ESTUDO LONGITUDINAL DO EFEITO DA IDADE E TEMPO ATÉ A MORTE EM GASTOS COM SAÚDE

RESUMO

O objetivo desse artigo é analisar os determinantes dos gastos individuais com assistência a saúde, desagregando por status de sobrevivência. A população avaliada é constituída de indivíduos com mais de 40 anos de idade, beneficiários de um plano de saúde administrado por uma autogestão do estado de São Paulo. A carteira e os gastos com assistência a saúde financiados pela autogestão foram acompanhamos longitudinalmente durante 6 anos. O acompanhamento permitiu observar os eventos de morte nessa população e retrospectivamente, calcular o tempo de vida restante para a fração da população não sobrevivente. A separação dos gastos entre sobreviventes e não sobreviventes, tendo em vista o efeito significativo da distancia até a morte sobre gastos tem sido amplamente discutido na literatura da última década em função da sua importância para projeções de gastos com saúde, a partir das mudanças demográficas da população. Como as taxas de mortalidade específicas por idade têm se alterado ao longo do tempo, é fundamental que perfis de gasto por idade projetados incorporem os ganhos de longevidade. Os resultados mostraram que a proximidade da morte tem um efeito importante sobre os gastos. O aumento de gasto é observado desde 3 anos antes da morte, chegando a um efeito marginal de até 430% no ano da morte. Com relação a atributos demográficos, chamou atenção a importante diferença de gasto entre homens e mulheres, que fica evidenciado quando incorporamos o grupo de variáveis associadas a morte. Em média, homens têm um gasto inferior ao das mulheres. Homens não sobreviventes, entretanto, apresentam um gasto mais elevado.

1. INTRODUÇÃO

Nas últimas décadas, os gastos com saúde cresceram em todo o mundo (Gerdtham, 2000; Gottret, Schieber, 2006; Dormont et al, 2007). Uma das explicações para esse aumento é a mudança observada no perfil epidemiológico da população, com o aumento das doenças não transmissíveis como causa da mortalidade em detrimento das doenças transmissíveis (Omran, 2005; Gottret, Schieber, 2006). Esse padrão da morbidade e mortalidade se verifica tanto em países desenvolvidos, que já estão em fase avançada da transição epidemiológica, como em países em desenvolvimento. No Brasil, a transição epidemiológica tem ocorrido de forma distinta daquela observada em países desenvolvidos. Observa-se a sobreposição de etapas, na qual há predominância de doenças transmissíveis e não transmissíveis, caracterizando um duplo perfil epidemiológico (Schramm et al, 2004).

Concomitante à mudança epidemiológica, observa-se também uma mudança no perfil demográfico da população, resultante do envelhecimento. Essa mudança no perfil demográfico é resultante do processo de transição demográfica, o qual pode ser dividido em duas etapas. A primeira é caracterizada pelo declínio dos níveis de mortalidade, especialmente a infantil, e manutenção de altos níveis de fecundidade, resultando em uma população jovem e com rápido crescimento. A segunda etapa é caracterizada pela redução sustentada dos níveis de fecundidade, a qual determinou uma maior importância dos grupos mais idosos na pirâmide etária. Em países em desenvolvimento, especialmente latino-americanos e asiáticos, a mudança no perfil demográfico se dá de uma forma acelerada (Brito, 2007). No Brasil, a redução da mortalidade se iniciou na década de 1940, seguida, nos anos 1970, de uma redução importante nos níveis de fecundidade (Carvalho, 2004). Embora aconteça em ritmo e intensidade diferentes, a queda de fecundidade resultará em populações com perfil envelhecido e taxas de crescimento baixos, talvez negativas (Carvalho & Wong, 2008).

A combinação do envelhecimento populacional com o aumento das despesas segundo idade é frequentemente reportada como um dos maiores determinantes de gastos futuros com cuidado de saúde (Dormont et al, 2007). Empiricamente, entretanto, a hipótese de que o aumento da proporção de idosos responderia por grande parte da elevação dos gastos de saúde observados não foi corroborada (Getzen, 1992; Gerdtham et al, 1992; Hitiris, 1992; Barros, 1998). Zweifel e outros (1999) sugerem que a explicação para esse pequeno ou nulo impacto da composição sobre os gastos estaria na possível correlação espúria entre idade e gasto. O determinante relevante dos gastos não seria a idade por si, mas quanto as pessoas estão próximas da morte, hipótese que os autores denominaram *red herring*. Como as taxas de mortalidade específica são mais elevadas nas idades mais avançadas, a idade estaria captando o efeito da proximidade da morte, variável não observada prospectivamente.

Ao longo da última década, a hipótese de Zweifel foi testada por diversos autores (Seshamani, Gray, 2004; Zweifel et al, 2004; Werblow et al, 2007), que aprofundaram a discussão apresentando alternativas metodológicas e mensurando o impacto da proximidade da morte sobre os gastos a partir de diferentes fontes de dados. De modo geral, os resultados demonstraram que o gradiente do efeito da idade sobre os gastos é pouco pronunciado e que os gastos do último ano de vida tendem a diminuir, sobretudo a partir dos 80 anos.

Diante desse novo arcabouço, Stearns e Norton (2004) avaliaram as implicações empíricas da incorporação do tempo até a morte sobre projeções de gasto nos Estados Unidos. Os resultados mostraram que considerar a idade isoladamente aumenta de 9 a 15% os gastos projetados, num período de 20 anos. Mais recentemente, Shang e Goldman (2008) testaram o poder preditivo da expectativa de vida em substituição ao tempo até a morte, reportando também pouco poder preditivo da idade após a inclusão da expectativa de vida. Segundo as simulações apresentadas pelos autores, a partir de informações do Medicare, projeções de 2040 baseadas em gastos por idade chegam a ser 9% mais altas que aquelas baseadas na expectativa de vida; em 2080, a diferença chega a 22%. Essas diferenças preditivas reforçam a conclusão de que tempo até a morte deve ser incluído em modelos de gastos de saúde e que essa incorporação é fundamental para projeções de gasto a partir de mudanças demográficas na população, uma vez que vivemos importantes aumentos de longevidade.

O objetivo deste trabalho é analisar o comportamento dos gastos com serviços de saúde financiados por uma operadora de plano de saúde, distinguindo atributos associados a indivíduos não sobreviventes. Como vimos, no Brasil ainda persiste um duplo perfil epidemiológico, o que pode determinar perfis de gasto controlados pela proximidade da morte diferentes daqueles já avaliados para economias que completaram a transição epidemiológica. Além disso, a transição demográfica brasileira, já em curso, exerce pressões sobre previsões de demanda que logrem responder aos gestores do sistema de saúde como o envelhecimento pode pressionar gastos agregados em saúde.

