

Custos Assistenciais de uma Operadora de Planos de Saúde e o Efeito da Idade, do Estado de Sobrevivência e da Proximidade à Morte

Healthcare Expenditure of a Health Maintenance Organization and the Effect of Age, Survival Status and Proximity to Death

Rômulo Alves Soares

Mestre em Administração e Controladoria pela Universidade Federal do Ceará – UFC
Graduado em Ciências Atuariais pela Universidade Federal do Ceará – UFC
Professor substituto do Departamento de Administração da UFC
Avenida da Universidade, 2486 – Benfica, Fortaleza – CE
CEP: 60020-180

E-mail: romuloalves61@gmail.com

Cleber de Souza Gondim

Graduado em Ciências Atuariais pela Universidade Federal do Ceará – UFC
Avenida da Universidade, 2486 – Benfica, Fortaleza – CE
CEP: 60020-180

E-mail: cleversg@gmail.com

Alane Siqueira Rocha

Professora da Universidade Federal do Ceará – UFC
Avenida da Universidade, 2486 – Benfica, Fortaleza – CE
CEP: 60020-180

E-mail: alane.siqueira@yahoo.com.br

Resumo

A pesquisa teve como objetivo avaliar o efeito da idade, do estado de sobrevivência e da proximidade à morte sobre os custos assistenciais em uma operadora de planos de saúde (OPS). Para isso, foram utilizadas informações de 326.094 beneficiários de uma OPS do estado do Ceará, entre os anos de 2014 e 2015. As informações levantadas diziam respeito aos custos mensais desses usuários no período analisado, sua idade, estado de sobrevivência, sexo e quantidade de meses até o falecimento daqueles que vieram a óbito. Foram utilizados os testes T e de Wilcoxon para a comparação dos custos anuais de beneficiários sobreviventes e não sobreviventes, bem como a regressão quantílica, para verificar a influência das variáveis independentes sobre o custo assistencial. Empregou-se ainda a árvore de classificação e regressão e o *random forest*, para verificar o grau de importância das variáveis na determinação do custo. Os resultados mostram que, apesar de significativa, a idade tem uma importância reduzida na determinação dos custos da OPS quando considerada em conjunto com o estado de sobrevivência e a proximidade à morte. Face aos resultados obtidos, sugere-se incorporar, nas projeções de custos assistenciais das OPS, o estado de sobrevivência e a proximidade à morte de seus beneficiários, com vistas à geração de estimativas mais acuradas e permitindo um melhor gerenciamento de sua carteira.

Palavras-chave: Custos assistenciais, Idade, Estado de sobrevivência, Proximidade à morte.

Abstract

This research aimed to evaluate the effect of age, survival status and proximity to death over the healthcare expenditure of Health Maintenance Organizations (HMOs). In order to achieve

this goal, information about 326.094 beneficiaries of a HMO in Ceara state, between 2014 and 2015, were used. The information used covered their monthly expenditure, age, survival status, gender and months left to death. For the analysis, it was used T and Wilcoxon tests, in order to compare the annual expenditure of survival and non-survival beneficiaries, and quantile regression, in order to verify the influence of the independent variables over the health expenditure. Classification and regression tree and random forest were used to access the level of importance of each variable to determine the health expenditure. Results showed that, even though age was significant, it has a minor importance to determine health expenditure when analyzed together with survival status and proximity to death. These results point to the need of including those variables, survival status and proximity to death, in health expenditure projections of HMOs, allowing them to be more accurate and ensuring a better management.

Keywords: Health expenditure, Age, Survival status, Proximity to death.

1 Introdução

Ao longo das últimas décadas o sistema brasileiro de saúde pública e privada passou por transformações, seja no que tange ao seu ordenamento para o provimento de saúde, como também na sua inserção em um contexto de transição demográfica e epidemiológica (MENDES *et al.*, 2012). No Sistema Único de Saúde (SUS), sistema público de saúde brasileiro, tem-se, sobretudo, maior ênfase ao cuidado primário e à preocupação com a universalização do seu acesso. Enquanto o sistema de saúde suplementar passa, nos últimos tempos, por amplo processo de regulamentação. Esse processo tem como marco a criação da Agência Nacional de Saúde Suplementar (ANS), com a Lei 9.656/98, que permitiu uma maior organização e sustentabilidade ao setor (ANDRADE; MAIA; RODRIGUES, 2013).

Na saúde suplementar, foco deste trabalho, diversos são os fatores que afetam os custos assistenciais, tais como a transição demográfica, configurada por um envelhecimento populacional, novas tecnologias em saúde, novos medicamentos, judicialização do setor, intervencionismo estatal. Nesse cenário, para que se tenha melhores processos gerenciais e de tomada de decisão, a gestão de risco de planos de saúde busca compreender os efeitos do maior número possível de variáveis que afetam seus custos assistenciais.

A transição demográfica brasileira em curso, retratada por um constante aumento na proporção de idosos na população, exerce pressões sobre estimativas da demanda por serviços de saúde que respondam com êxito aos gestores do sistema de saúde como o envelhecimento pode afetar os gastos com saúde (MAIA; ANDRADE; FERES, 2012). Rodrigues e Afonso (2012) apontam que o envelhecimento populacional é um dos fatores mais discutidos quando se trata do aumento da demanda por serviços de saúde, e isso ocorre porque, segundo os autores, a relação entre idade e gastos com saúde é positiva e exponencial.

Rodrigues e Afonso (2012) destacam dois argumentos que colocam em cheque o papel de grande impulsionador dos custos com saúde que se atribui ao envelhecimento da população. O primeiro deles, segundo Rodrigues e Afonso (2012), está ligado à elevação proporcionalmente maior na idade média com que as doenças começam a ocorrer que o aumento observado na expectativa de vida. Fries (1980; 2000) e Cutler, Ghosh e Landrum (2013) mostram evidências, em países desenvolvidos, de adiamento do aparecimento de doenças crônicas não-transmissíveis, determinando a presença dessas doenças por um período menor de vida, fenômeno chamado de compressão da morbidade.

O segundo argumento está relacionado com a influência que a proximidade à morte pode ter sobre o padrão de utilização de serviços médicos. McGrail *et al.* (2000) ressaltam que a principal causa pela qual o dispêndio com saúde parece ser maior para os idosos é o fato de mais pessoas estarem perto da morte, recebendo tratamentos mais intensivos, na tentativa de

fazer com que este momento se afaste. Uma vez que as taxas de mortalidade são maiores nas idades mais avançadas, pode haver uma sobreposição entre a idade e a proximidade à morte que pode afetar a estimação dos gastos com saúde (MAIA; ANDRADE; FERES, 2012; RODRIGUES; AFONSO, 2012).

