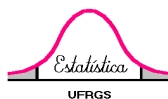




UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
DEPARTAMENTO DE ESTATÍSTICA



Aplicação do modelo de riscos competitivos em pacientes diagnosticados com câncer no ano de 2006 no Hospital de Clínicas de Porto Alegre

Autora: Natalia Elis Giordani
Orientadora: Professora PhD Suzi Alves Camey
Coorientadora: Professora PhD Luciana Neves Nunes

Porto Alegre, 16 de Janeiro de 2013.

Universidade Federal do Rio Grande do Sul
Instituto de Matemática
Departamento de Estatística

Aplicação do modelo de riscos competitivos em
pacientes diagnosticados com câncer no ano de
2006 no Hospital de Clínicas de Porto Alegre

Autora: Natalia Elis Giordani

Monografia apresentada para obtenção
do grau de Bacharel em Estatística.

Banca Examinadora:
Professora PhD Suzi Alves Camey (orientadora)
Professora PhD Luciana Neves Nunes (coorientadora)
Professor Dr. Álvaro Vigo

Porto Alegre, 16 de Janeiro de 2013.

Dedico este trabalho aos meus pais, Volmir e Marta, meus grandes exemplos, às minhas irmãs Geovanna e Samantha e ao meu namorado Lucas, meus maiores incentivadores.

Agradecimentos

O caminho percorrido foi longo e certamente sem o apoio de algumas pessoas teria sido muito mais difícil.

Agradeço primeiramente à minha família por ter sempre me incentivado e apoiado em todas as decisões, entendendo minha ausência e nunca me deixando desistir. Obrigada pai e mãe por aceitarem minhas escolhas e por terem sempre me amparado. Sem dúvida alguma vocês são essenciais na minha vida e sou muito grata por tudo. Às minhas irmãs Geo e Sami, obrigada pelas ligações, e-mails e fotos que me enviaram tantas vezes fazendo diminuir minha saudade de irmã do meio, nada acostumada a viver sem vocês duas. À minha avó, Diles, por ser presença constante em minha vida e por me querer tão bem. E aos meus sobrinhos lindos, motivo de tanta alegria, Ana Carolina e Augusto Henrique.

Ao meu namorado Lucas, por ter entendido e ter sido tão paciente nas inúmeras vezes que precisei abrir mão de sua companhia para estudar ou me dedicar a esse trabalho. E, por, além disso, ter se mostrado um grande companheiro e meu melhor amigo, sempre dando um jeitinho de me ajudar quando eu necessitava.

Às colegas que viraram muito mais do que simplesmente colegas: Pricila, Tássia e Sabrina. Foi muito bom descobrir em vocês pessoas maravilhosas com quem sei que posso contar. Dividimos momentos fantásticos, que certamente fizeram nossos dias mais leves e alegres, impulsionando a vontade de continuar.

Aos amigos que conheci no NAE, Bárbara e Tiago, vocês também não poderiam deixar de estar nos meus agradecimentos. Obrigada por terem tornado minhas tardes, durante um ano e meio, mais divertidas. Bárbara, obrigada pelas conversas, desabafos, idas e vindas no Agronomia. Obrigada Tiago, por ter me acompanhando tantas vezes até em casa depois das aulas de noite no campus do Centro!

As minhas orientadoras Suzi e Lu, pela atenção que me deram ao longo do desenvolvimento desse trabalho. Foi muito bom ter trabalhado com vocês, professoras que eu já admirava e que agora, sem dúvidas, tenho ainda mais respeito.

Por fim, agradeço a todos que de uma forma ou outra acompanharam minha trajetória e torceram para que mais essa etapa fosse concluída. Certamente todos os momentos, bons ou ruins, fáceis ou difíceis, me tornaram uma pessoa mais forte e madura, porém com os mesmos valores do dia em que saí da casa dos meus pais e com a certeza de que sozinha tudo isso não teria sentido.

Resumo

Na oncologia é comum se utilizar a análise de sobrevida para avaliação das taxas de letalidade. Porém, as abordagens mais tradicionais – Kaplan-Meier e Cox - consideram apenas um desfecho, ignorando eventos ditos competitivos. Desprezar tal característica resulta na superestimação da função de distribuição acumulada e os efeitos ou associações com as covariáveis podem ser estimados indevidamente. Como alternativa, propõe-se o modelo de riscos competitivos. Este trabalho tem como objetivo descrever esse modelo através de sua aplicação na sobrevida de pacientes a partir do diagnóstico de câncer atendidos pelo Hospital de Clínicas de Porto Alegre. Para tanto foi utilizada uma coorte de 1.364 pacientes diagnosticados com câncer em 2006, acompanhados até 2011, sendo que destes pacientes 600 foram a óbito. Foram comparados os resultados da abordagem tradicional e do modelo de riscos competitivos para três situações distintas. Os resultados mostram que desprezar a natureza competitiva dos eventos resulta em estimativas superestimadas da função de distribuição acumulada.