A base de dados utilizada no trabalho consiste de informações longitudinais dos beneficiários vinculados à carteira de um plano de saúde do estado de São Paulo. O acompanhamento é de seis anos, e as informações provêm de registros administrativos da operadora. Nessa carteira, os beneficiários do plano de saúde têm cobertura médico-hospitalar de todos os procedimentos previstos pela modalidade ambulatorial e hospitalar com obstetrícia, conforme rol de procedimentos estabelecido pela Agência Nacional de Saúde Suplementar (ANS). Os registros administrativos da operadora contabilizam todos os gastos referentes aos procedimentos utilizados pelos beneficiários. A maior vantagem de uma base de dados restrita a uma carteira de beneficiários é a ausência de diferenciais de acesso aos serviços de saúde cobertos pelo plano de saúde. Compilações mais representativas da população podem perder precisão na observação da demanda, na medida em que diferenças de acesso se traduzem no uso dos serviços. No Brasil, há poucas informações disponíveis, em

nível individual, sobre gastos com cuidado de saúde. O banco de dados mais difundido é o Sistema de Informações Hospitalares do SUS, com registro dos pagamentos das internações em hospitais que integram a rede do sistema público; as maiores limitações dessa base de dados são a ausência de um acompanhamento longitudinal das informações e a defasagem no ajuste dos valores remunerados.

Na literatura nacional, esta abordagem é inédita, excetuando-se a análise descritiva apresentada em Andrade et al (2010), que distingue gastos segundo perfil etário e *status* de sobrevivência.

Para estimar o efeito da morte e da proximidade da morte sobre os gastos com saúde, adotamos o método das equações de estimação generalizada, também denominado modelo de média populacional, e restringimos a análise à população acima de 40 anos de idade.

Os resultados mostraram uma grande importância do conjunto de variáveis associadas à morte, reforçando a necessidade das projeções de gasto com saúde distinguirem sobreviventes e não sobreviventes, sobretudo se as mudanças demográficas prevêem quedas nas taxas de mortalidade específica. O crescimento dos gastos é bastante significativo a medida que se aproxima a morte e, além disso, existe um diferencial importante de gastos entre homens e mulheres segundo status de sobrevivência. Sem a incorporação destas variáveis, homens apresentam um gasto, controlado por idade, levemente maior do que mulheres. Com a incorporação, homens apresentam um gasto inferior ao das mulheres, se sobreviventes, relação que se inverte entre não sobreviventes no painel. Quanto ao efeito da idade, mesmo com a inclusão das variáveis associadas a morte, ela continua apresentando um gradiente significativo sobre os gastos, que tende a ser pouco menos importante entre não sobreviventes.

2. BASE DE DADOS E METODOLOGIA

2.1 Base de dados e tratamento das informações

A base de dados utilizada neste trabalho é um painel anual de gastos individuais com serviços de saúde computado entre 2004 e 2009. Os gastos correspondem aos serviços cobertos por uma operadora de autogestão do estado de São Paulo. A referida operadora é a Sabesprev, que administra a carteira de beneficiários da Sabesp, a qual oferece a seus funcionários, dependentes e familiares uma cobertura de plano de saúde na forma de autogestão.

Os empregados com vínculo ativo com a empresa, titulares do plano de saúde, e seus respectivos dependentes têm seus prêmios parcialmente subsidiados pela empresa. A parcela do prêmio paga é taxada na forma de alíquota sobre os rendimentos do titular, independentemente do número de dependentes. A adesão ao plano da empresa é compulsória para empregados, titulares do plano, e facultada, sem custo adicional, a dependentes. Os empregados da ativa ainda podem agregar familiares à carteira da seguradora, mediante o pagamento de um prêmio de risco individual que é determinado segundo a faixa etária do beneficiário. A adesão de dependentes e familiares de empregados inativos (aposentados e ex-empregados) é opcional, e o prêmio é cobrado individualmente segundo a faixa etária, obedecendo às normas da legislação vigente.

Para constituir o painel, contabilizamos todos os gastos efetuados pela operadora com cada beneficiário anualmente, a partir de 2004. Originalmente, usamos dois bancos de registros administrativos da operadora. O primeiro, denominado despesa, tem periodicidade mensal e consiste de todos os registros de procedimentos realizados pelos beneficiários da carteira. Esses procedimentos são registrados com a finalidade de pagar ao provedor e consistem da informação desagregada de cada item do cuidado, data de utilização, provedor e

o respectivo valor pago pela operadora. O segundo banco de dados também tem periodicidade mensal e é denominado cadastro. Neste, registram-se atributos do beneficiário, como data de nascimento e sexo, além de informações a respeito do vínculo na carteira (titular, dependente ou agregado), data de entrada no plano, data de saída e motivo de desligamento, para inativos.

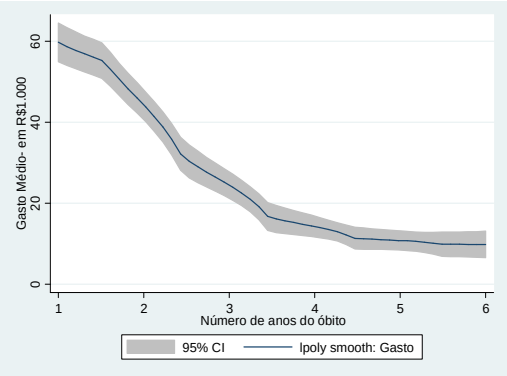
Primeiramente, reorganizamos o banco de despesa. Embora a Sabesprev pague mensalmente sua rede de provedores, há uma diferença entre o mês/ano de uso do cuidado e o mês/ano do pagamento. Essa diferença requer a reorganização das informações segundo a data de utilização. Além disso, como as informações estão registradas a preços correntes, é preciso corrigir os valores nominais por um índice de preços. Depois de reorganizar o banco de despesa segundo o mês e o ano da utilização, corrigimos mensalmente todos os valores pagos pelo Índice de Preços ao Consumidor (IPC) (grupo saúde) da Fundação Instituto de Pesquisas Econômicas (FIPE), tendo como base dezembro de 2008. Seguiu-se a organização dos 72 meses de cadastro, de forma a manter as informações sobre os atributos dos beneficiários e sobre seu vínculo na carteira mensalmente. Associada ao motivo de desligamento para os inativos, a condição de ativo ou inativo em dado mês/ano de referência nos permitiu aplicar dois filtros. Mantivemos no banco de dados todos beneficiários com vínculo ativo durante os 72 meses, como também os beneficiários faleceram ao longo desse período. Durante os seis anos, observamos 1.209 saídas devido a óbito, e um total de 43.163 beneficiários se manteve todo o tempo.