Assim, o objetivo do trabalho foi investigar o efeito da idade, do estado de sobrevivência e da proximidade à morte na determinação dos custos assistenciais de um plano de saúde. Do decorrer do estudo foram realizadas, também, as seguintes análises: i) comparação dos custos assistenciais entre beneficiários sobreviventes e não sobreviventes; ii) análise do comportamento dos custos assistenciais por faixa etária entre sobreviventes e não sobreviventes; e iii) mensuração da importância de diferentes variáveis independentes para a determinação dos custos assistenciais.

Foram utilizados registros administrativos provenientes de custos assistenciais da carteira de beneficiários de uma cooperativa médica de grande porte do estado do Ceará, tendo como referência o biênio 2014/2015. As técnicas estatísticas utilizadas para o alcance dos objetivos traçados foram os testes T e de Wilcoxon, a regressão quantílica, as árvores de classificação e regressão (CART) e o *random forest* (RF).

Esta pesquisa justifica-se, num primeiro momento, pela pouca evidência empírica encontrada sobre o tema sob análise no Brasil. Miranda, Mendes e Silva (2016) apontam que o envelhecimento populacional brasileiro tem ocorrido de modo diferente do que se observa para os países desenvolvidos, onde o processo ocorreu de modo mais lento. Ademais, o estudo também utiliza técnicas estatísticas ainda não aplicadas a esse contexto, o que permite comparar os resultados aqui encontrados com estudos anteriores.

O trabalho está dividido em seis partes, sendo a primeira delas esta introdução. Na segunda parte é feita uma revisão da literatura dos principais trabalhos que já trataram do tema. Em seguida descreve-se a metodologia utilizada para a consecução dos objetivos traçados. Na quarta parte são apresentados os resultados, os quais são discutidos na quinta parte do trabalho. Por fim, são feitas algumas conclusões acerca do trabalho realizado.

2 Revisão da literatura

O mercado de saúde suplementar apresenta características diferenciadas dos demais setores da economia. Arrow (1963) levanta algumas características que o distinguem dos demais bens e serviços fornecidos pelo mercado. Destacam-se as seguintes particularidades: a demanda por serviços de saúde não é estável em sua origem, mas sim irregular e imprevisível (CHUBAROVA, 2010); os serviços médicos, em geral, só são capazes de proporcionar satisfação na presença de doenças (LOCKER; DUNT, 1978); e a utilização de serviços médicos está, em geral, associada a uma probabilidade de morte ou prejuízo das capacidades funcionais do corpo (ARROW, 1963).

Arrow (1963) já destacava as dificuldades de predição intrínseca ao mercado de saúde por tratar com indivíduos, cada qual com seu organismo vivo e irreparável. As incertezas sobre a utilização de serviços de saúde e seus custos relacionados são uma preocupação para o setor de saúde suplementar que busca solvência em seus planos, a partir dos recursos das contraprestações pecuniárias dos usuários.

Cutler, Lleras-Muney e Deaton (2006) evidenciaram que para cada ano adicional na expectativa de vida de um indivíduo, o custo médio deste com saúde torna-se maior. Assim, o comportamento futuro da mortalidade ocupa papel importante quando se busca estimar os gastos no mercado de saúde.

Dentre os fatores que entram em discussão na literatura sobre a explicação do aumento observado dos custos com saúde, destacam-se dois: idade e proximidade à morte. Considerando uma maior utilização dos serviços de saúde, a população em idades mais avançadas tem sido vista como a causadora das variações nos custos na saúde, ou seja, a longevidade é vista como

determinante nesse aumento (MILLER, 2001). No entanto, uma série de estudos têm questionado essa visão, sugerindo que os custos na saúde são mais influenciados pela proximidade da morte de um indivíduo do que pela idade em si (GRAY, 2005).

A idade como fator de impacto nos custos da saúde já motivou vários estudos (BARER, EVANS; HERTZMAN, 1995; GERDTHAM *et al.*, 1992; MARZOUK, 1991; SMITH, HEFFLER & FREELAND, 1999; MILLER, 2001; SESHAMANI & GRAY, 2003). Contudo, uma das primeiras tentativas de analisar tal impacto foi feita em 1956, por Abel-Smith e Titmuss. Os autores projetaram um incremento nos custos hospitalares de 10,6%, a partir das mudanças demográficas ocorridas no período analisado (1951 a 1971). Entretanto, um estudo realizado pela *Royal Commission on the National Health System* de 1979 revelou que o aumento observado no período foi de 71%, e a contribuição da idade para esse aumento, segundo os autores do estudo, teve pouca relevância (GRAY, 2005).

A maioria dos cálculos para se chegar a essas análises partia do custo per capita ponderado pela população projetada em cada grupo etário, e os trabalhos tinham como motivação subsidiar iniciativas políticas e planejar investimentos em longo prazo na área da saúde.

Na linha de contestação da relevância do envelhecimento populacional na explicação da evolução dos custos com saúde, citam-se os estudos de Burner *et al.* (1992), Getzen (1992) e Seshamani e Gray (2002, 2003). Para os EUA, Burner *et al.* (1992) verificaram que o impacto do envelhecimento populacional seria de apenas 0,5% ao ano dos custos com a saúde, de um total de 8,4% projetado para o aumento dos custos anuais. Getzen (1992) analisando dados de 20 países membros da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), entre os anos de 1960 e 1988 constatou, assim como Seshamani e Gray (2002, 2003), que não havendo outras variáveis à equação, não há nenhuma associação discernível entre a estrutura etária e os custos na saúde.

Uma vez que as intervenções de cuidados de saúde são comuns nos últimos estágios de vida e que uma alta proporção dos indivíduos que morrem está sob atenção hospitalar, e que estudos mostram que a maioria dos custos individuais com saúde ocorre no último ou nos últimos dois anos de vida do indivíduo, independentemente de sua idade, os pesquisadores começaram a ter na variável proximidade à morte seu objeto de investigação (BERENSTEIN, 2009).

A idade não foi desconsiderada nos estudos que se seguiram, o que ocorreu foi um movimento para abranger outras variáveis que permitissem entendimento mais amplo do problema. Se o entendimento comum era que, em média, os idosos gastam mais que os jovens, considerando suas taxas de morbidade e mortalidade elevadas, e que tratamentos em pacientes terminais costumam ser mais elevados, a variável proximidade à morte poderia ser o complemento da variável idade, explicando assim o aumento dos custos com saúde ao longo da vida de um indivíduo (FELDER *et al.*, 2000).

Vários estudos passaram a incorporar a variável proximidade à morte para explicar ou simular a distribuição de gastos ligados à saúde. Conforme afirmam Yang, Norton e Stearns (2003), se o efeito da proximidade à morte é ignorado, então o aparente efeito da idade nos custos da saúde será exagerado, levando a uma superestimação dos custos.

Talvez o estudo mais influente utilizando a variável proximidade à morte seja o de Zweifel, Felder e Meier (1999). Com base em dados longitudinais de uma seguradora suíça, os autores analisaram a relação entre os custos com a saúde e a idade ao longo dos últimos oito trimestres de vida de indivíduos que morreram entre o período de 1983-1992. Eles concluíram que o envelhecimento da população se tratava de algo que desviava a atenção do real problema, a proximidade à morte. Os autores usaram a expressão *red herring* para descrever o fenômeno, que significa algo que desvia a atenção daquilo que é importante.

