-as mais confortáveis e seguras. Por outro lado, o avanço medicinal estimula o crescimento econômico através de novas empresas com tecnologia de ponta e geração de empregos (FENASAÚDE, 2018).

Tabela 1: Receitas e despesas, por modalidade das operadoras, por tipo, segundo a modalidade da operadora em milhões.

Modalidade da Operadora	Receita de contraprestações	Outras receitas operacionais	Despesa assistencial	Despesa administrativa	Despesa de comercialização	Outras despesas operacionais
Total	132.002,352414	12.335,33899	111.462,927243	13.981,463.63	3.995,862353	12.527,076548
Operadoras médico-hospitalares	129.826,936818	12.256,19890	110.446,620226	13.464,53299	3.775,058368	12.602,579220
Autogestão	15.632,514271	846,227644	14.790,182600	1.703,294547	1.613,393	704,889513
Cooperativa Médica	42.428,951926	7.985,790750	35.428,488854	4.702,187692	673,872682	7.774,395401
Filantropia	1.810,062069	2.461,725105	1.444,265634	914,832155	22,717203	1.838,806914
Medicina de Grupo	40.675,6606	956,726336	32.835,043842	4.593,524107	1.441,526790	1.602,738262
Seguradora Especializada em Saúde	29.279,747919	5.729,065	25.948,639296	1.550,694498	1.635,328300	681,749130

Operadoras exclusivam ente odontológic as	2.175,415596	79,140097	1.016,307016	516,930631	220,803986	-75,502672
Cooperativa odontológic a	424,036769	53,544942	274,637284	127,771716	18,372781	45,170180
Odontologi a de grupo	1.751,378827	25,595155	714,669732	389,158915	202,431205	-120,672852

Fontes: Elaboração própria através de dados do Caderno de Informações da Saúde Suplementar – dezembro/2017

DIOPS/ANS/MS – 21/11/2017.

Notas: 1. Dados preliminares, sujeitos à revisão;

2. Dados referentes ao primeiro trimestre.

3. Não inclui dados da operadora exclusivamente odontológica com até 20.000 beneficiários, dispensadas do envio do DIOPS nos três primeiros trimestres.

A tabela 1 mostra os valores monetários de receitas e despesas geradas pelas operadoras de saúde (ANS, 2017). Dentre as suas modalidades, a que mais se destaca em termos de receita e despesa é a médico-hospitalar, chegando a ser responsável, no ano de 2017, por 78,5% do total das despesas de operadoras de saúde.

2.2 Revisão de Literatura

2.2.1 Custos Médico-Hospitalares

O custo médico hospitalar é constituído por despesas geradas pela assistência dada aos beneficiários quando estes fazem o uso do plano através da realização de exames, consultas e outros procedimentos cobertos; calculado pelo IESS geralmente em períodos anuais, trata-se da razão entre o total de despesas assistenciais (em volume monetário) e o número de pessoas expostas ao plano (LARA; REIS; MINAMI, 2018).

De acordo com Sá (2012), as inovações tecnológicas auxiliam de forma inegável na melhoria da saúde, a partir dos serviços prestados pelas operadoras. No entanto, grandes avanços tecnológicos exigem custos elevados, o que contribui para que os gastos com saúde no Brasil representando em torno de 57,6% do total de gastos com saúde no Brasil, somando um valor de R\$314,6 bilhões podendo ter impacto nos custos que são passados para seus clientes, identificados entre 2010 e 2015 (LARA, 2018).

Entende-se que diante de custos (e preços) tão elevados, a demanda pelos serviços de saúde privada, *ceteris paribus*, tende a diminuir. Outro fator redutor dessa demanda é a fase recessiva do ciclo econômico e a consequente redução do nível de emprego e renda (IESS, 2019).

No entanto, não é apenas isso que causa a elevação do custo: um fator importante e de impacto significativo é a faixa etária. Por exemplo, no período de dezembro de 2017 a dezembro de 2018, a variação em pontos percentuais para pessoas com idade igual ou superior a 59 anos foi igual a 2,1, ou seja, mais pessoas nessa faixa etária passaram a utilizar dos serviços sem que houvesse um aumento nos prêmios cobrados pelas operadoras (IESS, 2019).

Tendo em vista que muitos fatores podem influenciar no custo, como idade, sexo, etc., é interessante notar que perfis podem ser traçados, a fim de entender quem gera mais custo. Dessa forma, Aldridge e Kelley (2015), a partir de dados próprios do MEPS (conjunto de pesquisas em larga escala de famílias e indivíduos) nos EUA, no ano de 2011, com base na *National Health Expenditure Accounts* e *National Center for Health Statistics*, para estimar as despesas geradas pela população estudada, identificaram que pessoas com um perfil de alto custo assistencial podem ser encaixadas em três grupos: no primeiro, estão as pessoas que passaram por uma situação que gerou um alto custo, mas que foi normalizada e por isso voltaram a gerar custos mais baixos (como, por exemplo, um indivíduo que sofreu ataque isquêmico vascular); no segundo grupo estão as pessoas que possuem doenças crônicas e devido a isso geram custos anuais elevados; e o terceiro, onde estão as pessoas que geram custos elevados por estarem no fim de suas vidas.

De forma semelhante, o IESS (sd) objetivou por meio de uma análise, caracterizar os beneficiários brasileiros de maior custo assistencial, a partir de dados de 76,1 mil beneficiários no ano de 2015. Dentre seus principais resultados, tem-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 geram os custos mais altos. Destes, 56,8% são mulheres e 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 com dados do Sistema de Informações de Produtos (SIP), através de um índice de valor que compara o produto 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 se compara 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) e não pelo aumento da frequência (associada a fatores sociais, demográficos e epidemiológicos).

Além disso, uma forma usual criada pelo IESS para estudar os custos assistenciais foi a Variação de Custos Médico-Hospitalares (VCMH), que capta a variação de preços e a frequência de utilização, pelos beneficiários, dos serviços cobertos (LARA; REIS; MINAMI, 2018).

Um comparativo de relatórios das consultorias Aon Hewitt, Mercer e Willis Towers Watson, contemplando mais de 230 operadoras de saúde em mais de 91 países, realizado por Silva e Carneiro (2018), mostrou que além dos fatores já citados nesta revisão, o uso inadequado dos planos de saúde pode impactar na VCMH, dado que a utilização desnecessária de serviços como procedimentos médicos e exames, sem a efetiva necessidade, gera o custo assistencial.

Comumente se comparam as VCMH às inflações tradicionais como o IPCA. Em 2016 a variação da inflação (IPCA/IBGE) foi igual a 6,3%, enquanto a VCMH foi de 20,4% (LARA; REIS; MINAMI, 2018). Segundo Reis e Lara (2019), essa discrepância ocorre porque a VCMH é a soma dos preços com a variação da frequência mais o efeito das variações dos mesmos. Sendo assim, só teria como a inflação ser igual à VCMH se não houvesse variação na frequência – desta forma, o IPCA se torna um comparativo inviável para a variação de custos assistenciais.

A variação nos custos também depende de diversos fatores, que nem sempre estão ligados aos indicadores de preços; na série histórica apresentada pelo IESS de dezembro de 2007 a dezembro de 2018, a VCMH teve aumentos significativos, em dezembro de 2011, os gastos em percentuais foram de 9,2% em consultas, 6,0% em exames, 10,7% em terapias, 11,8% em OSA, 13,5% em internação e 85,8% em outros, gerando uma VCMH de 12,9%, note que 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, gerando uma VCMH de 17,3%, mostrando que a diversas formas de se gerar o custo assistencial vem aumentando cada vez mais ao longo dos anos. (IESS, 2019).

Para o cálculo da VCMH são utilizadas informações de grande parte das operadoras,

fazendo com que a VCMH/IESS seja o índice mais preciso de variação. Para a sua composição, leva-se em consideração a ocorrência de sinistro, a data do ocorrido e a quantidade de expostos ao risco no período de um ano; para cálculos aplicados em cada operadora, os custos provem das despesas geradas pela utilização dos serviços prestados (WATSON, 2018).

Exposto o foco de análise dos custos médico-hospitalares encontrados na literatura, reitera-se a contribuição da presente pesquisa, que objetiva adicionar a informação biométrica na análise dos referidos custos.

2.2.2 Análise de Sobrevivência em Custos

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 irão influenciar 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 “análise de sobrevivência em custos”, “sobrevivência em custos”, “sobrevivida em custos” e “costs survival analysis”.

Jackman et al. (2017), 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, a partir de dados referentes a 160 pacientes *prepathway* (antes da intervenção) e 210 *postpathway* (depois da intervenção), 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 algumas covariáveis (sexo, raça, perfil genérico, distância entre residência e clínica, e a 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 que com a regressão ajustada, o custo foi de U\$\$ 69.122 antes do *prepathway* e U\$\$ 52.037 *postpathway*, obtendo uma economia de U\$\$ 17,085 mil.

Herrmann (2011) objetivou analisar o custo-efetividade em um grupo de 318 pacientes, em que 219 eram do sexo masculino e 153 já haviam sido internado pelo menos uma vez. O banco de dados foi do ambulatório de insuficiência cardíaca do Hospital de Clínicas de Porto Alegre (HCPA), onde o evento de interesse era a morte do paciente por insuficiência cardíaca. Inicialmente foi feito o uso do método de Kaplan-Meier para estimar a função de sobrevivência para todo o grupo. Em seguida, dividiram-se os pacientes em dois grupos: o grupo 1 foi constituído pelos que já foram internados pelo menos uma vez; e o grupo 2 pelos que nunca foram internados. Assim, concluiu-se que o modelo mais adequado para ajustar o comportamento de morte do grupo 1 é o log-normal, e que a probabilidade de sobreviver a 4000 dias é de 0,254; para o grupo 2, o modelo de regressão de Weibull foi considerado mais adequado, resultando em uma probabilidade de sobreviver 4000 dias igual a 0,495.