Na organização final do banco de dados, computamos retrospectivamente o número de meses até o óbito, quando era o caso. Depois desse tratamento inicial, organizarmos o banco de dados, até então com periodicidade mensal, de modo que o período de referência passasse a ser anual, para o quê, consideramos o tipo de informação de cada variável. No caso dos gastos, interessava o valor despendido durante todo o ano, de forma que somamos todos os gastos mensais. Quanto às informações mensais sobre tamanho da família e idade do beneficiário, calculamos a média anual da idade e o tamanho médio da família durante o ano. A respeito do vínculo do beneficiário com o plano (agregado, dependente ou titular) e do tipo de plano (pleno, básico ou especial), como pode haver mudanças ao longo do ano – devido, por exemplo, a alguma transição de plano por aposentadoria ou maioridade de dependente –, mantivemos os vínculos e planos que vigoraram na maior parte do ano.

Embora a base de dados utilizada aqui tenha características específicas, provavelmente atributos associados à inserção do titular no mercado de trabalho, a taxa de mortalidade foi bastante parecida com a do estado de São Paulo. Quanto à população com mais de 60 anos, a taxa de mortalidade (média dos seis anos de acompanhamento) foi de 3,07%, contra 3,6% no estado de São Paulo em 2010, segundo dados da Fundação Seade. Entre a população da carteira que esteve ativa durante os seis anos, a taxa de mortalidade (média dos seis anos) foi de 0,46%.

O gráfico 1 mostra o comportamento do gasto médio anual segundo o número de anos até a morte, obtido por regressão polinomial local ponderada, estimada por funções de Kernel. A área sombreada em torno da linha corresponde ao intervalo de confiança de 95% da média. Verifica-se um aumento significativo do gasto, à medida que aumenta a proximidade da morte.

GRÁFICO 1 – Gasto médio anual segundo número de anos até a morte 2004-2009 – carteira Sabesprev



fonte: Elaborado pelas autoras a partir de dados da Sabesprev.

Para entender melhor a variável proximidade da morte, é preciso observar todos os indivíduos incluídos na análise durante um tempo mínimo. Entre não sobreviventes, a maior limitação é não observarmos completamente todos os períodos que antecederam a morte, criando uma censura à esquerda. Em um painel de seis anos, observamos, em média, três anos antes da morte, com um máximo de seis anos para aqueles que morreram no último ano. A censura, entretanto, é mais importante para os indivíduos sobreviventes. Nesse caso, há uma censura à direita na informação de mortalidade (Zweifel et al, 2004). A censura ocorre porque não observamos a data de morte para indivíduos sobreviventes durante o período de tempo analisado. A informação retrospectiva de tempo até a morte só está disponível para indivíduos que morreram no período observado. Assim, quando comparamos sobreviventes e não sobreviventes, assumimos implicitamente que sobreviventes persistem nessa condição no período seguinte ou, pelo menos, num tempo futuro previsível. Se não for tratado, esse tipo de censura pode sobreestimar o efeito da idade sobre os gastos, na medida em que incluímos como sobrevivente um grupo de indivíduos que já está próximo da morte. Uma estratégia advogada por alguns autores para lidar com a censura consiste na exclusão dos sobreviventes da análise (Seshamani; Gray, 2004), mas ela tem a desvantagem de inviabilizar a análise do gasto segundo a idade, entre sobreviventes.

Adotamos aqui uma forma alternativa de atenuar esse problema, excluindo do painel o período do qual desconhecemos a ocorrência de morte dentro de algum tempo predeterminado (Steimnmann et al, 2007). Do ponto de vista prático, essa estratégia demanda um acompanhamento longitudinal e implica a exclusão ao menos das observações do último ano. Como dispomos de um acompanhamento de seis anos de uma carteira com mais de 40 mil vidas, temos mais de 250 mil observações. Se optássemos por indicadores de tempo até a morte desde seis anos até a morte, teríamos que excluir as últimas cinco observações, o que reduziria drasticamente o número de observações. Como solução intermediária, excluímos os últimos três períodos de tempo, e adotamos quatro indicadores de tempo ate a morte: ano da morte, um ano antes da morte, dois anos antes da morte e três anos antes da morte.

Além disso, uma segunda restrição, inerente à base de dados utilizada, foi a baixa frequência de mortes até os 40 anos de idade. Nessa faixa etária, registramos 104 óbitos, inviabilizando uma estimativa consistente para esse grupo. Assim, optamos por estimar o modelo de gasto na população acima de 40 anos.

Na tabela 1, apresentamos a estatística média de sobreviventes e não sobreviventes (todos com mais de 40 anos). As diferenças de idade e de gasto médio por ano são significativas. Os não sobreviventes apresentam uma idade média de 69 anos, com gasto médio 11 vezes superior ao dos sobreviventes, que têm em média 51 anos. Quanto às

diferenças de sexo/gênero, mulheres sobreviventes apresentaram um gasto médio maior que homens sobreviventes, relação que se inverte entre não sobreviventes. Fatores biológicos (sexo) e comportamentais (gênero) explicam as diferenças entre a saúde do homem e da mulher. Mulheres reportam maior prevalência de doenças crônicas (Koopmans; Lamers, 2007) e também tendem a apresentar um comportamento mais preventivo que homens (Vegda, 2009). Entre não sobreviventes, provavelmente homens tendem a descobrir o problema de saúde em estágios mais avançados da doença, o que resulta em gasto maior.

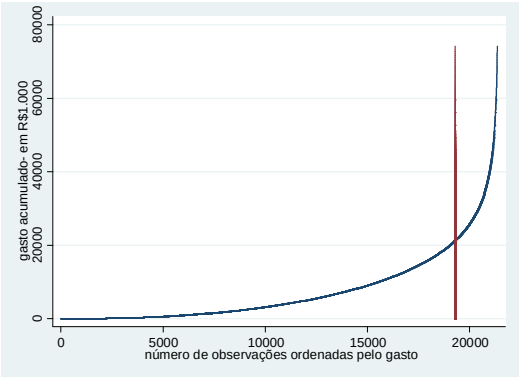
TABELA 1 – Estatística descritiva segundo *status* de sobrevivência 2004-2006 – carteira Sabesprev selecionada

média (desvio padrão)	**homens**	**mulheres**	**total**
sobreviventes			
idade	50,68	52,59	51,62
número de pessoas	10.361	10.066	20.427
número de pessoas período	31.083	30.198	61.281
gasto médio anual (em R$ de dez. de 2008)	2.104 (8.169)	2.713 (9.245)	2.405 (8.721)
prob. (gasto > 0)	0,93	0,94	0,94
não sobreviventes			
idade	65,45	72,26	68,93
número de pessoas	542	563	1105
número de pessoas período	1347	1414	2.761
gasto médio anual (em R$ de dez. de 2008)	30.532 (90.882)	24.442 (57.141)	27.414 (75.557)
prob. (gasto > 0)	0,90	0,88	0,89

fonte: Elaborada pelas autoras a partir de dados da Sabesprev.