Felder (2001) concluiu que ao controlar a demanda de custos na saúde pela proximidade à morte, não há efeito da idade. Ao analisar a interação entre as variáveis idade e proximidade à morte, replicando o estudo de Zweifel, Felder e Meier (1999), Seshamani e Gray (2004) verificaram que se faz necessário uma atenção na escolha do modelo a ser aplicado, pois em um primeiro momento eles não encontraram evidências positivas no relacionamento das variáveis proximidade à morte e idade na explicação dos custos hospitalares.

Revisitando o estudo de Zweifel, Felder e Meier (1999), que após sua publicação recebeu muitas críticas quanto à modelagem utilizada, Zweifel, Felder e Werblow (2004) propõem a mesma análise, com uma nova abordagem metodológica, além da inserção de custos de indivíduos vivos. Ao final, eles corroboram a conclusão do estudo anterior, porém ressaltam o fato de que a idade, para os indivíduos vivos, pode ser um fator importante na explicação dos custos.

Mais recentemente, Yu, Wang e Wu (2015), com base em dados do *Taiwan National Health Insurance* no período de 2005 a 2009, mostraram que estimativas de envelhecimento populacional diminuem e se tornam insignificantes ao longo da curva de custos dos indivíduos analisados, sobretudo para a população de 65 anos ou mais. Para essa população, os coeficientes de proximidade à morte, estimados por meio de regressão quantílica, são significativamente positivos na maioria dos quantis adotados.

No Brasil, Berenstein (2009) utilizou dados dos custos de internações hospitalares do SUS, para o estado de Minas Gerais, entre os anos de 2004 e 2005. Além da idade e a proximidade à morte, utilizou também o sexo e o tempo de permanência do paciente no hospital, e concluiu que quanto mais próximo da morte, mais elevado será o custo do paciente, porém, quanto mais elevada a idade, menor o impacto da proximidade à morte nos custos.

Rodrigues e Afonso (2012), com base em dados de 1995 a 2007 do Sistema de Informações Hospitalares do DATASUS (SIH/DATASUS), analisaram a evolução dos gastos com saúde por grupos etários, segregando usuários sobreviventes e não sobreviventes. Os autores observaram que para os sobreviventes, o gasto apresentava uma tendência de crescimento com a elevação da idade, enquanto no grupo de não sobreviventes observou-se uma tendência de queda. Além disso, os autores também simularam a evolução dos gastos com saúde para o ano de 2050. Quando não se levava em consideração o estado de sobrevivência na simulação, o aumento anual estimado foi de cerca de 9% ao ano, enquanto o cenário que incorporou o estado de sobrevivência apresentou um crescimento de 1,6% ao ano.

Andrade, Maia e Rodrigues (2013), analisando indicadores de custos por estado de sobrevivência de uma operadora de autogestão de São Paulo no ano de 2009, mostram que o dispêndio daqueles que faleceram é cerca de 40 vezes superior ao daqueles que permaneceram vivos. Além disso mostram também que a maior parcela de utilização dos custos se concentra em procedimentos de internação, sendo observado nesta categoria a maior diferença entre sobreviventes e não sobreviventes.

Com base na evidência empírica encontrada anteriormente, diferente do que fora encontrado por Felder (2001), Zweifel, Felder e Meier (1999) e Zweifel, Felder e Werblow (2004), mas alinhado com os achados de Rodrigues e Afonso (2012), Seshamani e Gray (2004, 2004b) e Yu, Wang e Wu (2015), busca-se verificar, para beneficiários de plano privados de assistência à saúde, as seguintes hipóteses de trabalho: a) espera-se que tanto a idade como o estado de sobrevivência, para todos os beneficiários, exerçam influência significativa sobre os custos assistenciais de todos os beneficiários (Hipótese 1); b) espera-se que a idade e a proximidade à morte, para aqueles que não sobreviveram, exerçam influência sobre os custos assistenciais (Hipótese 2); e c) espera-se também que haja uma maior importância da proximidade à morte para determinar os custos assistenciais dos beneficiários não sobreviventes (Hipótese 3).

3 Metodologia

Nesta seção são apresentadas as características da base de dados utilizada para a extração das variáveis do estudo, bem como a descrição das técnicas de análise estatísticas aplicadas às variáveis.

3.1 Base de dados

A base de dados utilizada na pesquisa compreende informações dos custos assistenciais dos beneficiários de uma carteira de uma OPS, caracterizada como uma cooperativa médica de grande porte, localizada no estado do Ceará. As informações são oriundas de registros administrativos mensais dos anos de 2014 e 2015, gerados a partir da utilização de serviços cobertos pelo plano de saúde. Além dos custos assistenciais, outras informações levantadas foram a data de nascimento, sexo, data de entrada no plano, data de exclusão e motivo de desligamento junto à OPS.

Para os beneficiários que permaneceram ativos durante todo o período, foram utilizadas as informações referentes aos 12 meses de 2015. Para os usuários desligados do plano, durante o ano de 2015, foram considerados os 12 meses anteriores ao desligamento, sendo este o motivo da base utilizada compreender o período de 2014 a 2015. Segundo Andrade, Maia e Rodrigues (2013), esse tipo de análise, que segue o conceito de beneficiário-ano, permite corrigir possíveis erros de subestimação das observações censuradas no ano de 2015. Também foi aplicado à pesquisa o conceito de beneficiário sinistrado, isto é, foram considerados apenas aqueles que geraram custos à operadora em algum momento do período analisado. Considerando as delimitações utilizadas, ao final havia registros de utilização de 326.094 beneficiários, dos quais 1.261 tinha seu estado de sobrevivência classificado como óbito, ou seja, 0,39% dos beneficiários sinistrados vieram a óbito ao final do período analisado.

Existe, nessa análise, uma limitação que diz respeito aos beneficiários sobreviventes, uma vez que não se possui a data de falecimento destes beneficiários. Assim, quando se compara sobreviventes e não sobreviventes, considera-se que os sobreviventes permanecerão nessa condição no período seguinte ao da análise ou, pelo menos, em um tempo futuro previsível (ZWEIFEL; FELDER; WERBLOW, 2004).

Outra restrição identificada na base de dados se refere ao fato dos custos considerados serem tão somente os ocorridos dentro da rede de atendimento da OPS, ou seja, não foram computadas as demandas por procedimentos que foram solucionadas/atendidas por meio de atenção médico-hospitalar privada ou no SUS. Também podem ocorrer vieses nos custos totais dos beneficiários decorrentes de sinistros ocorridos e não avisados. Vale salientar, também, que os custos dos beneficiários não foram deflacionados por nenhum índice, portanto são custos observados na época do seu pagamento.