Como um exemplo do uso de riscos competitivos em atuária, Santos Júnior (2018), em sua tese, estudou os critérios de elegibilidade e a escolha das premissas atuariais para analisar, mediante estudo de caso, a aplicabilidade de modelos de sobrevivência pouco explorados para estimação de variáveis relevantes a essas premissas. Para isso, ele fez o uso da análise de sobrevivência, através de modelos paramétricos, não paramétricos e semiparamétricos, com abordagens de riscos competitivos, chamando atenção para o modelo de Cox, dada o bom ajuste e a facilidade de interpretação.

Como se vê, os materiais acima não discutiram efetivamente o conteúdo buscado, mesmo que seus títulos o sugerissem, Mesmo assim optou-se por apresentar, mesmo que timidamente, os materiais encontrados para destacar a escassez e a necessidade de novos estudos.

3 ASPECTOS METODOLÓGICOS

No presente estudo, o evento de interesse é a geração de custo assistencial por parte do segurado, ou seja, o indivíduo falha quando gera custo, enquanto a censura é determinada quando o período delimitado de estudo se encerra e o segurado não gera custo ao plano até aquela data. Como se considera a fonte geradora de custo, isto é, o “procedimento” realizado, caracteriza-se o uso de modelos de riscos competitivos em análise de sobrevivência. Além disso, considera-se a presença de covariáveis na modelagem do tempo até a geração de custo assistencial, a saber.

3.1 Tipos de pesquisa, universo, coleta de dados e variáveis

Esse estudo de caso, de caráter quantitativo, faz o uso de dados longitudinais, é descritivo quanto aos objetivos e bibliográfico/documental quanto aos procedimentos.

Para esse estudo foram utilizados dados disponibilizados por uma operadora de saúde (que optou pelo anonimato), contendo informações de 160.572 beneficiários no período de 2013 a 2015 com periodicidade diária.

Como foram disponibilizados apenas os dados referentes aos segurados que utilizaram o plano durante o período observado, não existem tempos de censura, só existem tempos de falha, isto é, trata-se de uma análise de sobrevivência sem censura.

A variável resposta em análise de sobrevivência é o tempo de falha. Como a falha (ou evento de interesse) do corrente trabalho é a geração de custo assistencial, logo, a variável resposta é o tempo, desde a entrada do segurado no plano, até a geração de custo assistencial.

Listam-se no Quadro 1 as variáveis disponibilizadas. Quanto aos níveis da variável Procedimento, “materiais e medicamentos” englobam todos os materiais e medicamentos utilizados nos procedimentos por pessoa na sua primeira utilização do plano; “honorários” são relativos aos honorários médicos, honorários não médicos, procedimentos terapêuticos, hemoterapia e gasoterapia; “Diárias e taxas” correspondem às diárias, taxas e internações em casa; “SADT” é formado por serviços de apoio ao diagnóstico e terapias.

Quadro 1: Variáveis 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

Variáveis	Tipo	Níveis
		R\$108,56 Moderado – valores entre R\$108,56 e R\$453,53 Alto – valores maiores ou iguais a R\$453,72
Idade	Quantitativa	-
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.

A subtração entre a data de uso e a data de adesão ao plano define a variável resposta “tempo até a primeira utilização do plano pelo segurado”; a subtração entre a data de uso e a data de nascimento representa a “idade do segurado na ocorrência da falha”, ou seja, no seu primeiro uso do plano.

“Procedimento”, aqui, indica as causas de geração de custo assistencial; é usada num primeiro momento (na ausência de riscos competitivos) como covariável e num segundo momento (na presença de riscos competitivos) como múltiplos decrementos, isto é, forças tendentes a retirar o participante de seus status inicial “não gerador de custos assistenciais”.

Na ausência de riscos competitivos, isto é, quando a causa de falha é única (abandonar-se o status inicial de “não gerador de custo assistencial” apenas pela causa geral “gerou custo assistencial ao plano”), “sexo”, “idade”, “custo” e “procedimento” são covariáveis no modelo de regressão que estima o risco de geração de custo assistencial; na presença de riscos competitivos, ou seja, quando mais de uma causa concorre para provocar a falha (é possível abandonar o status inicial de “não gerador de custo assistencial” por mais de uma causa, não concomitantes, “gerou custo assistencial ao plano por meio do procedimento Materiais e medicamentos”, “gerou custo assistencial ao plano por meio do procedimento Honorário”, etc.) apenas “sexo”, “idade” e “custo” são covariáveis no modelo de regressão que estima o risco de geração de custo assistencial, já que procedimento redefine as causas de falha.

3.2 Análise de sobrevivência

3.2.1 Análise baseada em risco

O presente trabalho faz uso dos modelos de sobrevivência, especializados na investigação do tempo de falha. Dentre seus principais objetivos, tem-se a estimação das funções de sobrevivência e de risco.

Segundo Herrmann (2011), a função de sobrevivência é a probabilidade de sobrevivência do indivíduo até o instante t (tempo de falha), dado pela equação (1):

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

A função de risco, por sua vez, é obtida através da divisão entre a probabilidade de ocorrência da falha em um período de tempo e o tamanho do intervalo, sem que tenha ocorrido a mesma até o instante t (COLOSIMO; GIOLO, 2006), conforme expressão (2):

$$h(t) = \frac{S(t) - S(t + \Delta t)}{\Delta t S(t)} = \frac{f(t)}{S(t)}. \quad (2)$$

A função de sobrevivência pode ser estimada não paramétrica, paramétrica e semiparametricamente. Aqui, serão apresentadas, sucintamente, a primeira e a terceira abordagens, tanto na ausência quanto na presença de riscos competitivos.

Na ausência de riscos competitivos, de acordo com Colosimo e Giolo (2006), o estimador Kaplan-Meier para a função de sobrevivência considera tanto o intervalo de tempo quanto o número de falhas e é escrito conforme (1):

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

$t_1 < t_2 \dots < t_k$, 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 custo assistencial no intervalo de tempo $[t_{j-1}, t_j)$, feita através da divisão do número de falhas em t_j , pelo número de indivíduos expostos ao risco em t_{j-1} .

O modelo semiparamétrico de Cox, por sua vez, estima, inicialmente, a função de risco. Também conhecido como o modelo de riscos proporcionais, o modelo de regressão de Cox, de forma genérica é dado por (4):

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

Onde $\lambda(t)$ e λ_0 são funções de taxa de falhas e g é uma função que deve ser especificada.

Desta forma, o modelo de Cox, consoante Colosimo e Giolo (2006), é composto pelo produto de duas funções, uma paramétrica não especificada e a outra não paramétrica especificada que torna o modelo mais flexível; sabe-se que β é o vetor associado às covariáveis. Para a sua utilização, é necessária a verificação do pressuposto de proporcionalidade dos riscos. Dado isso, por meio desse modelo, é possível estimar o risco de gerar custo assistencial, considerando-se os efeitos das covariáveis.

Por fim, pretende-se construir uma tábua biométrica sob a ótica dos riscos competitivos, intitulada “geração de custo assistencial” por meio de técnicas de eventos competitivos. De acordo com Carvalho et al. (2011), na presença de múltiplos decrementos, o indivíduo está exposto ao risco de ocorrer k eventos, ou seja, k forças atuam sobre os indivíduos para tirá-los de seu estado inicial; nesse caso, trataremos como decrementos as diferentes formas de utilização do plano pelo usuário, ou seja, as diferentes formas de geração de custo assistencial (Materiais e medicamentos, Honorário, Diárias e taxas, Outros, SADT).

Existem diversas maneiras de modelar os riscos, conforme Carvalho et al. (2011): pela sobrevivência livre de eventos, onde é considerado o tempo até o primeiro evento, independente de qual evento ocorra (o tempo até que o segurado gere o custo, independentemente do procedimento); pelo risco específico por causa, onde a intenção é estimar o efeito das covariáveis sobre as falhas específicas (mede o efeito covariáveis sobre o risco de falha por tipo de procedimento); pelo risco de subdistribuição, uma noção alternativa (ao risco específico por causa) de risco. Aqui, são abordadas as duas primeiras maneiras.

O modelo livre de eventos corresponde à abordagem tradicional já apresentada, na ausência de riscos competitivos. A segunda abordagem, isto é, a de risco específico por causa, assume que cada risco específico por causa segue um modelo de Cox (1972).

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

em que β_{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 de base não especificada e não negativa.

Na ausência de risco competitivo, os modelos se reduzem a um modelo de sobrevivência de Cox padrão, o que permite facilitar a interpretação; na presença de riscos concorrentes, no entanto, a interpretação é um pouco mais complicada – “o aumento da

ocorrência de uma falha não é necessariamente resultado do aumento do respectivo risco específico por causa; ele pode ser provocado pela diminuição do risco específico da causa concorrente” (SANTOS JÚNIOR, 2018, p. 59).

A análise é realizada em ambiente R, viabilizada pelos pacotes survival (THERNEAU, 2015), mvna (ALLIGNOL; BEYERSMANN; SCHUMACHER, 2008), etm (ALLIGNOL; SCHUMACHER; BEYERSMANN, 2011) e cmprsk (GRAY, 2014).

3.2.2 Avaliação e seleção de modelos

No processo de modelagem, é necessário adotar alguns procedimentos para avaliar adequabilidade do modelo e posteriormente, dentre os adequados, selecionar aquele considerado “melhor”. Neste sentido, o presente trabalho utiliza-se de três ferramentas: a) comparação gráfica entre os ajustes (para a função de sobrevivência, isto é, a probabilidade de não gerar o primeiro custo assistencial) realizados pelo modelo de Cox e o de Kaplan-Meier, considerando-se adequado(s) o(s) modelo(s) de Cox com forma(s) e nível(is) parecido(s) com o de Kaplan-Meier, de acordo com Colosimo e Giolo (2006); b) O Critério de Informação de Akaike (AIC), cujo menor valor indica o melhor dentre os modelos comparados; c) o Desvio Quadrático Médio (DQM), cujo menor valor também indica o melhor dentre os modelos comparados; d) Teste de Schoenfeld ; e) Teste de Logrank

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 o termo $2k$, respectivamente. Dessa forma, AIC faz uma ponderação entre a adequação aos dados e a complexidade do modelo (SOBRAL; BARRETO, 2016),

O Desvio Quadrático Médio (DQM) permite avaliar a distância entre os dados observados (d_o) e esperados (d_e). De acordo com Pinheiro (2007), quanto menor DQM, mais aderente é o ajuste. É definida como $DQM_t = \left(\frac{d_o - d_e}{d_e}\right)^2$.