Na distribuição dos gastos da população da carteira com mais de 40 anos, observamos uma grande assimetria. Em 2005, 10% dos beneficiários com maior gasto foram responsáveis por 72% do gasto total da carteira. O gráfico 2 mostra a distribuição do gasto acumulado em 2005, com linha vertical delimitando 10% dos beneficiários que mais gastaram no ano, à direita. O gasto total, desse grupo de indivíduos, nesse ano foi de aproximadamente 75 milhões de reais, sendo 21 milhões gastos por 90% dos usuários, e o restante, pelos 10% com maior gasto.

GRÁFICO 2 – Gasto acumulado em 2005 – carteira Sabesprev selecionada



fonte: Elaborado pelas autoras a partir de dados da Sabesprev.

2.2 Método de estimação

As informações individuais sobre custos/gastos com cuidados de saúde têm características particulares e relevantes para a modelagem econométrica. A primeira se refere à natureza não negativa da variável, e a segunda, a sua distribuição. Em termos populacionais, a distribuição dos gastos com saúde apresenta tipicamente uma fração de zeros e uma cauda longa, assimétrica à direita (Jones, 2010). Os zeros representam a fração da população que não demandou nenhum cuidado de saúde. A cauda longa e assimétrica à direita decorre de eventos de saúde associados a despesas altas. Esses eventos são relativamente raros,

entretanto, são normalmente responsáveis por uma grande proporção dos gastos. A acomodação dessa estrutura é fundamental para a obtenção de estimativas precisas.

Como ponto de partida dos principais métodos de estimação, a primeira característica relevante é o tratamento das observações com gasto zero. Quando os dados apresentam uma massa significativa de observações com gasto zero, duas alternativas são usualmente contempladas: o modelo de duas partes e o modelo de Heckmam (Deb, Manning, Norton, 2010). Em ambos os casos, modelamos separadamente as observações com gasto zero. Quando não há uma fração significativa de observações com gasto zero, é possível adotar diretamente métodos como modelos lineares generalizados (MLG) ou mínimos quadrados ordinários (Deb, Manning; Norton, 2010; Jones, 2010), assumindo que os gastos são uma função exponencial das variáveis explicativas (Manning; Mallahy, 2001).

Uma característica que chamou atenção nos nossos dados foi que apenas 6% das observações apresentaram gasto zero, e, longitudinalmente, apenas 1,5% dos indivíduos apresentaram algum período com gasto zero. Esse baixo percentual nos conduziu a opção de não modelar explicitamente os casos de gasto zero. Estratégia similar foi adotada em estudo recente por Shang e Goldman (2008).

Diante desse primeiro filtro, optamos pelos modelos lineares generalizados, que têm apresentado boa aplicabilidade na literatura de economia da saúde (Mullahy, 1998; Blough et al,1999; Manning; Mallahy, 2001; Buntin, 2004; Manning et al, 2005; Manning, 2006). Em comparação com os modelos tradicionais de regressão linear, que empregam a transformação logarítmica dos gastos como forma de tornar a variável mais simétrica, criando problemas de retransformação para a escala econômica original (Duan, 1983; Jones, 2010), nos modelos lineares generalizados, as predições de gasto podem ser feitas diretamente na escala real. Além disso, a heterocedasticidade do modelo é contemplada pela família distributiva, que especifica a relação entre a variância e a média condicional.

Nos MLG, modelamos diretamente as funções de média e variância na escala original de gastos. A função de média condicionada é especificada diretamente por:

$$E[y_i \mid x_i] = \mu_i = f(x_i^{'}\beta)$$

Onde y_i é o nível de custos do indivíduo *i*, x_i é o vetor de características individuais, e *f* a função link, usualmente função log em modelos de gasto, assumindo então a seguinte forma:

$$E[y_i \mid x_i] = f(x_i^{'}\beta) = \exp(x_i^{'}\beta)$$

O segundo componente do modelo é a distribuição usada para especificar a relação entre a variância e média condicional:

$$\mathrm{var}[y_i \mid x_i] = \mu^{v}$$

Onde, se $v = 0$, a variância é constante (família distributiva gaussiana); se $v = 1$, a variância é proporcional à média (Poisson); se $v = 2$, a variância é proporcional ao quadrado da média (*gamma*); se $v = 3$, a variância é proporcional ao cubo da média (inversa gaussiana). Outras distribuições usam a função quadrática da média: Bernoulli, $\mu(1-\mu)$, e binomial, $n\mu(1-\mu)$ (Jones, 2010).

Em aplicações empíricas, uma forma útil de escolher a distribuição, sugerida por Manning e Mullahy (2001), é o teste de Park para estimar a relação entre a média e a variância (Park,1966), que consiste em regredir o quadrado dos resíduos (ou do MLG, ou do

MQO com transformação log) sobre os valores preditos de *y* do mesmo modelo, ambos log transformados (Buntin, 2004):

$$\ln((y_i - \hat{y}_i)^2) = \lambda_0 + \lambda_1 \ln(\hat{y}_i) + \varphi_i$$

O coeficiente λ_1 indica que função de variância é apropriada aos dados.

2.3 Estratégia empírica

Como explicado acima, dispomos de um acompanhamento longitudinal de três anos, constituído por indivíduos sobreviventes por no mínimo 36 meses e por indivíduos não sobreviventes, para os quais calculamos retrospectivamente o tempo de vida restante. Uma das principais vantagens dos dados em painel é a possibilidade de observar a mesma unidade *cross-section* em diferentes pontos do tempo, o que permite tratar o efeito individual não observado.

O método dos MLG foi construído originalmente para aplicações de dados *cross-section* (Hilbe, 2011). Quando as informações estão em painel, entretanto, é possível incorporar o tratamento da heterogeneidade não observada via efeito populacional, de forma que, o efeito não observado é uma média da população. O método que incorpora o efeito populacional é denominado equações de estimação generalizada (EEG).

As EEG são um método proposto inicialmente por Liang e Zeger (1986) que avalia a relação entre a variável resposta e as variáveis preditoras num contexto populacional e cuja característica essencial da especificação do modelo é a fatoração da função de variância para incluir uma estrutura de correlação parametrizada no painel. Para acomodar as características da variável gasto, implementamos a função log-link e aplicamos o teste de Park para escolher a função distribuição que modela a relação entre a média e a variância. O coeficiente λ_1 estimado a partir do teste de Park foi de 1,25. Portanto, a função de variância apropriada está entre a distribuição de Poisson e a distribuição *gamma*, sendo mais próxima da distribuição de Poisson. Na distribuição de Poisson, a variância $V(y \mid x)$ é proporcional à média $E(y \mid x)$.