Quadro 1 – Variáveis utilizadas no estudo

Variável	Tipo	Operacionalização
Custo mensal	Variável dependente quantitativa	Custos de utilização do plano em um mês
Custo anual	Variável dependente quantitativa	Somatório dos últimos 12 meses dos custos de utilização
Estado de sobrevivência	Variável independente qualitativa	1 em caso de desligamento por falecimento, 0 caso contrário
Proximidade à morte	Variável independente qualitativa	Indica a quantidade de meses até o falecimento
Idade	Variável independente quantitativa	Idade, em anos completos, do beneficiário
Sexo	Variável de controle qualitativa	1 caso o beneficiário seja homem e 0 caso seja mulher

Fonte: Elaborado pelos autores.

Efeito da Idade, do Estado de Sobrevivência e da Proximidade à Morte nos Custos Assistenciais de uma Operadora de Planos de Saúde

De posse das informações levantadas, foram computadas as variáveis necessárias para a realização do estudo. A descrição e operacionalização dessas variáveis são apresentadas no Quadro 1.

4 Estratégia de análise

As variáveis foram utilizadas para a estimação de dois modelos. O primeiro deles (Equação I) segue os trabalhos de Andrade, Maia e Rodrigues (2013) e Rodrigues e Afonso (2012) e busca estimar a relação entre o custo anual e o estado de sobrevivência dos beneficiários, tendo a idade e o sexo como variáveis de controle. O segundo modelo (Equação II) segue as abordagens de Zweifel *et al.* (1999) e Yu *et al.* (2015) e busca estimar a relação entre o custo mensal e a proximidade à morte dos beneficiários não sobreviventes, novamente com idade e sexo como variáveis de controle.

$$CAA_i = \beta_0 + \beta_1 Idade_i + \beta_2 Sexo_i + \beta_3 Sobrevivência_i + \varepsilon_i \quad (I)$$

$$CAM_i = \beta_0 + \beta_1 Idade_i + \beta_2 Sexo_i + \beta_3 Proximidade_i + \varepsilon_i \quad (II)$$

Em que CAA_i , $Idade_i$, $Sexo_i$ e $Sobrevivência_i$ representam, respectivamente, a raiz quadrada do custo assistencial anual, a raiz quadrada da idade, o sexo e o status de sobrevivência do i -ésimo beneficiário. No segundo modelo, CAM_i e $Proximidade_i$ são respectivamente a raiz quadrada do custo assistencial mensal e o mês antes do óbito do i -ésimo beneficiário que veio a óbito.

Os dados levantados foram submetidos a uma série de análises estatísticas. Inicialmente, foi conduzida uma análise descritiva dos dados, com o intuito de facilitar a visualização das características dos beneficiários. Em conjunto com as análises descritivas, foram conduzidos testes de hipótese para comparar os custos assistenciais entre beneficiários sobreviventes e não-sobreviventes. Foram realizados o teste t e o de Wilcoxon para a comparação de duas amostras independentes.

Em seguida, para verificar se as variáveis independentes e de controle dos modelos são capazes de influenciar de modo significativo os custos assistenciais dos beneficiários, foi utilizada a regressão quantílica. A escolha dessa técnica se deu, em um primeiro momento, em razão da sua maior robustez em relação a *outliers* quando comparada com o método dos mínimos quadrados ordinários (MQO) (JONH, 2015). Essa robustez é importante uma vez que bancos de dados ligados a custos assistenciais de saúde, não raro, apresentam valores extremos (LUMLEY *et al.*, 2002). Além disso, o MQO baseia-se principalmente na distribuição central dos dados, enquanto a regressão quantílica permite o estudo em toda sua distribuição (YU; WANG; WU, 2015). Por fim, outra vantagem da utilização do modelo de regressão quantílico é sua robustez na ausência de normalidade e homocedasticidade da variável dependente (KOENKER; BASSETT JR., 1982).

Para além da significância estatística, buscou-se também atribuir um grau de importância para as variáveis independentes e de controle na determinação dos custos assistenciais. Para isso foram utilizadas a árvore de classificação e regressão (CART) e o *random forest* (RF).

A CART, proposta por Briman, Friedman, Olshen e Stone em 1984, é um dos tipos de árvore de decisão mais eficaz e mais conhecido que existe (TUFFÉRY, 2011). A lógica de uma árvore de decisão é segregar os indivíduos de uma amostra original, chamada de nó pai, em subamostras menores, com médias mais distintas possível e variância cada vez menor. O RF é um algoritmo, proposto por Breiman em 2001, que é capaz de gerar, de maneira aleatória, diferentes árvores de decisão em diferentes subespaços amostrais obtidos por meio do método de amostragem *bootstrap* (NISBET, ELDER; MINER, 2009). Para casos de regressão, a CART e o RF se apresentam como forma de medir a importância de uma variável na redução média

de impureza (*average impurity reduction*), que mensura o efeito de determinada variável sobre a pureza dos nós gerados (GRÖMPING, 2009). O RF apresenta ainda uma forma alternativa de medir a importância das variáveis, baseada no aumento do quadrado médio do erro (QME) quando os valores de uma variável são permutados (GENUER, POGGI; TULEAU-MALOT, 2010).

Todas as análises estatísticas foram executadas com o *software* R (R CORE TEAM, 2016). Para as estatísticas descritivas e os testes T e de Wilcoxon foi utilizado o pacote *stats*. Para a regressão quantílica foi utilizado o pacote *quantreg* (KOENKER, 2016). Para verificar a importância das variáveis por meio da CART, usou-se o pacote *rpart* (THERNEAU, ATKINSON; RIPLEY, 2015), enquanto para o *random forest*, foi feito uso do pacote *randomForest* (LIAW; WIENER, 2002).

5 Resultados

Seguindo a estratégia metodológica traçada, a análise dos dados inicia-se com o cálculo de estatísticas descritivas, com o intuito de compreender melhor a amostra segundo as suas características. Os resultados das estatísticas descritivas podem ser vistos na Tabela 1.

Tabela 1 – Estatística descritiva para a amostra completa por grupos de sobrevivência

		Beneficiários	Sobreviventes	Não sobreviventes
Sexo	Homens	135.244	134.738	506
	Mulheres	190.850	190.095	755
	Total	326.094	324.833	1.261
Idade	Média	35,96	35,80	76,58
	Mediana	34,00	34,00	81,00
	Desvio	22,21	22,08	17,89
	Mínimo	0,00	0,00	0,00
	Máximo	109,00	109,00	105,00
Custo (em R\$ 1,00)	Média	2.905,80	2.621,84	76.052,64
	Mediana	741,00	737,08	47.511,83
	Desvio	13.918,03	11.879,69	91.513,87
	Mínimo	1,40	1,40	30,00
	Máximo	1.008.070,00	1.008.070,00	634.986,90

Fonte: Elaborada pelos autores.