O Teste de Schoenfeld é usado para verificar a suposição de riscos proporcionais no modelo de Cox, tendo como hipótese nula que seus riscos são proporcionais, através de seus resíduos; Desta forma q_i é dado por:

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

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

Teste de Logrank verifica se as populações possuem propriedades de riscos proporcionais. 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).

4 ANÁLISE E DISCUSSÃO DE RESULTADOS

4.1 Análise descritiva

A Tabela 2 mostra as estatísticas descritivas referentes ao tempo, idade e custo que leva para o segurado gerar o primeiro custo assistencial. É possível observar que o menor dos tempos é 1, ou seja, algum dos usuários utilizou o plano já no primeiro dia de cobertura; a menor idade é de um ano e o menor custo foi de R\$0,02.

Tabela 2: Estatísticas descritivas referentes ao tempo que leva para o indivíduo gerar o primeiro custo assistencial e a idade do indivíduo na respectiva ocorrência entre 2013 e 2015

Variável	Mínimo	Máximo	Média	Mediana	3rd Qu
Tempo (dias)	1	31,40	17,49	16,17	24,40
Idade (anos)	1	115,5784	41,1626	37,2868	60,0602
Custo	0,02	53857,40	108,56	70,00	113,50

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

A variável custo, para fins de regressão, foi categorizada para que fosse possível obter uma melhor interpretação dos resultados. Nesse caso, ela é formada por 3.571 casos considerados como de custo alto, 119.774 como de custo baixo e 37.227 como de custo moderado.

A Tabela 3 exhibe as frequências de geração de custo assistencial por Procedimento. De um total de 160.572 primeiros usos do plano (segurados), a maior parte (97229) das frequências se refere ao pagamento de honorários.

Tabela 3: Frequência da geração de primeiros custos assistenciais por tipo de Procedimento entre 2013 e 2015

Procedimento	Quantidade
Diárias e Taxas	7000
Honorários	97229
Materiais Medicamentos	9323
Outros	13178
SADT	33842

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

O banco de dados é formado por 93.207 mulheres e 67.365 homens.

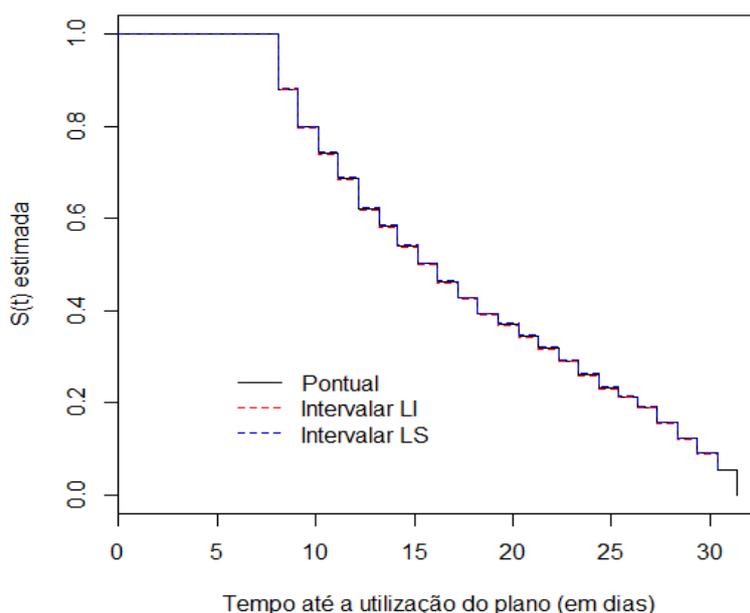
4.2 Análise de sobrevivência na ausência de riscos competitivos

Inicialmente, na análise não paramétrica, estimam-se as funções de sobrevivência relativas a todo o grupo de segurados a) sem a consideração de covariável b) considerando diferentes funções de sobrevivência por níveis de variáveis categóricas (sexo e procedimento). Em seguida, averíguam-se, por meio do teste logrank, a igualdade de curvas entre os níveis de covariáveis comparados.

Depois, semiparametricamente, estima-se a função de risco relativa à geração do primeiro custo assistencial gerado pelo segurado, considerando o efeito de três covariáveis (sexo, idade e procedimento). Além disso, verifica-se o pressuposto de proporcionalidade dos riscos, por meio do teste de Schoenfeld.

Assim, a Figura 3 apresenta as funções de sobrevivência estimadas pelo estimador não paramétrico de Kaplan-Meier para o conjunto de segurados. Nota-se que as estimativas inferior e superior são muito próximas das pontuais.

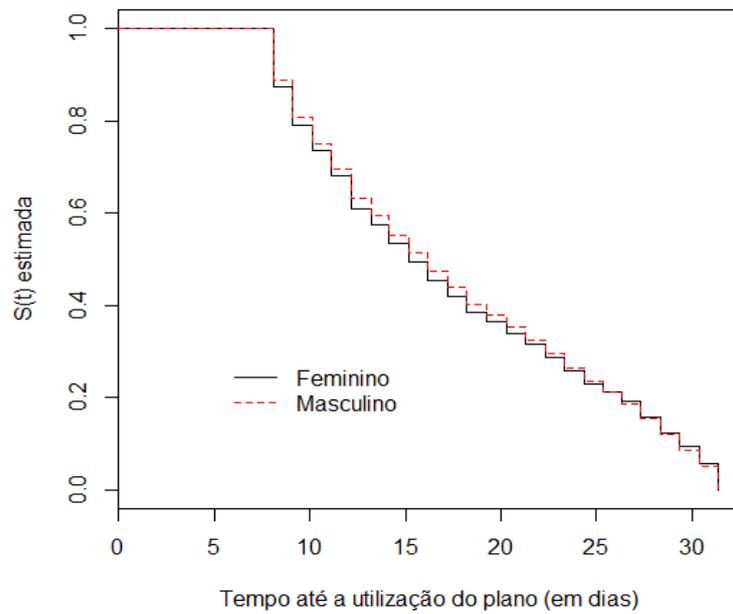
Figura 2: Sobrevivência estimada por Kaplan-Meier para o tempo até o segurado gerar o primeiro custo assistencial



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

A Figura 4 apresenta as funções de sobrevivência estimadas pelo estimador não paramétrico de Kaplan-Meier para o conjunto de segurados. Nota-se que curvas de sobrevivência estimadas para homens e mulheres parecem não diferir entre si.

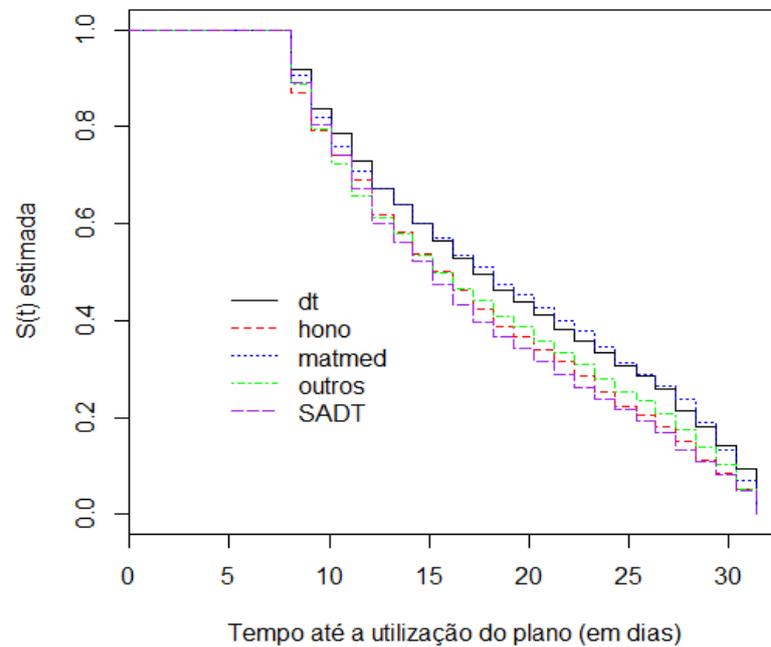
Figura 3: Sobrevivência estimada por Kaplan-Meier, por níveis da covariável sexo, para o tempo até o segurado gerar o custo assistencial



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

A Figura 5 apresenta as funções de sobrevivência estimadas pelo estimador não paramétrico de Kaplan-Meier para o conjunto de segurados. Nota-se que algumas das curvas de sobrevivência estimadas, por tipo de procedimento, parecem diferir entre si.