A regressão de Poisson com log link estimada é:

$$Log\,(E(y_{it} \mid X_{it})) = x_{it}^{'}\beta$$

A função de variância do modelo geral, estimada por GLM, é:

$$V(\lambda_{it}) = [D(V(\lambda_{it}))^{1/2} R_{n_i X n_i} D(V(\lambda_{it}))^{1/2}]_{n_i X n_i}$$

Onde $V(\lambda_{it})$ é a função de variância do GLM definida em termos da média. No binomial negativo, a função de variância é $\lambda + \alpha\lambda^2$. *D* é uma matriz diagonal com as funções de variância de y_i na diagonal, e *R* é a matriz de correlação de trabalho. Se assumirmos independência entre os painéis, ou seja, ausência de correlação entre as observações para um mesmo indivíduo i em diferentes períodos de tempo, temos que $R_{n\,Xn} = I_{n\,Xn}$; nesse caso, estrutura de correlação é independente.

A vantagem adicional da estimação por EEG é que a matriz de correlação da função de variância do GLM pode ser arbitrariamente parametrizada. Entre as possíveis estruturas de correlação, podemos ter: permutável, não estruturada e auto regressiva.

Na estrutura permutável, assume-se que a correlação entre as observações de um mesmo indivíduo ao longo do tempo é sempre a mesma. Na não estruturada, para cada observação do indivíduo ao longo do tempo, há uma correlação diferente, que é livremente

estimada a partir dos dados. Na estrutura de correlação autorregressiva (AR), assume-se que existe um decréscimo do valor do coeficiente de correlação à medida que aumentam os intervalos de tempo no painel.

Optamos por replicar o modelo de Poisson com log link contemplando duas estruturas de correlação: não estruturada e AR de primeira ordem. O quadro 1 descreve as variáveis de controle.

QUADRO 1 – Definição das variáveis explicativas

variável	**descrição**
	associadas a morte
ano da morte	*dummy* = 1: morreu naquele período
1 ano antes da morte	*dummy* = 1: está a um ano da morte (morreu no período seguinte)
2 anos antes da morte	*dummy* = 1: está a dois anos da morte
3 anos antes da morte	*dummy* = 1: está a três anos da morte
morte	*dummy* = 1: não sobrevivente
	demográficas
idade	idade em anos
idade*idade	idade em anos ao quadrado
sexo	*dummy* = 1: homem
	demográficas associadas a morte
morte*sexo	interação dummy se não sobrevivente e sexo
morte*Idade	interação dummy se não sobrevivente e idade em anos
	relativas ao plano
agregado	*dummy* = 1: agregado ao plano
dependente	*dummy* = 1: dependente do plano
básico	*dummy* = 1: possui plano básico
especial	*dummy* = 1: possui plano especial
tamanho da família	nº de pessoas da mesma família vinculadas à carteira Sabesprev
	tempo
ano 2005	*dummy* de ano 2005
ano 2006	*dummy* de ano 2006

fonte: Elaborado pelas autoras.

3. RESULTADOS

Nesta seção, apresentamos os resultados encontrados a partir do modelo EEG. Para visualizar o comportamento das variáveis demográficas a partir da inclusão das variáveis associadas a morte, estimamos quatro especificações separadamente.

O quadro 2 explicita as variáveis usadas em cada especificação/modelo. No modelo 1, especificamos apenas as variáveis de idade, idade ao quadrado e sexo; no modelo 2, o grupo de variáveis associadas a morte; no modelo 3, as variáveis associadas a morte e as demográficas e as interações; por fim, no modelo 4, a especificação completa.

QUADRO 2 – Variáveis usadas em cada modelo

variável	modelo 1	modelo 2	modelo 3	modelo 4
ano da morte		X	X	X
1 ano antes da morte		X	X	X
2 anos antes da morte		X	X	X
3 anos antes da morte		X	X	X
morte		X	X	X
idade	X		X	X
idade*idade	X		X	X
homem	X		X	X
morte*homem			X	X
morte*idade			X	X
agregado				X
dependente				X
básico				X
especial				X
tamanho da família				X
ano de 2005	X	X	X	X
ano de 2006	X	X	X	X

fonte: Elaborado pelas autoras.

Na tabela 2, apresentamos a semielasticidade e o efeito marginal do modelo estimado com estrutura de correlação autorregressiva de primeira ordem (AR1). Na tabela 3, a semielasticidade e o efeito marginal do modelo com estrutura de correlação permutável. De modo geral, os dois modelos apresentaram pouca diferença dos efeitos estimados. Todas as variáveis são significativas, independentemente da especificação. Para facilitar a discussão, nos reportaremos aos resultados do modelo estimado por AR1.

O modelo 1 evidenciou um efeito marginal de R$ 303 para cada ano de idade adicional, com uma tendência de queda dessa diferença ao longo da idade e homens apresentando um gasto 5% maior que o das mulheres. No modelo 2, em que só incluímos controles associados ao evento de morte, observamos que não sobreviventes ao painel gastaram 147% a mais que os sobreviventes. O tempo até a morte apresenta um efeito sempre crescente com a proximidade da morte, chegando a 170% no último ano de vida.

No modelo 3, incluímos as variáveis demográficas, as associadas a morte e as interações. Os resultados apontaram um efeito similar ao do modelo 2 nas *dummies* de tempo até a morte: à medida que se aproxima a morte, os gastos aumentam a taxas cada vez maiores, partindo de um aumento de 18% a três anos da morte e chegando a 181% no ano da morte. Já a dummy de morte teve seu efeito marginal quase duplicado, o que significa que, controlando para idade e sexo, não sobreviventes do painel apresentam um gasto marginal 250% superior aos sobreviventes. A este efeito deve ser adicionado o quão próximo o indivíduo se encontra da morte, no ano da morte, o gasto marginal chega a uma elevação de 430%. Esse conjunto de resultados reforça a necessidade da distinção entre sobreviventes e não sobreviventes para fins de projeção de gastos. Todas as variáveis associadas a morte apresentaram um efeito marginal elevado sobre os gastos, de forma que, em grupos de idade com taxas de mortalidade mais elevada, o gasto médio aumenta, sobretudo em função daqueles próximos da morte. Em outras palavras, caso as taxas de mortalidade específica se alterem, o gasto médio na faixa etária também se altera.