É possível perceber, num primeiro momento, que há uma pequena parcela de beneficiários que vieram a óbito no período analisado, totalizando 1.261, o que representa 0,39% de toda a amostra utilizada. Apesar disso, o custo total anual dos beneficiários que vieram a óbito representa 10,12% do total da amostra, indicando a relevância desse fator para os custos. Isso também é corroborado com os valores obtidos para as médias e medianas do custo quando se separam os grupos de sobreviventes e não sobreviventes. A média do último grupo é 29 vezes maior que a média dos sobreviventes, enquanto a mediana mais de 60 vezes superior. A observação dos valores máximos, bem como a comparação das médias e medianas dentro dos grupos revela ainda que os *outliers* são de fato uma preocupação no banco de dados.

Dando continuidade à análise, dada a evidência de que há uma diferença entre o custo dos beneficiários sobreviventes e não sobreviventes, foram realizados testes T e de Wilcoxon para verificar se esta diferença é estatisticamente significativa. Aproveitou-se ainda para realizar a comparação considerando grupos etários diferentes: inicialmente, os testes foram realizados com a amostra completa; em seguida, foram excluídos os beneficiários com menos de 60 anos

e realizado um novo teste; no passo seguinte, foram considerados apenas aqueles com idade igual ou superior a 70 anos; o processo foi repetido até que se tivesse apenas usuários com 90 anos ou mais. Os resultados são expostos na Tabela 2.

Tabela 2 – Teste T e de Wilcoxon para os custos dos sobreviventes e não sobreviventes

	Custo					Estatística de Teste	
	Média	Mediana	Desvio	Mínimo	Máximo	T	Wilcoxon
Beneficiários	2.905,80	741,00	13.918,03	1,40	1.008.070,00	28,49 ***	3,7·10 ⁹ ***
Sobreviventes	2.621,84	737,08	11.879,69	1,40	1.008.070,00		
Não sobreviventes	76.052,64	47.511,83	91.513,87	30,00	634.986,90		
Beneficiário (≥ 60)	7.867,50	1.967,10	26.781,15	2,60	1.008.070,00	26,22 ***	5·10 ⁷ ***
Sobreviventes	6.502,19	1.922,96	22.097,31	2,60	1.008.070,00		
Não sobreviventes	74.126,81	48.844,50	84.935,39	30,00	634.986,90		
Beneficiário (≥ 70)	9.879,40	2.168,44	30.434,08	2,6	634.986,90	23,93 ***	2,2·10 ⁷ ***
Sobreviventes	7.759,30	2.082,09	24.394,69	2,60	604.263,70		
Não sobreviventes	71.268,48	47.853,44	81.376,98	30,00	634.986,90		
Beneficiário (≥ 80)	12.540,10	2.118,68	35.906,30	2,60	634.986,90	19,14 ***	6,3·10 ⁶ ***
Sobreviventes	9.041,38	1.955,62	27.620,00	2,60	564.614,40		
Não sobreviventes	67.416,59	45.060,80	79.690,93	48,62	634.986,90		
Beneficiário (≥ 90)	16.060,10	2.137,97	36.435,75	2,60	392.195,80	12,45 ***	4,5·10 ⁵ ***
Sobreviventes	10.277,15	1.697,89	26.536,75	2,60	292.746,10		
Não sobreviventes	57.642,51	38.872,58	62.065,42	48,62	392.195,80		

Fonte: Elaborada pelos autores.

Níveis de significância: * < 0,1; ** < 0,05; *** < 0,01.

Os testes mostram que a diferença entre os grupos é significativa a 1%, independente do recorte etário utilizado. Destaque-se que as estatísticas de teste foram reduzindo à medida que as amostras eram compostas por usuários mais velhos. Cabe ressaltar também o comportamento das médias dos grupos de sobreviventes e não sobreviventes dentro dos grupos etários abordados. A média dos custos dos sobreviventes aumentam à medida que os beneficiários mais novos vão sendo excluídos, enquanto para o grupo de não sobreviventes ocorre o comportamento inverso, isto é, a média vai diminuindo.

Após a verificação de que existem diferenças estatisticamente significantes nos custos de sobreviventes e não sobreviventes, buscou-se verificar, por meio da regressão quantílica, como o estado de sobrevivência, a idade e o sexo dos beneficiários é capaz de influenciar o seu custo anual. Os resultados podem ser vistos na Tabela 3.

Tabela 3 – Resultados das regressões quantílicas e MQO para a amostra completa

Variável	Percentil 10	Percentil 25	Percentil 50	Percentil 75	Percentil 90	MQO
Intercepto	6,07***	7,66***	9,04***	9,78***	10,02***	8,46***
Estado (Vivo)	-1,95***	-2,84***	-3,42***	-3,45***	-3,17***	-2,92***
Idade	0,17***	0,21***	0,21***	0,22***	0,26***	0,22***
Sexo (Masculino)	-0,52***	-0,54***	-0,43***	-0,34***	-0,20***	-0,38***
Goodness-of-fit	0,04	0,06	0,09	0,10	0,11	-
R ²	-	-	-	-	-	0,15
R ² -ajustado	-	-	-	-	-	0,15
Teste F	-	-	-	-	-	1.913,00***

Fonte: Elaborada pelos autores.

Níveis de significância: * < 0,1; ** < 0,05; *** < 0,01.

Os resultados mostram que ao longo de toda a distribuição dos custos, as três variáveis exercem influência significativa a 1%. Para o estado de sobrevivência observa-se que usuários

que permanecem vivos tem custos menores quando comparados aos beneficiários que faleceram. Para a idade, percebe-se uma influência positiva sobre os custos, enquanto para o sexo, os homens têm custo anual inferior ao das mulheres.

Dado que as variáveis selecionadas se mostraram significantes para a determinação do custo anual dos beneficiários, aplicou-se em seguida a árvore de regressão (CART) e o *random forest* (RF) para mensurar a importância de cada uma dessas variáveis. Os resultados podem ser vistos na Tabela 4.

Tabela 4 – Importância das variáveis para a determinação do custo anual da amostra completa

Variável	CART (Pureza)	RF (Pureza)	RF (QME)
Estado	$6,77 \cdot 10^{12}$ (82,86%)	$1,99 \cdot 10^{12}$ (78,77%)	50,33%
Idade	$1,38 \cdot 10^{12}$ (16,90%)	$5,22 \cdot 10^{11}$ (20,70%)	26,35%
Sexo	$1,94 \cdot 10^{10}$ (0,24%)	$2,52 \cdot 10^{10}$ (0,53%)	3,19%

Fonte: Elaborada pelos autores.

Tanto no critério de pureza da CART, bem como nos dois critérios do *random forest*, foi atribuído ao estado de sobrevivência do beneficiário uma maior importância em relação à idade. Destaca-se ainda que o sexo, segundo as técnicas aplicadas, apesar de ser significativo no modelo de regressão, tem uma importância marginal para a determinação dos custos.