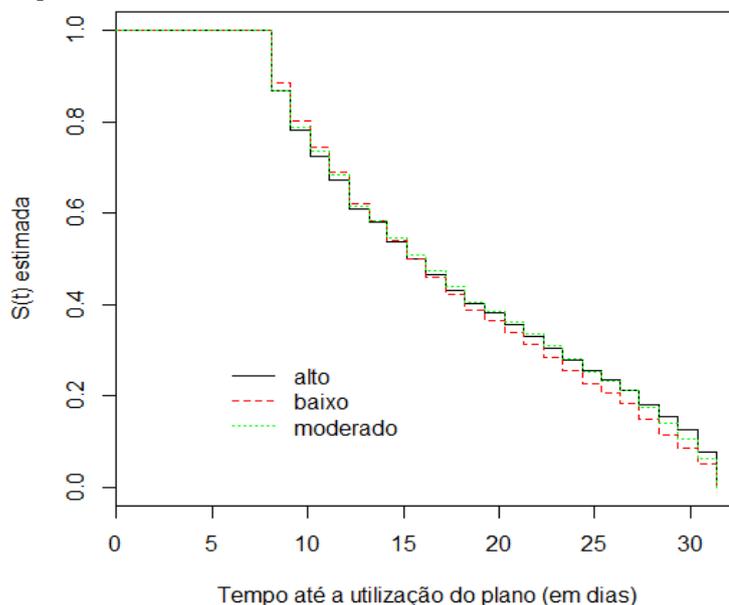
Figura 4: Sobrevivência estimada por Kaplan-Meier, por níveis da covariável sexo, para o tempo que leva o segurado a gerar o custo



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

A Figura 6 apresenta as funções de sobrevivência estimadas pelo estimador não paramétrico de Kaplan-Meier para o conjunto de segurados. Nota-se que algumas das curvas de sobrevivência estimadas, por nível de custo, parecem estar alinhadas a curva de sobrevivência.

Figura 5: Sobrevivência estimada por Kaplan-Meier, por níveis da covariável custo, para o tempo que leva o segurado a gerar o custo por nível



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

Para corroborar a interpretação dos gráficos foi realizado o teste de logrank, assumindo um nível de significância $\alpha = 0,05$. Seus resultados são apresentados na Tabela 4 e demonstram que se rejeitam as hipóteses de que: os tipos de procedimentos possuem sobrevivências estimadas iguais; homens e mulheres possuem sobrevivências estimadas iguais.

Tabela 4: Resultados dos testes logrank utilizados para comparações das sobrevivências estimadas em cada nível (grupo), por covariáveis (sexo, procedimento e custo), considerados o tempo que leva para o segurado gerar o primeiro custo assistencial

Covariável	Valor p
Sexo	0,05
Procedimento	<0,001
Custos	<0,001

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

Para melhor entender quais dos níveis da covariável procedimento se diferem, foi realizado o teste logrank para cada combinação possível (em duplas). Conforme a Tabela 5, a

única dupla que não diferiu quanto às sobrevivências estimadas foi “materiais e medicamentos” e “diárias e taxas”.

Tabela 5: Resultados dos testes logrank utilizados para as comparações dos níveis do grupo causa considerados no tempo que leva para o segurado gerar o custo assistencial

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 através de dados de uma operadora que optou pelo anonimato.

Foi realizado o teste logrank também para cada combinação possível (em duplas) de níveis de custo. Conforme a Tabela 6, a única dupla que não diferiu quanto às sobrevivências estimadas foi “Moderado” e “Alto”.

Tabela 6: Resultados dos testes logrank utilizados para as comparações dos níveis do grupo custo considerados no tempo que leva para o segurado gerar o custo assistencial por nível

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

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

Na análise semiparamétrica foi utilizado o modelo de Cox para estimar a função de risco relativa à geração do primeiro custo assistencial por parte do segurado, considerando três covariáveis (sexo, idade, procedência e custo). Realizaram-se ajustes para todas as combinações possíveis dessas covariáveis, ou seja:

$$Cox\ 1 = \textit{sexo} + \textit{idade} + \textit{procedência},$$

$$Cox\ 2 = \textit{sexo} + \textit{idade},$$

$$Cox\ 3 = \textit{idade} + \textit{procedência},$$

$$Cox\ 4 = \textit{sexo} + \textit{procedência},$$

$$Cox\ 5 = \textit{sexo},\ Cox\ 6 = \textit{procedência},$$

$$Cox\ 7 = \textit{idade},$$

$$Cox\ 8 = \textit{sexo} + \textit{idade} + \textit{procedência} + \textit{custo},$$

Cox 9 = sexo + idade + custo,

Cox 10 = sexo + procedência + custo,

Cox 11 = sexo + custo,

Cox 12 = idade + procedência + custo,

Cox 13 = idade + custo,

Cox 14 = procedência + custo,

Cox 15 = custo.

Para a seleção de modelos, optou-se pela combinação dos métodos subjetivo e objetivo, isto é, a visualização de gráfico (Figura 6) e a observação do *Akaike Information Criterium* (AIC). Desta forma foram calculados os AIC de cada ajuste, de modo que os cinco modelos com menores AIC foram os modelos Cox 1, Cox 3, Cox 8, Cox 12 e Cox 14, sendo o Cox 8 o menor deles, conforme a Tabela 7.

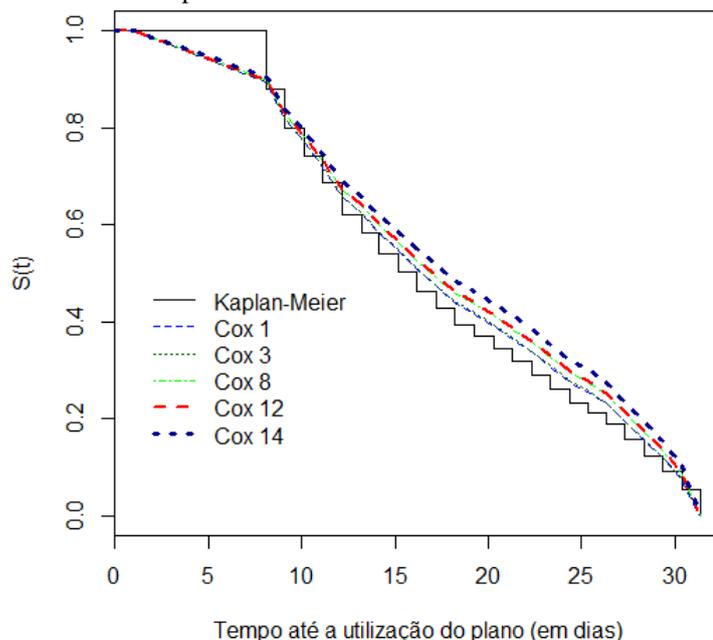
Tabela 7: AIC dos setes modelos ajustados para explicar o risco de geração do primeiro custo assistencial por segurado do plano

Modelo ajustado	AIC
Cox 1	3527458
Cox 2	3528120
Cox 3	3527460
Cox 4	3527607
Cox 5	3528260
Cox 6	3527607
Cox 7	3528125
Cox 8	3527333
Cox 9	3528033
Cox 10	3527459
Cox 11	3528152
Cox 12	3527336
Cox 13	3528039
Cox 14	3527459
Cox 15	3528154

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

Quanto à observação da Figura 7, os cinco modelos com menores AIC não destoam da sobrevivência estimada pelo estimador de Kaplan-Meier. No entanto, dentre eles, Cox 1 e Cox 3, sobrepostos neste caso, parecem estimar valores mais próximos daquele, indício de ajustes mais adequados.

Figura 6: Curvas de sobrevivência comparadas ao modelo de Cox



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

Para escolha do modelo mais adequado, utilizou-se ainda, do Desvio Quadrático Médio (DQM) para verificar qual dos cinco modelos ajustados por Cox estimam valores (esperados) mais próximos aos observados (ocorridos), traçado por Kaplan-Meier. Esse resultado se baseou nas estimativas dispostas no Apêndice A e é apresentado na Tabela 8.

Tabela 8: Desvio Quadrático Médio de cinco modelos ajustados para estimar a função de sobrevivência, ou seja, a probabilidade de não gerar o custo assistencial pela primeira vez

Tempo	Desvio quadrático médio				
	Cox 8	Cox 12	Cox 1	Cox 3	Cox 14
1,0000	0,0000%	0,0000%	0,0000%	0,0000%	0,0000%
8,1000	0,0457%	0,0485%	0,0239%	0,0258%	0,0809%
9,1000	0,1411%	0,1497%	0,0741%	0,0799%	0,2484%
10,1333	0,2440%	0,2588%	0,1286%	0,1385%	0,4280%
11,1333	0,3802%	0,4031%	0,2011%	0,2165%	0,6642%
12,1667	0,6061%	0,6423%	0,3218%	0,3465%	1,0535%
13,2000	0,7632%	0,8086%	0,4062%	0,4373%	1,3227%
14,1333	0,9737%	1,0314%	0,5197%	0,5594%	1,6818%
15,1667	1,2107%	1,2821%	0,6483%	0,6977%	2,0837%
16,1667	1,4876%	1,5748%	0,7992%	0,8599%	2,5504%
17,2000	1,7895%	1,8940%	0,9646%	1,0376%	3,0566%
18,2000	2,1250%	2,2484%	1,1491%	1,2359%	3,6161%
19,2333	2,3653%	2,5022%	1,2818%	1,3784%	4,0153%
20,2667	2,6854%	2,8402%	1,4592%	1,5689%	4,5451%
21,2667	3,0452%	3,2198%	1,6594%	1,7839%	5,1381%
22,3000	3,4914%	3,6903%	1,9087%	2,0514%	5,8708%
23,3000	4,0311%	4,2592%	2,2116%	2,3763%	6,7529%
24,3333	4,6690%	4,9312%	2,5716%	2,7623%	7,7890%

25,3667	5,1434%	5,4306%	2,8408%	3,0508%	8,5551%
26,3000	5,8119%	6,1339%	3,2221%	3,4592%	9,6304%
27,3333	6,9769%	7,3584%	3,8912%	4,1753%	11,4918%
28,3333	8,5484%	9,0080%	4,8029%	5,1504%	13,9720%
29,3667	10,4968%	11,0495%	5,9514%	6,3769%	16,9966%
30,3667	14,2499%	14,9697%	8,2216%	8,7939%	22,6688%

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

Assim, opta-se pela apresentação dos resultados referentes a dois modelos: Cox 8, que possui menor AIC, e Cox 1, que apresenta menor DQM. Seus resultados são exibidos nas Tabelas 9, 10, 11 e 12.