Com relação as variáveis demográficas, não observamos uma mudança significativa do coeficiente da idade. Vale reforçar que a incorporação das variáveis associadas a morte é

fundamental para a distinção de gastos entre sobreviventes e não sobreviventes, informação fundamental sobretudo para modelos de projeção de gastos, onde taxas específicas de mortalidade tendem a mudar ao longo do tempo. O fato do coeficiente se manter relativamente constante com a incorporação das informações sobre morte significa apenas que não houve mudança do efeito médio da idade com a incorporação dos novos controles. Por outro lado, a diferença entre homens e mulheres, que foi positiva no modelo 1, assume um valor negativo, evidenciando um gasto marginal 20% menor entre homens. A explicação para esta queda fica evidenciada quando analisamos as interações. A interação da dummy de morte e da dummy de sexo mostrou que homens não sobreviventes apresentaram um gasto maior que mulheres não sobreviventes (22%). Segue, portanto, que as diferenças observadas entre os modelos nos coeficientes associados ao sexo evidenciam um comportamento bastante distinto entre homens e mulheres sobreviventes e não sobreviventes, corroborando evidências que reportam gastos mais elevados entre homens não sobreviventes do que entre mulheres. Homens tendem a gastar mais quando têm problemas graves de saúde associados a mortalidade, o que se explica ou por fatores biológicos, ou por diferenças no estilo de vida e na condução de um cuidado preventivo ao longo da vida. Por último, observamos que a idade interagida com morte apresentou semielasticidade de -2%; como a estimativa da idade ainda é superior (10%), a interpretação que segue é que, entre não sobreviventes, a variação do gasto à medida que a idade aumenta é positiva (cerca de 8%) e a taxas decrescentes.

No modelo 4, de estimativa completa, a inclusão das variáveis associadas à condição do beneficiário no plano não alteraram significativamente os coeficientes já discutidos. Esse grupo de variáveis, entretanto, apresentou um efeito marginal relevante sobre os gastos. Destacaram-se as variáveis de plano: beneficiários do plano especial têm um gasto 19% maior que beneficiários do plano denominado pleno, categoria de referência no modelo, enquanto beneficiários do plano básico têm um gasto 9% menor que a categoria de referência. O plano básico se distingue dos demais quanto ao padrão de acomodação do cuidado hospitalar, que é de enfermaria e implica menos gastos para o agente segurador.

No gráfico 3, mostramos o gasto predito segundo a idade, separando sobreviventes e não sobreviventes e mantendo todas as médias observadas entre os dois grupos. Para facilitar a visualização, reportamos os gastos por regressão polinomial local ponderada, estimada por funções de Kernel. A área em sombreada em torno da linha corresponde ao intervalo de confiança de 95% da média. A linha superior representa o gasto esperado na população não sobrevivente do painel. A linha contínua, do meio, representa os gastos observados, e a linha de baixo, os gastos esperados entre sobreviventes.

O primeiro aspecto que chama atenção é a importância da incorporação dos indicadores de mortalidade: uma significativa parcela da elevação dos gastos observada com o aumento da idade se deve à proximidade da morte, corroborando a necessidade de distinção entre os dois grupos.

O gráfico também ilustra que, em relação aos indicadores de mortalidade, a idade teve um efeito bem menos proeminente, embora ainda significativo, diferindo das evidencias de *red herring* postuladas por Zweifel e outros (1999). Esse resultado corrobora evidencias reportadas na literatura internacional (Seshamani, Gray, 2004; Werblow et al, 2007).

No gráfico 4, apresentamos o gasto predito para sobreviventes, separando homens e mulheres: a linha superior mostra mulheres sobreviventes, e a linha inferior, homens sobreviventes. Observamos a diferença de nível entre os dois grupos e um gasto crescente com a idade, com tendência a cair nas idades muito avançadas.

TABELA 2 – Resultados da regressão de gasto por EEG com estrutura de correlação autorregressiva de primeira ordem

Função link log com família Poisson												
estrutura de correlação auto regressiva de primeira ordem												
variáveis/modelo	**modelo 1**			**modelo 2**			**modelo 3**			**modelo 4**		
	semielasticidade	**efeito marginal**		**semielasticidade**	**efeito marginal**		**semielasticidade**	**efeito marginal**		**semielasticidade**	**efeito marginal**	
ano da morte				1,75	5.867,99	***	1,81	6.075,26	***	1,81	6.070,50	***
1 ano antes da morte				1,29	4.330,30	***	1,34	4.489,94	***	1,34	4.490,60	***
2 anos antes da morte				0,61	2.055,63	***	0,65	2.185,02	***	0,65	2.194,77	***
3 anos antes da morte				0,15	518,20	***	0,18	6.08,304	***	0,18	613,24	***
morte				1,47	4.945,42	***	2,39	8.014,6		2,39	8.018,13	
idade	0,09	303,79	***				0,10	362,53	***	0,10	352,96	***
idade*idade	0,00	-1,20	***				0,00	-2,23	***	-0,00	-2,24	***
homem	0,05	161,94	***				-0,20	-671,58	***	-0,22	-753,37	***
morte*homem							0,42	1.399,83	***	0,40	1.368,95	***
morte*idade							-0,02	-74,17	***	-0,02	-73,68	***
agregado										0,03	116,11	***
dependente										-0,04	-161,58	***
básico										-0,09	-320,37	***
especial										0,19	659,44	***
tamanho da família										0,00	10,24	***
ano de 2005	0,13	428,65	***	0,05	158,56	***	0,02	52,56	***	0,01	56,73	***
ano de 2006	0,19	640,41	***	0,11	384,15	***	0,06	187,25	***	0,01	198,58	***
nº de observações					63.874							
nº de indivíduos					21.348							
categorias de referência: 2004, mulher, titular do plano de saúde e plano pleno												
*** significativo a 1%												

fonte: Elaborada pelas autoras a partir de dados da Sabesprev.