Nas seções anteriores detectou-se uma maior importância do estado de sobrevivência para a determinação dos custos e um custo médio do grupo de beneficiários não sobreviventes consideravelmente superior ao dos sobreviventes. Assim, buscou-se um melhor entendimento do comportamento dos custos do grupo de beneficiários não sobreviventes, incluindo uma variável adicional que indica a proximidade ao óbito. Em um primeiro momento, foram feitas análises de regressão, com resultados apresentados na Tabela 5.

Tabela 5 – Resultados das regressões quantílicas e MQO para a amostra não sobrevivente

Variável	Percentil 10	Percentil 25	Percentil 50	Percentil 75	Percentil 90	MQO
Intercepto	15,78***	48,20***	124,32***	281,08***	369,30***	155,25***
Meses p/ óbito (1)	-4,37**	-21,04***	-12,40	40,52***	57,88***	9,41***
Meses p/ óbito (2)	-6,14***	-30,57***	-56,73***	-28,93***	13,63	-24,88***
Meses p/ óbito (3)	-7,63***	-32,86***	-68,54***	-74,59***	-47,02***	-45,95***
Meses p/ óbito (4)	-8,10***	-35,48***	-75,84***	-96,43***	-74,28***	-57,32***
Meses p/ óbito (5)	-8,26***	-36,02***	-79,10***	-107,66***	-103,97***	-64,39***
Meses p/ óbito (6)	-8,20***	-35,63***	-79,07***	-106,36***	-109,44***	-67,96***
Meses p/ óbito (7)	-7,83***	-35,58***	-79,48***	-111,78***	-116,29***	-69,18***
Meses p/ óbito (8)	-8,24***	-36,73***	-82,69***	-115,12***	-119,27***	-71,31***
Meses p/ óbito (9)	-8,28***	-36,99***	-83,44***	-117,02***	-131,88***	-75,30***
Meses p/ óbito (10)	-8,25***	-36,54***	-82,93***	-121,24***	-128,97***	-74,55***
Meses p/ óbito (11)	-8,23***	-36,99***	-84,00***	-121,08***	-134,76***	-76,73***
Meses p/ óbito (12)	-8,25***	-37,45***	-83,56***	-122,27***	-133,91***	-76,45***
Idade	0,03*	-0,05	-2,53***	-13,24***	-16,61***	-4,86***
Sexo (Masculino)	-0,03	-0,73**	-1,59***	-5,01**	-5,25	-2,68**
<i>Goodness-of-fit</i>	0,01	0,03	0,09	0,17	0,18	-
R ²	-	-	-	-	-	0,17
R ² -ajustado	-	-	-	-	-	0,17
Teste F	-	-	-	-	-	154,50***

Fonte: Elaborada pelos autores.

Níveis de significância: * < 0,1; ** < 0,05; *** < 0,01.

Diferentemente da regressão que considera a amostra completa e o custo anual, quando foi considerado o custo mensal dos beneficiários que faleceram, para alguns percentis da distribuição de custos e em algumas das variáveis não se verificou significância estatística. Em algumas situações, o nível de significância foi superior a 1%.

Destaca-se que a idade apresenta uma influência negativa sobre o custo nos percentis de 50 a 90, com significância de 1%. Entretanto, a idade revelou baixa significância (10%) no percentil 10 e não foi significativa no percentil 25. Quanto à variável meses para o óbito, em geral, observa-se um aumento do custo assistencial mensal à medida que se aproxima do óbito. Além disso, é possível perceber que esse aumento do custo ocorre de modo mais acentuado a partir do quarto mês que antecede o falecimento.

O sexo não demonstrou significância nos pontos mais extremos utilizados para análise, os percentis 10 e 90, enquanto nos percentis 25 e 75 sua significância foi de 5%. Apesar disso, o comportamento observado é semelhante ao da análise da seção anterior, uma vez que os homens apresentaram custo menor que o das mulheres.

Assim como foi feito para a amostra completa, foram utilizadas a árvore de regressão e o *random forest* para mensurar a importância das variáveis para a determinação do custo mensal dos beneficiários que vieram a óbito, o que é apresentado na Tabela 6.

Tabela 6 – Importância das variáveis para a determinação do custo mensal dos não sobreviventes

Variável	CART (Pureza)	RF (Pureza)	RF (QME)
Meses p/ óbito	$5,54 \cdot 10^{11}$ (99,70%)	$4,12 \cdot 10^{11}$ (80,31%)	115,54%
Idade	$1,36 \cdot 10^9$ (0,24%)	$9,53 \cdot 10^{10}$ (18,54%)	49,50%
Sexo	$2,87 \cdot 10^8$ (0,05%)	$5,85 \cdot 10^9$ (1,15%)	3,93%

Fonte: Elaborada pelos autores.

Em todos os critérios, a quantidade de meses para o óbito dos beneficiários tem uma maior importância do que a idade e o sexo. Destaca-se que, para os usuários que faleceram, a idade tem uma importância relativa menor quando se compara com os resultados da amostra completa. O sexo novamente, apesar de demonstrar alguma significância estatística, tem uma importância marginal na determinação do custo dos usuários não sobreviventes.

Com base nos resultados obtidos, observa-se evidências para suportar as três hipóteses levantadas nesta pesquisa. Quando se consideraram todos os beneficiários, tanto a idade quanto o estado de sobrevivência se mostraram significantes (H_1) para explicar o custo. Quando a amostra passou a contar apenas com beneficiários que faleceram, a idade e a proximidade à morte foram significantes para a determinação do custo (H_2). Observou-se ainda que a proximidade à morte foi considerada a variável mais importante na determinação do custo mensal (H_3).

6 Discussão

De modo geral, as análises realizadas permitem concluir que, ao serem controlados o estado de sobrevivência dos beneficiários e a proximidade à morte para aqueles que vieram a óbito, a idade têm sua importância estatística reduzida no que diz respeito à determinação dos custos ligados à saúde. Há de se destacar também que a idade apresenta uma relação positiva com os custos da amostra como um todo, mas passam a ter uma relação negativa uma vez que são considerados apenas os beneficiários não sobreviventes.

A comparação dos custos dos beneficiários vivos e mortos vai em encontro dos resultados apresentados por Andrade, Maia e Rodrigues (2013) que mostram que o gasto de beneficiários que vieram a óbito foi cerca de 40 vezes maior que os de beneficiários que permaneceram vivos no plano de saúde oferecido pela SABESP. As autoras mostram que os

custos mais elevados dos beneficiários que vem a óbito podem ser atribuídos principalmente a internações.

As regressões realizadas para a amostra com todos os beneficiários mostram que tanto a idade quanto a proximidade à morte influenciam de modo significativo, no entanto a importância do primeiro indicador é superior à do último, concordando com os resultados encontrados por Seshamai e Gray (2004). Os resultados encontrados nesta pesquisa ainda convergem parcialmente com o que foi encontrado por Yu, Wang e Wu (2015), Zweifel, Felder e Meier (1999) e Zweifel, Felder e Werblow (2004), uma vez que estes trabalhos demonstram que, ao ser controlado o efeito da proximidade à morte, a idade perde sua significância para explicar os custos.