Câncer; Sobrevida; Modelo de riscos competitivos

Abstract

In oncology is common to use survival analysis to evaluate lethality rates. However, the more traditional approaches - Kaplan-Meier and Cox - consider just one outcome, they ignore competing events. When we neglect this characteristic the results can show overestimation of the cumulative incidence and wrong estimation for the covariates' effects. As alternative, we propose the competing risks model. This work aims to describe this technique through application on patients with the diagnosis of cancer attended at Hospital de Clínicas de Porto Alegre. We studied a cohort of 1,364 patients diagnosed with cancer in 2006 and followed until 2011, among them occurred 600 deaths. We compared the results of the traditional approach and the competing risks model for three different situations. The results show that when we neglect the competitive nature of the events, we obtain cumulative distribution overestimated.

Cancer; Survival; Competing risks model

Sumário

1. Introdução	9
2. Material e métodos	11
2.1. <i>Modelo de Riscos Competitivos</i>	11
2.2. <i>Delineamento experimental e população em estudo</i>	13
2.3. <i>Identificação dos órbitos</i>	13
2.4. <i>Aplicando Modelo de Riscos Competitivos no software R</i>	15
2.4.1 <i>Organização dos dados</i>	15
2.4.2 <i>Códigos R</i>	15
3. Resultados	18
4. Discussão	21
Referências bibliográficas	23

Este artigo será submetido à Revista Cadernos de Saúde Pública
Autores: Natalia Elis Giordani, Luciana Neves Nunes, Suzi Alves Camey, Isaias
Prestes e Jair Ferreira.

1. Introdução

O câncer tornou-se uma doença comum em todo o mundo. Em países desenvolvidos, em média, uma a cada cinco pessoas tem como causa do óbito algum tipo dessa doença ^[1]. De acordo com dados da Organização Mundial da Saúde (OMS), 27% da população mundial com idade inferior a 70 anos tem como causa de morte o câncer ^[2]. No Brasil, entre 2006 e 2010, a mortalidade pela doença representou 15% das causas de falecimento ^[3].

O câncer de pulmão constitui a neoplasia mais incidente no mundo, sendo também a causa de morte por câncer mais frequente ^[4]. O segundo tipo de maior incidência mundial é o câncer de mama que corresponde a 22% dos novos casos diagnosticados a cada ano, sendo mais comum entre as mulheres ^[5]. No Brasil, esse cenário não é diferente ^[3,6,7]: o câncer de pulmão é o tumor maligno com maior índice de mortalidade entre os homens e o segundo maior entre as mulheres – estima-se cerca de 17 mil novos casos durante este ano –; seguido pelo câncer de mama, que é a causa de morte mais frequente entre as mulheres em todo o país, representando cerca de 11 mil óbitos por ano – espera-se, para 2012, cerca de 52 mil novos casos.

Uma publicação recente relata que no Brasil as taxas de mortalidade pelos cânceres de pulmão e colorretal vêm aumentando tanto entre os homens quanto entre as mulheres. Especificamente entre os homens, a taxa de mortalidade por câncer de próstata também tem aumentado. Entre as mulheres nota-se o mesmo comportamento para o câncer de mama. Já as taxas de mortalidade por câncer gástrico diminuíram em ambos os sexos ^[8].

O Rio Grande do Sul é um dos estados que apresenta maior taxa bruta de mortalidade do país, considerando todas as neoplasias entre os anos de 2006 e 2010. Em relação aos homens gaúchos, o câncer de pulmão é o mais comum como causa de óbito. Já entre as mulheres, o câncer de mama constitui a causa de morte mais frequente ^[4].

Na área oncológica, a sobrevida é o parâmetro mais utilizado para avaliação das taxas de letalidade ^[9]. A taxa de sobrevida – ou seja, a porcentagem de pacientes que se espera que sobreviva ao câncer – varia de acordo com o tipo e estágio da doença. Porém, segundo informações disponibilizadas pelo Instituto Nacional de Câncer (INCA), a sobrevida relativa para todos os tipos de câncer é de 50% em cinco anos ^[10]. Especificamente para o câncer de pulmão, a sobrevida média em cinco anos varia entre 13% e 21% em países desenvolvidos e entre 7% e 10% nos países em desenvolvimento. Em relação ao câncer de mama, a sobrevida média esperada após cinco anos é de 85% nos países desenvolvidos e de 60% nos países em

desenvolvimento ^[6]. Cabe salientar que a diferença entre tais números (sobrevida do câncer de pulmão e mama) deve-se ao fato de que o câncer de mama, se diagnosticado e tratado da forma correta, apresenta grandes chances de cura enquanto que o câncer de pulmão normalmente é detectado em estágios mais avançados, tornando-se uma doença altamente letal ^[5,11].