Em relação a Cox 8, verifica-se, por meio da Tabela 9, que as covariáveis sexo, idade, 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 primeiro custo assistencial (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 à 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; quanto ao custo, o risco de gerar baixo custo é maior que o de alto custo (em $1,069 - 1 = 0,069 = 6,9\%$), enquanto que não houve efeito significativo do custo moderado.

Tabela 9: Resultados do modelo de Cox 8 ajustado para o tempo que leva o segurado a gerar o custo assistencial

Variável - Nível	$\hat{\beta}$	$e^{\hat{\beta}}$	Valor p
Sexo - Masculino	-0,01197	0,9881	0,018338
Idade	-0,001227	0,9988	<0,0001
Procedência - Honorários	0,2000	1,221	<0,0001
Procedência - Materiais e medicamentos	0,01617	1,016	0,30734
Procedência - Outros	0,1686	1,184	<0,0001
Procedência - SADT	0,2533	1,288	<0,0001
Custo baixo	0,06658	1,069	0,000117
Custo moderado	0,00005403	1,000	0,999756

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

Verificou-se, ainda, a suposição de proporcionalidade de riscos do modelo Cox 8 por meio do Teste de Schoenfeld. Seus resultados são apresentados na Tabela 10 e indicam rejeição de proporcionalidade dos riscos do modelo, já que o valor p global é inferior ao nível de significância de 5%. No entanto, analisando-se variável por variável, tem-se que sexo e

idade rejeitam a proporcionalidade, enquanto que alguns níveis de Procedência e de Custo não rejeitam.

Tabela 10: Teste de Schoenfeld para proporcionalidade de riscos do Modelo de Cox 8 ajustado para o tempo que leva o segurado a gerar o custo assistencial

Variável – Nível	r	χ	Valor p
Sexo - Masculino	-0,02515	101,52	<0,0001
Idade	-0,03377	184,66	<0,0001
Procedência - Honorários	-0,00349	1,96	0,161589
Procedência - Materiais e medicamentos	0,00590	5,59	0,018064
Procedência - Outros	-0,00568	5,18	0,022888
Procedência - SADT	-0,00361	2,08	0,149656
Custo baixo	0,00880	12,43	0,000423
Custo moderado	0,00441	3,12	0,07225
GLOBAL	NA	395,67	<0,0001

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

Em relação a Cox 1, verifica-se, por meio da Tabela 11, que as covariáveis sexo, idade e 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 primeiro custo assistencial (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 à procedência, como $e^{\hat{\beta}} < 1$ para todos os níveis explicitados, todos eles apresentam maior risco de falha em relação ao nível implícito, qual seja, diárias e taxas.

Tabela 11: Resultados do modelo de Cox 1 ajustado para o tempo que leva o segurado a gerar o custo assistencial

Variável - Nível	$\hat{\beta}$	$e^{\hat{\beta}}$	Valor p
Sexo - Masculino	-0,010622	0,989434	0,0363
Idade	-0,001327	0,998674	<0,0001
Procedência - Honorários	0,202084	1,223950	<0,0001
Procedência - Materiais e medicamentos	0,020411	1,020620	0,1971
Procedência - Outros	0,155815	1,168610	<0,0001
Procedência - SADT	0,242814	1,274831	<0,0001

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

Verificou-se, ainda, a suposição de proporcionalidade de riscos do modelo Cox 1 por meio do Teste de Schoenfeld. Seus resultados são apresentados na Tabela 12 e indicam rejeição de proporcionalidade dos riscos do modelo, já que o valor p global é inferior ao nível de significância de 5%. No entanto, analisando-se variável por variável, tem-se que sexo e idade rejeitam a proporcionalidade, enquanto que alguns níveis de Procedência não rejeitam.

Tabela 12: Teste de Schoenfeld para proporcionalidade de riscos do Modelo de Cox 1 ajustado para o tempo que leva o segurado a gerar o custo assistencial

Variável – Nível	<i>r</i>	χ	Valor p
Sexo - Masculino	0,02543	103,86	<0,0001
Idade	-0,03556	204,63	<0,0001
Procedência - Honorários	-0,00284	1,29	0,25569
Procedência - Materiais e medicamentos	0,00641	6,61	0,01016
Procedência - Outros	-0,00698	7,82	0,00516
Procedência - SADT	-0,00414	2,75	0,09762
GLOBAL	NA	364,51	<0,0001

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

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, no sentido em que gera uma melhor aproximação entre o modelo mal especificado e o modelo verdadeiro. Com base nessa ideia, optou-se pelo uso do modelo especificado e pela interpretação de seus efeitos sobre o risco estimado de geração do primeiro custo assistencial e a sobrevivência estimada.

4.3 Análise de sobrevivência na presença de riscos competitivos

Nesta subseção se estimam semiparametricamente as funções de risco específico por causa, ou seja, as funções de risco considerando-se cinco diferentes tipos de procedência, relativas à geração do primeiro custo assistencial gerado pelo segurado, 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 custo assistencial” são as gerações de custo oriundas de a) materiais e medicamentos, b) Honorário, c) Diárias e taxas, d) Outros, e) SADT.

Risco de gerar custo assistencial por sexo, idade e custo

A Tabela 13 exhibe 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 “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 custo assistencial oriundo de diárias e taxas: o risco do homem é 16,55% (= 1,1655 – 1) maior

que o da mulher ($e^{\hat{\beta}} > 1$); o aumento da idade tende a diminuir o risco ($e^{\hat{\beta}} < 1$); e o custo moderado é menos arriscado 46,77% ($= |0,5323 - 1|$) que o custo alto ($e^{\hat{\beta}} < 1$).

Em relação ao risco específico “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 custo assistencial oriundo de honorários: o risco do homem é 5,33% ($= |0,9467 - 1|$) menor que o da mulher o aumento da idade tende a diminuir o risco ($e^{\hat{\beta}} < 1$); o custo baixo é 332% ($= 4,32 - 1$) mais arriscado e o custo moderado é 225% ($= 3,25 - 1$) mais arriscado quando comparados ao custo alto ($e^{\hat{\beta}} < 1$).

Tabela 13: Resultados do efeito do sexo, idade e custo sobre as variáveis, para o risco do segurado a gerar o custo assistencial

Risco específico	Variável – Nível	$\hat{\beta}$	$e^{\hat{\beta}}$	Valor p
Procedência – Diárias e taxas	Sexo – Masculino	0,153223	1,165585	<0,0001
	Idade	-0,003451	0,996555	<0,0001
	Custo baixo	-0,068191	0,934082	0,357
	Custo moderado	-0,630456	0,532349	<0,0001
Procedência – Honorários	Sexo – Masculino	-0,0547202	0,9467500	<0,0001
	Idade	-0,00324228	0,9975790	<0,0001
	Custo baixo	1,4649118	4,3271617	<0,0001
	Custo moderado	1,180679	3,2565854	<0,0001
Procedência - Materiais e medicamentos	Sexo - Masculino	0,0580336	1,0597506	0,00558
	Idade	0,0044550	1,0044649	<0,0001
	Custo baixo	0,0996336	1,1047661	0,11960
	Custo moderado	-0,7768232	0,4598646	<0,0001
Procedência – Outros	Sexo - Masculino	0,2214392	1,2421669	<0,0001
	Idade	-0,0140222	0,986075	<0,0001
	Custo baixo	-0,9426422	0,3895971	<0,0001
	Custo moderado	-0,5354953	0,5853793	<0,0001
Procedência – SADT	Sexo - Masculino	-0,505307	0,9507245	<0,0001
	Idade	0,0060642	1,0060826	<0,0001
	Custo baixo	-1,0622361	0,3456820	<0,0001
	Custo moderado	-0,4933035	0,6106059	<0,0001

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

Quanto ao risco específico “Procedência – Materiais e Medicamentos”, 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 custo assistencial oriundo de materiais e medicamentos: o risco do homem é 5,9% ($= 1,0597506 - 1$) maior que o da mulher, o aumento da idade tende a aumentar o risco ($e^{\hat{\beta}} < 1$); o custo moderado é 54,01% ($= |0,4598646 - 1|$) menos arriscado quando comparados ao custo alto ($e^{\hat{\beta}} < 1$).

Em relação ao risco específico “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 custo assistencial oriundo de outros: o risco do homem é 24,21% ($= 1,2421669 - 1$)

maior que o da mulher, o aumento da idade tende a diminuir o risco ($e^{\hat{\beta}} < 1$); o custo baixo é 61,04% ($= |0,3895971 - 1|$) menos arriscado e o custo moderado é 41,46% ($= |0,5853793 - 1|$) menos arriscado quando comparados ao custo alto ($e^{\hat{\beta}} < 1$).

Quanto ao risco específico “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 custo assistencial oriundo de procedência: o risco do homem é 4,93% ($= |0,9507245 - 1|$) menor que o da mulher, o aumento da idade tende a aumentar o risco ($e^{\hat{\beta}} < 1$); o custo baixo é 65,43% ($= |0,3456820 - 1|$) menos arriscado e o custo moderado é 38,94% ($= |0,6106059 - 1|$) menos arriscado quando comparados ao custo alto ($e^{\hat{\beta}} < 1$).

5 CONCLUSÃO

Objetivou-se nesse trabalho, através de dados de uma operadora que optou pelo anonimato, analisar o comportamento biométrico de um grupo de segurados referente à geração de primeiros custos assistenciais pelo segurado, a partir de uma amostra (referente ao período de 2013 a 2015) de 160.572 primeiros usos (segurados).

A base conta com 93207 pessoas do sexo feminino e 67365 do sexo masculino; também distribuídos por causa tem-se a quantidade de pessoas que geraram o custo por causa específica onde pela causa “Diárias e Taxas” foram 7.000 pessoas, por “Honorários” foram 97229 pessoas, por “Materiais Medicamentos” foram 9.323 pessoas, por “Outros” foram 13.178 pessoas e por “SADT” foram 33.842 pessoas; O tempo médio que leva uma pessoa a gerar o custo assistencial é de aproximadamente 17 dias, sendo no mínimo 1 dia e no máximo aproximadamente 31 dias; Em relação à idade, em média, os segurados estão com aproximadamente 41 anos, quando fazem o uso do plano de saúde.