O resultado da pesquisa atesta a importância de serem levados em consideração o estado de sobrevivência dos indivíduos e a proximidade à morte para a estimação dos custos assistenciais ligados à saúde. Rodrigues e Afonso (2012) evidenciam essa importância ao simularem a evolução dos gastos com saúde no Brasil até 2050. A simulação baseada no perfil etário dos gastos médios revelou um aumento dos gastos de 380% entre 2007 e 2050, enquanto esse crescimento não passou de 70% ao ser levado em consideração o estado de sobrevivência. De modo semelhante, Seshamai e Gray (2004b) projetam um crescimento de 0,4% ao ano entre 2002 e 2026 para os gastos com saúde na Inglaterra quando a proximidade à morte é levada em consideração, taxa que cresce para 0,8% ao ano quando essa variável não é incluída na análise.

Serup-Hansen, Wickstrom e Kristiansen (2002) mostram-se em linha com esse pensamento quando afirmam que, quando é considerado o custo da morte, o efeito do envelhecimento sobre os custos é menor do que os métodos de estimação tradicionais que baseiam-se principalmente na idade. Os autores estimaram o aumento dos gastos com saúde para 19% da população dinamarquesa no período de 1995 a 2020, chegando a uma estimativa de 18,5%, para todo o período, usando métodos tradicionais de estimação. Ao levarem em consideração a elevação dos gastos no último ano de vida, os autores obtiveram um aumento estimado de 15,1%.

Nesticò *et al.* (2017) atribuem essa redução na projeção de gastos com saúde, ao se considerar a elevação dos gastos na iminência da morte, ao fato de, na fase terminal, pessoas idosas custam menos do que pessoas mais jovens. Os resultados da Tabela 2 desta pesquisa corroboram a afirmação de Nesticò *et al.* (2017), quando mostra que, à medida que os pacientes mais jovens são excluídos da amostra, o custo assistencial médio vai reduzindo. Segundo os autores, esse efeito da proximidade à morte sobre as projeções de gastos com saúde vem ganhando importância tanto no debate científico como no debate político, uma vez que redimensiona o efeito da idade na evolução desses gastos, considerando um cenário em que a expectativa de vida global também vem aumentando.

7 Conclusão

A pesquisa conduzida teve como objetivo analisar o efeito da idade, do estado de sobrevivência e da proximidade à morte na determinação dos custos assistenciais de um plano de saúde. Para tal utilizou-se dados referentes a 326.094 beneficiários no período de doze meses entre 2014 e 2015, dos quais 1.261 vieram a óbito durante o ano de 2015.

Os resultados obtidos mostram que todas as variáveis independentes elencadas no estudo foram capazes de influenciar os custos assistenciais. Apesar disso, observa-se uma menor importância da idade frente ao estado de sobrevivência e à proximidade à morte, o que indica que essas variáveis devem ser incluídas em projeções de custos, melhorando os processos gerenciais e a tomada de decisão por parte das OPS.

Também se observou que a idade apresenta um comportamento diferente ao se considerar apenas os beneficiários que faleceram. Quando são analisados todos os beneficiários,

a idade exerce influência positiva sobre os custos assistenciais. No entanto, quando é considerada apenas os beneficiários falecidos, essa influência passa a ser negativa.

O estudo corrobora, parcial ou integralmente, os resultados de outras pesquisas anteriores como as de Yu, Wang e Wu (2015) e Zweifel, Felder e Meier (1999) que mostram que a proximidade à morte é mais importante para explicar gastos com saúde do que a idade. Rodrigues e Afonso (2012) e Seshamani e Gray (2004b) também apresentam resultados convergentes em suas pesquisas quando mostram que a inclusão do estado de sobrevivência e da proximidade à morte são capazes de refinar as projeções futuras de gastos com saúde.

Em geral, os estudos mais difundidos para a estimação de custos com saúde dão destaque para as variáveis idade e sexo, com custo médio superior para idades mais avançadas e para as mulheres. Os resultados aqui encontrados dão suporte para a incorporação de outras variáveis, quais sejam estado de sobrevivência e tempo até a morte, nas estimativas de custos assistenciais com vistas a um melhor desvendamento do comportamento subjacente desses custos.

Dado que os recursos em uma OPS são finitos, financiadas quase que exclusivamente pelo mutualismo de seus beneficiários, o planejamento dessas entidades exige um esforço maior para garantir sua sustentabilidade. Dessa forma, faz-se necessária a criação de modelos que acompanhem os custos individuais dos beneficiários, os quais descrevam e encontrem, entre os serviços utilizados por esses beneficiários, similaridades que possam alcançar uma redução de custos para as OPS, ou ainda sua estimação de modo mais acurado.

O trabalho apresenta como limitações, inicialmente o fato de que, apensar de analisar um banco de dados bastante extenso, ele restringe-se a um contexto local, o que dificulta tentativas de generalização dos resultados. Além disso, a base de dados disponibilizada possuía poucas variáveis, sendo deixadas de fora outras variáveis que poderiam ser importantes na determinação do custo assistencial, tais como renda e escolaridade dos usuários.

Dada as limitações observadas, sugere-se para pesquisas futuras, a utilização de bancos de dados que compreendam contextos mais diversificados, além da inclusão de novas variáveis que possam afetar a dinâmica dos custos ligados à saúde. A segregação dos custos em grupos como internações, consultas e outros tipos de gasto também podem contribuir para o desenvolvimento do tema.

Referências

ABEL-SMITH, Brian; TITMUSS, Richard M. **The cost of the national health service in England and Wales**. Cambridge University Press, 2016.

ANDRADE, Mônica V.; MAIA, Ana C.; RODRIGUES, Cristina G. Indicadores de gastos com serviços médicos no setor de saúde suplementar no Brasil. **Revista Brasileira de Estudos de População**, p. 103-117, 2013.

ARROW, Kenneth J. Uncertainty and the welfare economics of medical care. **American Economic Review**, v. 53, n. 5, p. 941-973, 1963.

BARER, Morris L.; EVANS, Robert G.; HERTZMAN, Clyde. Avalanche or glacier?: health care and the demographic rhetoric. **Canadian Journal on Aging/La Revue canadienne du vieillissement**, v. 14, n. 02, p. 193-224, 1995.

BERENSTEIN, Cláudia K. **Os efeitos de idade e proximidade à morte sobre os gastos com internações no SUS: evidências com base no caso de Minas Gerais, 2004/2005**. 2009. Tese de Doutorado. Tese (Doutorado em Demografia) – Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte.