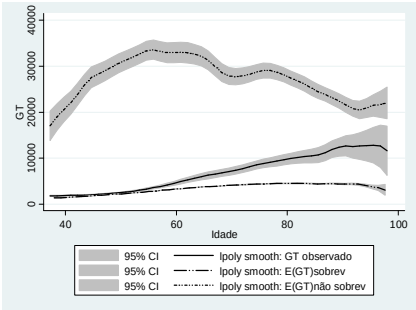
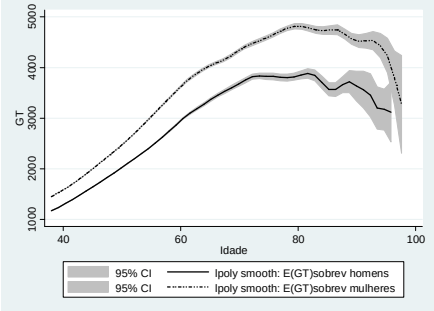
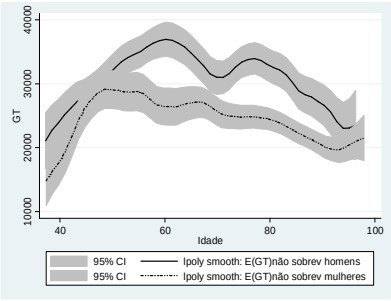
TABELA 3 – Resultados da regressão de gasto por EEG com estrutura de correlação permutável

Função link log com família Poisson

estrutura de correlação auto regressiva de primeira ordem

variáveis/modelo	modelo 1			modelo 2			modelo 3		modelo 4		
	semielasticidade	efeito marginal		semielasticidade	efeito marginal		semielasticidade	efeito marginal	semielasticidade	efeito marginal	
ano da morte				1,62	5.624,71	***	1,68	5.824,39	1,68	5.831,51	***
1 ano antes da morte				1,22	4.225,41	***	1,27	4.417,80	1,28	4.426,49	***
2 anos antes da morte				0,54	1.887,99	***	0,59	2.047,08	0,60	2.067,53	***
3 anos antes da morte				0,09	297,74	***	0,12	403,97	0,12	410,66	***
morte				1,53	5.312,67	***	2,39	8.271,96	2,36	8.183,18	***
idade	0,10	371,78	***				0,12	422,64	0,12	410,48	***
idade*idade	-0,00	-1,59	***				0,00	-2,71	0,00	-2,70	***
homem							-0,20	-691,36	-0,25	-856,44	***
morte*homem							0,35	1.218,21	0,35	1.222,47	***
morte*idade							-0,02	-71,62	-0,02	-70,20	***
agregado									0,06	216,96	***
dependente									-0,09	-304,06	***
básico									-0,07	-232,70	***
especial									0,18	622,21	***
tamanho da família									0,00	-10,21	***
ano de 2005	0,04	161,04	***	0,07	241,99	***	0,05	156,34	0,05	156,88	***
ano de 2006	0,10	368,08	***	0,14	476,36	***	0,09	308,25	0,09	312,87	***
nº de observações					64.042						
nº de indivíduos					21.537						
categorias de referência: 2004, mulher, titular do plano de saúde e plano pleno											
*** significativo a 1%											

fonte: Elaborada pelas autoras a partir de dados da Sabesprev.

GRÁFICO 3 – Gasto segundo idade, por *status* de sobrevivência	GRÁFICO 4 – Gasto segundo idade, por sexo – sobreviventes	GRÁFICO 5 – Gasto segundo idade, por sexo – não sobreviventes
		

fonte: Elaborado pelas autoras a partir de dados da Sabesprev.

4. CONSIDERAÇÕES FINAIS

Neste trabalho, analisamos os determinantes dos gastos individuais com saúde a partir de uma carteira de beneficiários de um plano de saúde do estado de São Paulo. A população avaliada é constituída de indivíduos com mais de 40 anos de idade, cujos gastos com saúde acompanhamos longitudinalmente durante 6 anos, o que nos permitiu observar os eventos de morte nessa população. Retrospectivamente, portanto, pudemos calcular o tempo de vida restante para a fração da população não sobrevivente. A incorporação do tempo até a morte como preditor de gastos tem sido amplamente discutida na literatura da última década (Seshamani, Gray, 2004; Zweifel et al, 2004; Werblow 2007). A consideração desse dado é fundamental para projeções de gastos com saúde em função das mudanças demográficas da população, visto que as taxas de mortalidade específicas por idade têm se alterado ao longo do tempo, alterando, consequentemente, perfis de gasto por idade.

Na análise econométrica, restringimos o painel aos três anos iniciais, de forma que observamos sobreviventes por no mínimo 36 meses e não sobreviventes. Os resultados mostraram que a proximidade da morte tem um efeito importante sobre os gastos. O aumento de gasto chega a um efeito marginal de até 430% no ano da morte. Nas variáveis demográficas, chamou atenção a importante diferença de gasto entre homens e mulheres, que fica evidenciado quando incorporamos o grupo de variáveis associadas a morte. Em média, homens têm um gasto inferior ao das mulheres. Homens não sobreviventes, entretanto, apresentam um gasto maior. Embora, de forma descritiva, esse seja um resultado presente na literatura (Owens, 2008), as principais abordagens econométricas que investigaram o comportamento dos gastos a partir da inclusão de variáveis associadas a morte não apresentaram esse resultado. Para discutir a diferença entre os gêneros, é fundamental que se façam estudos futuros capazes de representar a realidade brasileira e de medir a causa de morte.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, M. V., MAIA, A. C, RODRIGUES, C. G. Gastos com serviços médicos por status de sobrevivência no setor de saúde suplementar no Brasil. **Revista Brasileira de Estudos da População**. No prelo, 2011.

BARROS, P. P. The Black Box of Health Care Expenditure Growth Determinants. **Health Economics**. v.7(6), p.533-544, 1998.

BLOUGH, D. K., MADDEN, C. W., HORNBROOK, M. C. Modeling Risk Using Generalized Linear Models. **Journal of Health Economics**. v.18, p. 153-171, 1999.

BRITO, F. **A transição demográfica no Brasil**: as possibilidades e os desafios para a economia e a sociedade. Belo Horizonte: Cedeplar, 2007. (Texto para Discussão, 318)

BUNTIN, M. B., ZASLAVSKY, A. M. Too Much Ado about Two-part Models and Transformation? Comparing Methods of Modeling Medicare Expenditures. **Journal of Health Economics**. v.23, p. 525-542. 2004.

CARVALHO, J. A. M. **Crescimento populacional e estrutura demográfica no Brasil**. Belo Horizonte: CEDEPLAR, 2004. (Texto para discussão, 227)

CARVALHO, J. A. M., WONG, L. L. R. A transição da estrutura etária da população brasileira na primeira metade do século XXI. **Cadernos de Saúde Pública**, Rio de Janeiro, v.24, n.3, p.587-605, mar. 2008.

DEB, P., MANNING, W., NORTON, E. **Preconference Course**: Modeling Health Care Costs and Counts. ASHE – Cornell University Conference, 2010. Disponível em: <http://www.ispor.org/OpenSourceIndex/cached/Modeling_Health_Care_ Costs_and_Use_Files/ASHE2010_Minicourse_Cost_Use_slides.pdf>. Acesso em: 27 nov. 2011.

DORMONT, B., MARTINS, J. O., PELGRIN, F., SUHRCKE, M. **Health Expenditures, Longevity and Growth**. IX European Conference of the Fondazione Rodolfo Debenedetti on "Health, Longevity and Productivity" Limone sul Garda, 26 May 2007.