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 primeiro custo assistencial) estimadas pelo estimador de Kaplan-Meier, por nível de covariáveis (sexo, custo e procedência), são diferentes; a análise semiparamétrica permitiu, por meio do modelo de Cox selecionado (pelo menor AIC), estimar os riscos de geração de primeiro custo assistencial, tendo como principal resultado o fato de as covariáveis sexo, idade, custo (categorizado) e procedência exercerem efeitos significativos sobre a função de risco estimada (o homem apresenta menos risco que a mulher de gerar o primeiro custo assistencial; o risco de gerar o referido custo tende a diminuir com o aumento da idade; em relação à procedência, todos os níveis apresentam maior risco de falha em relação às diárias e taxas; quanto aos custos, o risco de gerar o custo assistencial pela categoria “baixo” é maior quando comparado a categoria “alto”, e pela categoria “moderado” se iguala ao risco da categoria “alto”).

Na presença de riscos competitivos se realizou somente a análise semiparamétrica, que permitiu, por meio do modelo de riscos específicos por causa, estimar os riscos de geração de primeiro custo assistencial para cada causa, 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 o custo assistencial pela categoria moderado é maior quando estes são oriundos de honorários; já pela categoria baixo, o risco é maior quando estes são oriundos de honorários.

Por fim, elencam-se as limitações deste trabalho e algumas sugestões para pesquisas futuras, a fim de idealizar a realização de pesquisas cada vez mais próximas dos problemas do mundo real:

- Seria interessante dar continuidade a pesquisa se utilizando de outras covariáveis, como escolaridade, renda, profissão, tipo de plano, etc.
- Além disso, poderiam ser ajustadas regressões paramétricas, bem como as semiparamétricas de subdistribuições de riscos, tanto no âmbito de ausência quanto no de presença de riscos competitivos.
- 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 primeiro custo assistencial por parte do segurado, incluindo também outros custos gerados, quais sejam, o segundo, o terceiro etc., ou seja, considerar a possibilidade de ampliação da matriz de transição. O modelo de eventos múltiplos seria capaz de modelar essa situação.

REFERÊNCIAS

ALDRIDGE, Melissa D.; KELLEY, Amy S.. Epidemiology of Serious Illness and High Utilization of Health Care. **Dying In America: Improving Quality And Honoring Individual Preferences Near The End Of Life.**, EUA, 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. *Regulação e saúde: estrutura, evolução e perspectivas da assistência médica suplementar*. Rio de Janeiro: ANS, 2002.

BRASIL. Constituição (1988). Lei nº 8.080, de 19 de setembro de 1990. . Disponível em: http://www.planalto.gov.br/ccivil_03/leis/18080.htm . Acesso em: 13 jul. 2019.

BRASIL. Constituição (1988). Lei nº 9961, de 28 de janeiro de 2000. Cria A Agência Nacional de Saúde Suplementar – Ans e Dá Outras Providências. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/19961.htm. Acesso em: 13 jul. 2019.

BRASIL. Lei nº 9656, de 03 de junho de 1998. Dispõe sobre os planos e seguros privados de assistência à saúde. Brasília, DF, 04 jun. 1998.

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

COLOSIMO, Enrico Antônio; GIOLO, Suely Ruiz. **ANALISE DE SOBREVIVÊNCIA APLICADA**. São Paulo: Edgard Blucher, 2006. 392 p.

CONASS. **Saúde suplementar**. 2011. Disponível em: https://www.conass.org.br/bibliotecav3/pdfs/colecao2011/livro_12.pdf. Acesso em: 28 jun. 2019.

CONASS. **Regulação em saúde**. Brasília: Copyright, 2007. 176 p. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232008000500032 . Acesso em: 08 jul. 2019

CONASS. **Saúde suplementar**. São Paulo: CONASS, 2015. Disponível em: https://www.conass.org.br/biblioteca/pdf/atualizacao-2015/L12_SAUDE-SUPLEMENTAR_jun2015.pdf . Acesso em: 01 jul. 2019.

FEDERAÇÃO NACIONAL DE SAÚDE SUPLEMENTAR (Brasil). **A dimensão social e econômica do setor de saúde suplementar**. sd. Disponível em: <http://midias.cnseg.org.br/data/files/06/E6/07/95/FBBA85108C5BC875A88AA8A8/Cartilha%20FenaSaude%20-%20Dimensao%20Social.pdf>. Acesso em: 21 jun. 2019.

GRAY, B. cmprsk: Subdistribution analysis of competing risks. R package version 2.2-7. n. 38, p. 1-13, 2014.

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

IESS. **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 . Acesso em: 15 ago. 2019.

IESS. **Conjuntura - Saúde Suplementar**. 35. ed. São Paulo: Instituto de Estudos de Saúde Suplementar, 2017. 9 p. Disponível em: https://www.iess.org.br/cms/rep/Conjuntura_julho2017.pdf . Acesso em: 08 ago. 2019.

IESS. **Série histórica da VCMH/IESS**. 2019. Disponível em: https://www.iess.org.br/cms/rep/historico_vcmh.pdf . Acesso em: 08 ago. 2019.

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

HJORT, Nils Lid. **On Inference in Parametric Survival Data Models**. 60. ed. Cidade: International Statistical Review, 1992. 355-387 p.

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.

LARA, Natalia. Texto para Discussão n° 72 - 2018 O Setor de Saúde na Perspectiva Macroeconômica. Período 2010 – 2015. São Paulo: Instituto de Estudos em Saúde Suplementar (IESS), 2018. 13 p. Disponível em: https://www.iess.org.br/cms/rep/td_72_saude_macroeconomia.pdf. Acesso em: 15 jul. 2019.

LARA, Natalia; REIS, Amanda; MINAMI, Bruno. **Texto para Discussão n° 71 - 2018 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. Acesso em: 15 jul. 2019

MONTONE, Januário. **O Impacto da Regulamentação no Setor de Saúde Suplementar**. Rio de Janeiro-rj: Ans, 2000. 46 p. Disponível em: http://www.ans.gov.br/images/stories/Materiais_para_pesquisa/Materiais_por_assunto/ProEditorialANS_Serie_ans_vol_1.pdf. Acesso em: 13 jul. 2019.

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).

REIS, Amanda; LARA, Natalia. **Nota sobre a Variação dos Custos Médico-Hospitalares (VCMH)**. São Paulo: Iess, 2019. 7 p. Disponível em: <https://www.iess.org.br/cms/rep/notaVCMHv2.pdf> . Acesso em: 16 ago. 2019.

REIS, Amanda. **Projeção das despesas assistenciais da Saúde Suplementar (2018-2030)**. Sao Paulo: Iess, 2018. 17 p.

RODRIGUES, José Ângelo. **Gestão de Risco Atuarial**. São Paulo: Saraiva, 2008. 442 p.

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.

SANTOS JÚNIOR, Luiz Carlos. **Análise de sobrevivência aplicada a premissas atuariais**: o caso da previdência pública municipal de Cabedelo/PB. 2018. 162 f. Tese (Doutorado) - Curso de Biometria, Bioestatística, Universidade Estadual Paulista, Botucatu, 2018.

SILVA, Amanda Reis Almeida; CARNEIRO, Luiz Augusto. **Tendências da variação de custos médico-hospitalares: comparativo internacional**. 2018. Disponível em: <https://www.iess.org.br/cms/rep/TD69.pdf> . Acesso em: 01 jul. 2019.

SILVA, R. P. Modelos flexíveis de sobrevivência com fração de cura: implementação computacional. 152 p. Dissertação (Mestrado) — Programa de Pós-Graduação em Matemática e Estatística Aplicada, Universidade Federal do Rio Grande do Norte, Natal, 2015.

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.sbmac.org.br/sbmac/article/view/1144/1157>. Acesso em: 12 mar. 2020.

SUPLEMENTAR, Agência Nacional de Saúde. **Caderno de Informação da Saúde Suplementar: Beneficiários, Operadoras e Planos**. 4. ed. Rio de Janeiro: Ans, 2017. 63 p. Disponível em: http://www.ans.gov.br/texto_lei_pdf.php?id=1762 . Acesso em: 26 ago. 2019.

THERNEAU, T. A package for survival analysis in s. n. 38, p. 1-5, 2015.

VISÃO SAÚDE: A REVISTA DOS PLANOS DE SAÚDE. São Paulo: Abrange, v. 1, 2016. Disponível em: <http://visaosaude.com.br/revista/visao-saude-v1.pdf> . Acesso em: 08 ago. 2019.

WATSON, Willis Towers. **Variação dos Custos Médico-Hospitalares das Operadoras de Saúde**. São Paulo: IESS, 2018. 14 slides, color.

APÊNDICES

APÊNDICE A – FUNÇÕES DE SOBREVIVÊNCIAS ESTIMADAS

A Tabela 14 apresenta os valores observados (empíricos), ou seja, “estimados por Kaplan-Meier (1958)” – \hat{S}_{KM} , e esperados (teóricos), isto é, “estimados por Cox (1972)”, $\hat{S}_{Cox\ 8}$, $\hat{S}_{Cox\ 12}$, $\hat{S}_{Cox\ 1}$, $\hat{S}_{Cox\ 3}$ e $\hat{S}_{Cox\ 14}$.