- BURNER, Sally T.; WALDO, Daniel R.; MCKUSICK, David R. National health expenditures projections through 2030. **Health Care Financing Review**, v. 14, n. 1, p. 1, 1992.
- CHUBAROVA, Tat'yana. The Health-Care System in Russia: Economic Problems of Theory and Practice. **Problems of Economic Transition**, v. 52, n. 11, p. 24-42, 2010.
- CUTLER, David M.; DEATON, Angus; LLERAS-MUNEY, Adriana. The determinants of mortality. **The Journal of Economic Perspectives**, v. 20, n. 3, p. 97-120, 2006.
- CUTLER, David M.; GHOSH, Kaushik; LANDRUM, Mary Beth. **Evidence for significant compression of morbidity in the elderly US population**. National Bureau of Economic Research, 2013.
- DUARTE MIRANDA, Gabriella Morais; GOUVEIA MENDES, Antonio da Cruz; ANDRADE DA SILVA, Ana Lucia. O envelhecimento populacional brasileiro: desafios e consequências sociais atuais e futuras. **Revista Brasileira de Geriatria e Gerontologia**, v. 19, n. 3, 2016.
- FRIES James F. Aging, natural death and the compression of morbidity. *NEJM* 1980; July 17: 130-5
- FRIES, James F. Compression of morbidity in the elderly. **Vaccine**, v. 18, n. 16, p. 1584-1589, 2000.
- GENUER, Robin; POGGI, Jean-Michel; TULEAU-MALOT, Christine. Variable selection using random forests. **Pattern Recognition Letters**, v. 31, n. 14, p. 2225-2236, 2010.
- GERDTHAM, Ulf-G. et al. An econometric analysis of health care expenditure: a cross-section study of the OECD countries. **Journal of health economics**, v. 11, n. 1, p. 63-84, 1992.
- GETZEN, Thomas E. Population aging and the growth of health expenditures. **Journal of gerontology**, v. 47, n. 3, p. S98-S104, 1992.
- GRAY, Alastair. Population ageing and health care expenditure. **Ageing Horizons**, n. 2, p. 15-20, 2005.
- GRÖMPING, Ulrike. Variable importance assessment in regression: linear regression versus random forest. **The American Statistician**, v. 63, n. 4, p. 308-319, 2009.
- JOHN, Onyedikachi O. Robustness of quantile regression to outliers. **American Journal of Applied Mathematics and Statistics**, v. 3, n. 2, p. 86-88, 2015.
- KOENKER, Roger. quantreg: quantile regression. R package version 5.05. **R Foundation for Statistical Computing**, Vienna, 2013. Disponível em <<http://CRAN.R-project.org/package=quantreg>>.
- KOENKER, Roger; BASSETT JR, Gilbert. Robust teste for heteroscedasticity based on regression quantiles. **Econometrica**, v. 50, n. 1, p. 43-61, 1982.
- LIAW, Andy; WIENER, Matthew. Classification and regression by randomForest. **R News**, v. 2, n. 3, p. 18-22, 2002.
- LUMLEY, Thomas *et al.* The importance of the normality assumption in large public health data sets. **Annual review of public health**, v. 23, n. 1, p. 151-169, 2002.
- MAIA, Ana Carolina; ANDRADE, Monica Viegas; FERES, Flávia Lúcia Chein. **Estudo longitudinal do efeito da idade e tempo até a morte em gastos com saúde**. REAP working paper, n. 37, 2012.

- MARZOUK, M. S. Aging, age-specific health care costs and the future health care burden in Canada. **Canadian Public Policy/Analyse de Politiques**, p. 490-506, 1991.
- MCGRAIL, Kimberlyn et al. Age, costs of acute and long-term care and proximity to death: evidence for 1987-88 and 1994-95 in British Columbia. **Age and Ageing**, v. 29, n. 3, p. 249-253, 2000.
- MENDES, Antonio C. G. M. et al. Assistência pública de saúde no contexto da transição demográfica brasileira: exigências atuais e futuras. **Cadernos de Saúde Pública**, v. 28, n. 5, p. 955-964, 2012.
- MILLER, Tim. Increasing longevity and medicare expenditures. **Demography**, v. 38, n. 2, p. 215-226, 2001.
- NISBET, R., MINER, G.; ELDER, J. **Handbook of statistical analysis and data mining applications**. Academic Press, 2009.
- NISTICÒ, F.; ALFIERI, W. De; TROIANO, G.; NANTE, N.; DEI, S.; PIACENTINI, P. Age-specific patterns of health care expenditure in dying people. **Public Health**, v. 152, p. 17-19, 2017.
- RODRIGUES, Cristina Guimarães; AFONSO, Luís Eduardo. O efeito do status de sobrevivência sobre gastos com internações hospitalares públicas no Brasil em uma perspectiva temporal. **Estudos Econômicos (São Paulo)**, v. 42, n. 3, p. 489-510, 2012.
- SERUP-HANSEN, Niels; WICKSTROM, Jannie; KRISTIANSEN, Ivar S. Future health care costs – do health care costs during the last year of life matter? **Health Policy**, v. 62, p. 161-172, 2002.
- SESHAMANI, Meena; GRAY, Alastair. Ageing and health care expenditure: the red herring argument revisited. **Health economics**, v. 13, n. 4, p. 303-314, 2004.
- SESHAMANI, Meena; GRAY, Alastair. Health care expenditures and ageing: an international comparison. **Applied health economics and health policy**, v. 2, n. 1, p. 9-16, 2002.
- SESHAMANI, Meena; GRAY, Alastair. The impact of ageing on expenditures in the National Health Service. **Age and Ageing**, v. 31, n. 4, p. 287-294, 2002.
- SESHAMANI, Meena; GRAY, Alastair. Time to death and health expenditure: an improved model for the impact of demographic change on health care costs. **Age and Ageing**, 2004b.
- SMITH, Sheila; HEFFLER, Stephen; FREELAND, Mark. The next decade of health spending: A new outlook. **Health Affairs**, v. 18, n. 4, p. 86, 1999.
- THERNEAU, Terry M. et al. rpart: Recursive partitioning. **R package version**, 2015.
- TUFFÉRY, Stéphane. **Data mining and statistics for decision making**. West Sussex: John Wiley & Sons, 2011.
- YANG, Zhou; NORTON, Edward C.; STEARNS, Sally C. Longevity and health care expenditures the real reasons older people spend more. **The Journals of Gerontology Series B: Psychological Sciences and Social Sciences**, v. 58, n. 1, p. 2-10, 2003.
- YU, Tiffany Hui-Kuang; WANG, David Han-Min; WU, Kuo-Lun. Reexamining the red herring effect on healthcare expenditures. **Journal of Business Research**, v. 68, n. 4, p. 783-787, 2015.
- ZWEIFEL, Peter; FELDER, Stefan; MEIERS, Markus. Ageing of population and health care expenditure: a red herring?. **Health economics**, v. 8, n. 6, p. 485-496, 1999.

ZWEIFEL, Peter; FELDER, Stefan; WERBLOW, Andreas. Population ageing and health care expenditure: new evidence on the “red herring”. **The Geneva Papers on Risk and Insurance Issues and Practice**, v. 29, n. 4, p. 652-666, 2004.