DUAN, N. Smearing Estimate: a Nonparametric Retransformation Method. **Journal of the American Statistical Association**. n.78, p.605-610, 1983.

GOTTRET, P.; SCHIEBER, G. **Health Financing Revisited**. Washington DC: The International Bank for Reconstruction and Development/The World Bank, 2006. 310p.

GERDTHAM, Ulf-G, JONSSON, B. International Comparisons of Health Expenditure: Theory, Data and Econometric Analysis. In: CULYER, A. J., NEWHOUSE, J. P. (Eds.). **Handbook of Health Economics**. New York: Elsevier. v.1A, p. 12-49, 2000.

GERDTHAM, Ulf-G, SOGAARD, J., ANDERSSON, F., JONSSON,B. An econometric analysis of health care expenditure: A cross-section study of the OECD countries. **Jornal of Health Economics**. v.11, p.63-84, 1993.

GETZEN, T. E. Population Aging and the Growth of Health Expenditures. **The Journal of Gerontology**. v.47, n.3, p.S98-S104, 1992.

HILBE, J. M. **Negative Binomial Regression**. 2nd ed. New York: Cambridge University Press, 2011.

HITIRIS, T., POSNETT, J. The Determinants and Effects of Health Expenditure in Developed Countries. **Journal of Health Economics**. v.11, n.2, p.173-181, Aug. 1992.

JONES, A. M. Models For Health Care. **Health, Econometrics and Data Group Working Paper**, 2010.

JONES, A. M. Health Econometrics. In: CULYER, A. J., NEWHOUSE, J. P. (Eds.). **Handbook of Health Economics**. New York: Elsevier. v.1A, p.265-336, 2000.

KOOPMANS, G. T., LAMERS, L. M. Gender and Health Care Utilization: The Role of Mental Distress and Help-seeking Propensity. **Social Science & Medicine**. Oxford. v.64, n.6, p.1.216-1.230, Mar. 2007.

LAYTE, R. An Analysis of the Impacto of Age and Proximity of Death on Health Care Costs in Ireland. **ESRI Working Paper**. n.193, 2007.

LIANG, K.-Y., ZEGER, S. Longitudinal Data Analysis Using Generalize Linear Models. **Biometrika**. v.73, p.13-22, 1986.

MANNING, W. G., MULLAHY, J. Estimating Log Models: To Transform or Not to Transform?. **Journal of Health Economics**. n.20, p.461-94, 2001.

MANNING, W. Dealing with Skewed Data on Costs and Expenditure. In: JONES, A. M. (Ed.). **The Elgar Companion to Health Economics**. Cheltenham: Edward Elgar, 2006. p.439-446.

MANNING, W. G., A. BASU, MULLAHY, J. Generalized Modeling Approaches to Risk Adjustment of Skewed Outcomes Dat. **Journal of Health Economics**. n.24, p.465-88, 2005.

MENDES, E. V. **Uma agenda para a saúde**. 2. ed. São Paulo: Hucitec, 2006. 300p.

MULLAHY, J. Econometric Modeling of Health Care Costs and Expenditures. A Survey of Analytical Issues and Related Policy Considerations. **Medical Care**. n.47, p.S104-S108, 2009.

MULLAHY, J. Much Ado about Two: Reconsidering Retransformation and the Two Part Model in Health Econometrics. **Journal of Health Economics**. n.17, p.247-81, 1998.

OMRAN, A. R. The Epidemiologic Transition: A Theory of the Epidemiologic of Population Change. **The Milbank Quarterly**. New York. v.38, n.4, p.731-57, 2005.

OWENS, G. M. Gender Differences in Health Care Expenditures, Resource Utilization, and Quality of Care. **Supplement to Journal of Managed Care Pharmacy**. v.14, n.3, 2008.

PARK, R. E. Estimation with Heteroscedastic Error. **Econometrica**. n.34, p.888, 1966.

RAITANO, M. The Impact of Death-related Costs on Health Care Expenditure: A Survey. **ENEPRI Research Report**. n.17, 2006.

SCHRAMM, J. A. et al. Transição epidemiológica e o estudo de carga de doença no Brasil. **Cien Saude Colet**. v.9, n.4, p.897-908, 2004.

SESHAMANI, M., GRAY, A. A Longitudinal Study of the Effects of Age and Time to Death on Hospital Costs. **Journal of Health Economics**. v.23, p.217-235, 2004.

SESHAMANI, M., GRAY, A. Time to Death and Health Expenditure: An Improved Model for the Impact of Demographic Change on Health Care Costs. **Age and Ageing**. London. v.33, n.6, p.556–561, Nov. 2004.

SHANG, B., GOLDMAN, D. Does Age or Life Expectancy Better Predict Health Care Expenditures? **Health Economics**. v.17, p.487-501, 2008.

STEARNS, S. C. NORTON, E. C. Time to Include Time to Death? The Future of Health Care Expenditure Predictions. **Health Economics**. v.13, p.315-327, 2004

STEIMNMANN, L., TELSER, H., ZWEIFEL, P. S. Aging an Future Healthcare Expenditure: A Consistent Approach. **Forum for Health Economics & Policy**. v.10, n. 2. Published by The Berkeley Electronic Press, 2007.

VAN DE VEN, W. P. M. M, ELLIS, R. P. Risk Adjustent in Competitive Health Plan Markets. In: CULYER, A. J., NEWHOUSE, J. P. (Eds.). **Handbook of Health Economics**. New York: Elsevier. v.1A, p.755-763, 2000.

VEGDA, K. et al. Trends in Health Services Utilization, Medication Use, and Health Conditions among Older Adults: A 2-year Retrospective Chart Review in a Primary Care Practice. **BMC Health Services Research**, Ann Arbor. v.9, n.217, p.1-7, Nov. 2009.

WERBLOW, A., FELDER, S., ZWEIFEL, P. Population Ageing and Health Care Expenditure: A School of "Red Herrings"? **Health Economics**. v.16, p.1.109-1.126, 2007

YANG, Z., NORTON, E. C., STEARNS, S. C. Longevity and Health Care Expenditures: The Real Reasons Older People Spend More. **The Journals of Gerontology**. v.58B, n.1, p.S2-S10, 2003.

ZWEIFEL, P., FELDER, S., WERBLOW, A. Population Ageing and Health Care Expenditure: New Evidence on the "Red Herring". **The Geneva Papers on Risk and Insurance**. v.29, n.4, p.652-666, 2004.

ZWEIFEL, P., FELDER, S., WERBLOW, A. Ageing of Population and Health Care Expenditure: A Red Herring?. **Health Economics**. v.8, p.485–496, 1999.