Tabela 14: Funções de sobrevivência estimadas pelo estimador de Kaplan-Meier e ajustadas por cinco modelos de Cox

Tempo (em dias)	Sobrevivência estimada					
	\hat{S}_{KM}	\hat{S}_{C8}	\hat{S}_{C12}	\hat{S}_{C1}	\hat{S}_{C3}	\hat{S}_{C14}
1,00	99,9994%	99,9995%	99,9995%	99,9995%	99,9995%	99,9995%
8,10	88,0452%	89,9685%	90,0277%	89,4282%	89,4819%	90,6226%
9,10	79,8003%	82,9149%	83,0119%	82,0337%	82,1211%	83,9862%
10,13	74,1817%	78,0368%	78,1577%	76,9409%	77,0496%	79,3742%
11,13	68,7044%	73,2193%	73,3621%	71,9296%	72,0575%	74,8005%
12,16	61,9716%	67,2037%	67,3711%	65,6987%	65,8479%	69,0601%
13,20	58,2692%	63,8469%	64,0266%	62,2356%	62,3952%	65,8416%
14,13	54,0960%	60,0184%	60,2108%	58,2989%	58,4693%	62,1567%
15,16	50,1532%	56,3540%	56,5572%	54,5449%	54,7241%	58,6143%
16,16	46,2615%	52,6876%	52,9001%	50,8032%	50,9898%	55,0536%
17,20	42,6506%	49,2373%	49,4571%	47,2956%	47,4880%	51,6871%
18,20	39,2129%	45,9046%	46,1300%	43,9210%	44,1174%	48,4206%
19,23	37,0401%	43,7721%	44,0002%	41,7690%	41,9673%	46,3222%
20,26	34,4469%	41,1981%	41,4288%	39,1796%	39,3794%	43,7805%
21,26	31,8686%	38,6054%	38,8374%	36,5809%	36,7811%	41,2098%
22,30	29,0648%	35,7436%	35,9759%	33,7241%	33,9237%	38,3592%
23,30	26,1378%	32,7040%	32,9349%	30,7039%	30,9013%	35,3149%
24,33	23,1971%	29,5911%	29,8187%	27,6275%	27,8210%	32,1774%
25,36	21,2995%	27,5468%	27,7712%	25,6171%	25,8071%	30,1049%
26,30	18,9572%	24,9792%	25,1979%	23,1045%	23,2887%	27,4874%
27,33	15,6023%	21,2029%	21,4101%	19,4364%	19,6092%	23,6040%
28,33	12,1553%	17,1776%	17,3680%	15,5668%	15,7237%	19,4109%
29,36	9,0632%	13,4069%	13,5760%	11,9877%	12,1251%	15,4207%
30,36	5,3216%	8,5486%	8,6799%	7,4609%	7,5649%	10,1580%
31,40	0,0000%	0,0025%	0,0026%	0,0014%	0,0015%	0,0053%

Interpreta-se a Tabela 13 do seguinte modo: em $t = 1$, tem-se $\hat{S}_{KM} = 99,9994\%$, ou seja, esta é a probabilidade (empírica) de o segurado pertencente ao plano de saúde analisado não gerar custo assistencial pela primeira vez no período observado. Seu complementar, isto é, $1 - 99,9994\%$, representa a probabilidade de gerar o referido custo. É possível fazer o mesmo tipo de interpretação para outros tempos (t) e outros modelos (de Cox) ajustados, tendo em vista que os modelos de Cox tentam captar o comportamento observado apontado pelo estimador de Kaplan-Meier.

APÊNDICE B - SCRIPT DA ANÁLISE EM LINGUAGEM R

###VIABILIZANDO LEITURA DOS DADOS

```
rm(list=ls(all=TRUE))
require(survival)
require(flexsurv)

setwd("C:/Users/Millys/Downloads/TCC II")
Dataset=read.table("jam.txt",header=TRUE)
summary(Dataset)
head(Dataset)
dim(Dataset)
attach(Dataset)
summary(SEXO)
names(Dataset)

summary(GRUPO)
is.factor(GRUPO)
tempo=Tempo
sle=STATUS
cau=GRUPO #Não é covariável
sex=SEXO
idade=IDADE
ca=CA
summary(ca)
shono
smatmed
sdt
sSADT
soutros

dados=data.frame(tempo,sle,cau,sex,ca,sSADT,soutros,shono,sdt,smatmed,idade)
head(dados)
summary(dados)
```

###AUSÊNCIA DE RISCOS COMPETITIVOS

##NÃO PARAMÉTRICO

```
#a) Kaplan-Meier
```

```
Surv(tempo,sle)
```

```
#Geral
```

```
ekm=survfit(Surv(tempo,sle)~1)
summary(ekm)
X11()
plot(ekm)
```

```
plot(ekm,main="",col=c("black","red","blue"),xlab="Tempo até a utilização do plano (em
dias)",ylab="S(t) estimada")
legend(5,0.3,col=c("black","red","blue"),c("Pontual", "Intervalar LI", "Intervalar
LS"),lwd=1,bty="n",lty=c(1,2,2))
```

```
#Geral por sexo
```

```
ekm1=survfit(Surv(tempo,sle)~sex)
```

```
summary(ekm1)
```

```
X11()
```

```
plot(ekm1,main="",col=c("black","red"),xlab="Tempo até a utilização do plano (em
dias)",ylab="S(t) estimada",lty=1:2)
```

```
legend(5,0.3,col=c("black","red"),c("Feminino", "Masculino"),lty=1:2,bty="n")
```

```
#Geral por causa
```

```
ekm2=survfit(Surv(tempo,sle)~cau)
```

```
summary(ekm2)
```

```
X11()
```

```
plot(ekm1,main="",col=c("black","red"),xlab="Tempo até a utilização do plano (em
dias)",ylab="S(t) estimada",lty=1:2)
```

```
legend(5,0.3,col=c("black","red"),c("Feminino", "Masculino"),lty=1:2,bty="n")
```

```
#Geral por custo
```

```
ekm3=survfit(Surv(tempo,sle)~ca)
```

```
summary(ekm3)
```

```
X11()
```

```
plot(ekm3,main="",col=c("black","red","green"),xlab="Tempo até a utilização do plano (em
dias)",ylab="S(t) estimada",lty=1:2)
```

```
legend(5,0.3,col=c("black","red","green"),c("alto", "baixo", "moderado"),lty=1:2,bty="n")
```

```
#comparando os custos
```

```
survdiff(Surv(c(tempo[ca=="MODERADO"],tempo[ca=="ALTO"]),c(sle[ca=="MODERAD
O"],sle[ca=="ALTO"])))~c(ca[ca=="MODERADO"],ca[ca=="ALTO"]),rho=0)
```

```
survdiff(Surv(c(tempo[ca=="BAIXO"],tempo[ca=="ALTO"]),c(sle[ca=="BAIXO"],sle[ca==
"ALTO"])))~c(ca[ca=="BAIXO"],ca[ca=="ALTO"]),rho=0)
```

```
survdiff(Surv(c(tempo[ca=="MODERADO"],tempo[ca=="BAIXO"]),c(sle[ca=="MODERA
DO"],sle[ca=="BAIXO"])))~c(ca[ca=="MODERADO"],ca[ca=="BAIXO"]),rho=0)
```

```
#comparando grupos
```

```
survdiff(Surv(c(tempo[cau=="dt"],tempo[cau=="hono"]),c(sle[cau=="dt"],sle[cau=="hono"])))
~c(cau[cau=="dt"],cau[cau=="hono"]),rho=0)
```

```
survdiff(Surv(c(tempo[cau=="dt"],tempo[cau=="matmed"]),c(sle[cau=="dt"],sle[cau=="mat
med"])))~c(cau[cau=="dt"],cau[cau=="matmed"]),rho=0)
```

```
survdiff(Surv(c(tempo[cau=="dt"],tempo[cau=="outros"]),c(sle[cau=="dt"],sle[cau=="outros"
]))~c(cau[cau=="dt"],cau[cau=="outros"]),rho=0)
```

```
survdiff(Surv(c(tempo[cau=="dt"],tempo[cau=="SADT"]),c(sle[cau=="dt"],sle[cau=="SADT
"])))~c(cau[cau=="dt"],cau[cau=="SADT"]),rho=0)
```

```
survdiff(Surv(c(tempo[cau=="hono"],tempo[cau=="matmed"]),c(sle[cau=="hono"],sle[cau==
"matmed"])))~c(cau[cau=="hono"],cau[cau=="matmed"]),rho=0)
```

```
survdiff(Surv(c(tempo[cau=="hono"],tempo[cau=="outros"]),c(sle[cau=="hono"],sle[cau==
"outros"])))~c(cau[cau=="hono"],cau[cau=="outros"]),rho=0)
```

```

survdifff(Surv(c(tempo[cau=="hono"],tempo[cau=="SADT"]),c(sle[cau=="hono"],sle[cau=="
SADT"])))~c(cau[cau=="hono"],cau[cau=="SADT"]),rho=0)
survdifff(Surv(c(tempo[cau=="matmed"],tempo[cau=="outros"]),c(sle[cau=="matmed"],sle[ca
u=="outros"])))~c(cau[cau=="matmed"],cau[cau=="outros"]),rho=0)
survdifff(Surv(c(tempo[cau=="matmed"],tempo[cau=="SADT"]),c(sle[cau=="matmed"],sle[ca
u=="SADT"])))~c(cau[cau=="matmed"],cau[cau=="SADT"]),rho=0)
survdifff(Surv(c(tempo[cau=="outros"],tempo[cau=="SADT"]),c(sle[cau=="outros"],sle[cau=
=="SADT"])))~c(cau[cau=="outros"],cau[cau=="SADT"]),rho=0)

```

##SEMI-PARAMÉTRICO

```

#regressao de cox
fit1=coxph(Surv(tempo,sle)~sex+idade+cau,data=dados)
summary(fit1)
#teste de schoenfeld
resid(fit1,type="scaledsch")
cox.zph(fit1, transform="identity")    ### g(t) = t
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht1=basehaz(fit1,centered=F)
tempo=Ht1$time
H01=Ht1$hazard
S01=exp(-H01)

```

```

#regressao de cox
fit2=coxph(Surv(tempo,sle)~sex+idade,data=dados)
summary(fit2)
#teste de schoenfeld
resid(fit2,type="scaledsch")
cox.zph(fit2, transform="identity")    ### g(t) = t
plot(cox.zph(fit2))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht2=basehaz(fit2,centered=F)
tempo2=Ht2$time
H02=Ht2$hazard
S02=exp(-H02)

```

```

#regressao de cox
fit3=coxph(Surv(tempo,sle)~cau+idade,data=dados)
summary(fit3)
#teste de schoenfeld
resid(fit3,type="scaledsch")
cox.zph(fit3, transform="identity")    ### g(t) = t
plot(cox.zph(fit3))
#para estimar a sobrevivência

```

```

#teste de schoenfeld
#ho= os riscos são proporcionais
Ht3=basehaz(fit3,centered=F)
tempo3=Ht3$time
H03=Ht3$hazard
S03=exp(-H03)

#regressao de cox
fit4=coxph(Surv(tempo,sle)~sex+cau,data=dados)
summary(fit4)
#teste de schoenfeld
resid(fit4,type="scaledsch")
cox.zph(fit4, transform="identity")    ### g(t) = t
plot(cox.zph(fit4))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht4=basehaz(fit4,centered=F)
tempo4=Ht4$time
H04=Ht4$hazard
S04=exp(-H04)

#regressao de cox
fit5=coxph(Surv(tempo,sle)~sex,data=dados)
summary(fit5)
#teste de schoenfeld
resid(fit5,type="scaledsch")
cox.zph(fit5, transform="identity")    ### g(t) = t
plot(cox.zph(fit5))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht5=basehaz(fit5,centered=F)
tempo5=Ht5$time
H05=Ht5$hazard
S05=exp(-H05)

#regressao de cox
fit6=coxph(Surv(tempo,sle)~cau,data=dados)
summary(fit6)
#teste de schoenfeld
resid(fit6,type="scaledsch")
cox.zph(fit6, transform="identity")    ### g(t) = t
plot(cox.zph(fit6))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht6=basehaz(fit6,centered=F)
tempo6=Ht6$time
H06=Ht6$hazard

```

```
S06=exp(-H06)
```

```
#regressao de cox
fit7=coxph(Surv(tempo,sle)~idade,data=dados)
summary(fit7)
#teste de schoenfeld
resid(fit7,type="scaledsch")
cox.zph(fit7, transform="identity")    ### g(t) = t
plot(cox.zph(fit7))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht7=basehaz(fit7,centered=F)
tempo7=Ht7$time
H07=Ht7$hazard
S07=exp(-H07)
```

```
#regressao de cox
fit8=coxph(Surv(tempo,sle)~idade+sex+cau+ca,data=dados)
summary(fit8)
#teste de schoenfeld
resid(fit8,type="scaledsch")
cox.zph(fit8, transform="identity")    ### g(t) = t
plot(cox.zph(fit8))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht8=basehaz(fit8,centered=F)
tempo8=Ht8$time
H08=Ht8$hazard
S08=exp(-H08)
```

```
#regressao de cox
fit9=coxph(Surv(tempo,sle)~idade+sex+ca,data=dados)
summary(fit9)
#teste de schoenfeld
resid(fit9,type="scaledsch")
cox.zph(fit9, transform="identity")    ### g(t) = t
plot(cox.zph(fit9))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht9=basehaz(fit9,centered=F)
tempo9=Ht9$time
H09=Ht9$hazard
S09=exp(-H09)
```

```
#regressao de cox
fit10=coxph(Surv(tempo,sle)~sex+cau+ca,data=dados)
summary(fit10)
```

```

#teste de schoenfeld
resid(fit10,type="scaledsch")
cox.zph(fit10, transform="identity")    ### g(t) = t
plot(cox.zph(fit10))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht10=basehaz(fit10,centered=F)
tempo10=Ht10$time
H10=Ht10$hazard
S10=exp(-H10)

#regressao de cox
fit11=coxph(Surv(tempo,sle)~sex+ca,data=dados)
summary(fit11)
#teste de schoenfeld
resid(fit11,type="scaledsch")
cox.zph(fit11, transform="identity")    ### g(t) = t
plot(cox.zph(fit11))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht11=basehaz(fit11,centered=F)
tempo11=Ht11$time
H11=Ht11$hazard
S11=exp(-H11)

#regressao de cox
fit12=coxph(Surv(tempo,sle)~idade+cau+ca,data=dados)
summary(fit12)
#teste de schoenfeld
resid(fit12,type="scaledsch")
cox.zph(fit12, transform="identity")    ### g(t) = t
plot(cox.zph(fit12))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht12=basehaz(fit12,centered=F)
tempo12=Ht12$time
H12=Ht12$hazard
S12=exp(-H12)

#regressao de cox
fit13=coxph(Surv(tempo,sle)~idade+ca,data=dados)
summary(fit13)
#teste de schoenfeld
resid(fit13,type="scaledsch")
cox.zph(fit13, transform="identity")    ### g(t) = t
plot(cox.zph(fit13))
#para estimar a sobrevivência

```

```

#teste de schoenfeld
#ho= os riscos são proporcionais
Ht13=basehaz(fit13,centered=F)
tempo13=Ht13$time
H13=Ht13$hazard
S13=exp(-H13)

#regressao de cox
fit14=coxph(Surv(tempo,sle)~cau+ca,data=dados)
summary(fit14)
#teste de schoenfeld
resid(fit14,type="scaledsch")
cox.zph(fit14, transform="identity")
plot(cox.zph(fit14))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht14=basehaz(fit14,centered=F)
tempo14=Ht14$time
H14=Ht14$hazard
S14=exp(-H14)

#regressao de cox
fit15=coxph(Surv(tempo,sle)~ca,data=dados)
summary(fit15)
#teste de schoenfeld
resid(fit15,type="scaledsch")
cox.zph(fit15, transform="identity")    ### g(t) = t
plot(cox.zph(fit15))
#para estimar a sobrevivência
#teste de schoenfeld
#ho= os riscos são proporcionais
Ht15=basehaz(fit15,centered=F)
tempo15=Ht15$time
H15=Ht15$hazard
S15=exp(-H15)

#ekm<-survfit(Surv(tempo,sle)~1)
time<-ekm$time

?par()
X11()
plot(ekm, conf.int=F, xlab="Tempo até a utilização do plano (em dias)", ylab="S(t)",pch=3)
lines(c(0,time),c(1,S01), lty=2,lwd=1,col="blue")
lines(c(0,time),c(1,S03), lty=3,lwd=1,col="darkgreen")
lines(c(0,time),c(1,S08), lty=4,lwd=1,col="green")
lines(c(0,time),c(1,S12), lty=2,lwd=2,col="red")
lines(c(0,time),c(1,S14), lty=3,lwd=3,col="darkblue")

```

```
legend(1,0.5,c("Kaplan-Meier","Cox 1","Cox 3","Cox 8","Cox 12","Cox
14"),bty="n",col=c("black","blue","darkgreen","green","red","darkblue"),lty=c(1,2,3,4,2,3),lwd=c(1,1,1,1,2,3))
```

```
#menor aic, melhor modelo
AIC(fit1,fit2,fit3,fit4,fit5,fit6,fit7,fit8, fit9,fit10,fit11,fit12,fit13,fit14,fit15)
AIC(fit8,fit12,fit1,fit3,fit14,fit10)
cbind(ekm$time,ekm$urv,S08,S12,S01,S03,S14,S10)
```

```
#####
```

```
#-----
```

```
###RISCOS COMPETITIVOS (CINCO CAUSAS DE SAÍDA)
```

```
### SEMIPARAMÉTRICO
```

```
## A) Risco Específico por Causa - eventos independentes
```

```
#sdt=stc
```

```
#Status = 1 se Causa = sdt
```

```
#Status = 0 caso contrário
```

```
#shono=sid
```

```
#Status = 1 se Causa = shono
```

```
#Status = 0 caso contrário
```

```
#smatmed=sinv
```

```
#Status = 1 se Causa = smatmed
```

```
#Status = 0 caso contrário
```

```
#soutros=sinv
```

```
#Status = 1 se Causa = outros
```

```
#Status = 0 caso contrário
```

```
#sSADT=sinv
```

```
#Status = 1 se Causa = sSADT
```

```
#Status = 0 caso contrário
```

```
Surv(tempo,sdt)
```

```
Surv(tempo,shono)
```

```
Surv(tempo,smatmed)
```

```
Surv(tempo,soutros)
```

```
Surv(tempo,sSADT)
```

```
#Efeito das 3 covariáveis sobre as causas
```

```
#Para Riscos Específicos
```

```
saida.espec1=coxph(Surv(tempo,cau=="dt")~sex+idade+ca,data=dados)
```

```
saida.espec1
```

```
cox.zph(saida.espec1)
par(mfrow=c(2,1))
plot(cox.zph(saida.espec1))
```

```
saida.espec2=coxph(Surv(tempo,cau=="hono")~sex+idade+ca,data=dados)
saida.espec2
cox.zph(saida.espec2)
par(mfrow=c(2,1))
plot(cox.zph(saida.espec2))
```

```
saida.espec3=coxph(Surv(tempo,cau=="matmed")~sex+idade+ca,data=dados)
saida.espec3
cox.zph(saida.espec3)
par(mfrow=c(2,1))
plot(cox.zph(saida.espec3))
```

```
saida.espec4=coxph(Surv(tempo,cau=="outros")~sex+idade+ca,data=dados)
saida.espec4
cox.zph(saida.espec3)
par(mfrow=c(2,1))
plot(cox.zph(saida.espec3))
```

```
saida.espec5=coxph(Surv(tempo,cau=="SADT")~sex+idade+ca,data=dados)
saida.espec5
cox.zph(saida.espec3)
par(mfrow=c(2,1))
plot(cox.zph(saida.espec3))
```