Da revisão bibliográfica sabe-se que a forma mais usual de se avaliar a sobrevida de pacientes com câncer é através da utilização de modelos de sobrevida tradicionais, tais como modelo de Cox e método de Kaplan-Meier. Para exemplificar, uma busca das palavras chave *cancer* e *survival analysis* no PUBMED retornou 2.501 artigos publicados na última década ^[12]. Tais modelos permitem considerar como variável resposta apenas um tipo de desfecho, normalmente o óbito por um tipo específico da doença. Essa condição torna esse tipo de abordagem bastante limitada, uma vez que um paciente com câncer pode falecer por causas diferentes do câncer.

Dessa forma, torna-se mais interessante avaliar não apenas o óbito, mas também a sua causa. Isso é possível através da aplicação de modelos de riscos competitivos, cujo objetivo é analisar diferentes eventos como desfecho de uma mesma situação de risco ^[13]. Dessa forma, este trabalho visa à aplicação de tal método para análise de sobrevida em pacientes com câncer diagnosticado no Hospital de Clínicas de Porto Alegre (HCPA) em 2006 com acompanhamento até 2011.

Assim, o objetivo principal deste trabalho é estudar a sobrevida de pacientes a partir do diagnóstico de câncer atendidos pelo HCPA através da aplicação do modelo de riscos competitivos. Além disso, como objetivo específico, pretende-se descrever tal técnica.

2. Material e métodos

2.1. Modelo de Riscos Competitivos

A análise de sobrevida é a técnica estatística utilizada em situações cujo interesse consiste em modelar o tempo até a ocorrência de determinado desfecho na presença ou não de censura. De forma geral, dados são classificados como censurados se ao término do período de observação não houve a ocorrência do desfecho de interesse ou em caso de perda de acompanhamento de um paciente.

Nesse tipo de análise, pode-se expressar a variável resposta através da função de sobrevida – probabilidade de não ocorrência do desfecho de interesse dentro do intervalo de tempo observado; ou da função de risco – risco de que o evento de interesse ocorra em um intervalo de tempo dado que o paciente sobreviveu até aquele instante; ou da função de risco acumulado – risco de um evento ocorrer até o tempo t ; ou ainda da função distribuição acumulada - probabilidade de que o desfecho de interesse ocorra em um intervalo instantâneo de tempo. Essas funções, chamadas de funções básicas de análise de sobrevida, são relacionadas entre si. Assim, ao se estimar a função de sobrevida obtêm-se também as funções risco e risco acumulado ^[14].

Existem diversas formas de se estimar a função de sobrevida, sendo as mais tradicionais o método de Kaplan-Meier – estimação não paramétrica – e o modelo de Cox – modelo semiparamétrico. Ambos permitem considerar apenas um desfecho de interesse, como por exemplo, o óbito por câncer. Tal condição ignora a possibilidade de ocorrência de outros desfechos, que no mesmo exemplo seria considerar que o paciente pode falecer por causas não relacionadas ao câncer.

Nessa situação, há o interesse de se modelar o tempo de sobrevida de eventos ditos competitivos, que se caracterizam por serem decorrentes de um mesmo fator de risco, onde a ocorrência de um dos eventos exclui a possibilidade de que os demais aconteçam. Assim, para este estudo, os eventos são classificados como competitivos uma vez que o interesse consiste em modelar a sobrevida de pacientes expostos ao fator de risco câncer primário, sujeitos a óbito por diferentes causas, que podem, ou não, estar relacionados à doença. A diferença entre os modelos tradicionais de sobrevida e o modelo de riscos competitivos, no contexto deste trabalho, está representada na Figura 1.

Eventos classificados como competitivos podem ser analisados de três formas ^[14]:

- i. Sobrevivência até o primeiro dos desfechos (event-free-survival): utilizada quando se estuda vários possíveis efeitos de um fator de risco. Considera-se o tempo até a ocorrência do primeiro evento, independente de qual dos eventos competitivos tenha ocorrido, sendo que o interesse consiste em estimar o efeito médio da variável independente sobre qualquer um dos eventos (esta abordagem não foi considerada nesse artigo).
- ii. Risco específico por causa: utilizado quando se deseja estimar o efeito de uma variável explicativa sobre um desfecho específico, considerando os demais possíveis desfechos como censura não informativa (mesmo que não o sejam).
- iii. Modelagem da subdistribuição do risco: baseada na subdistribuição (ou funções de incidência acumulada) dos riscos, esse método pondera e leva em conta não apenas o desfecho de interesse, mas a totalidade dos desfechos. Desse modo, ao sofrer um dos eventos competitivos o indivíduo não é censurado, mas recebe um peso com decaimento decrescente.

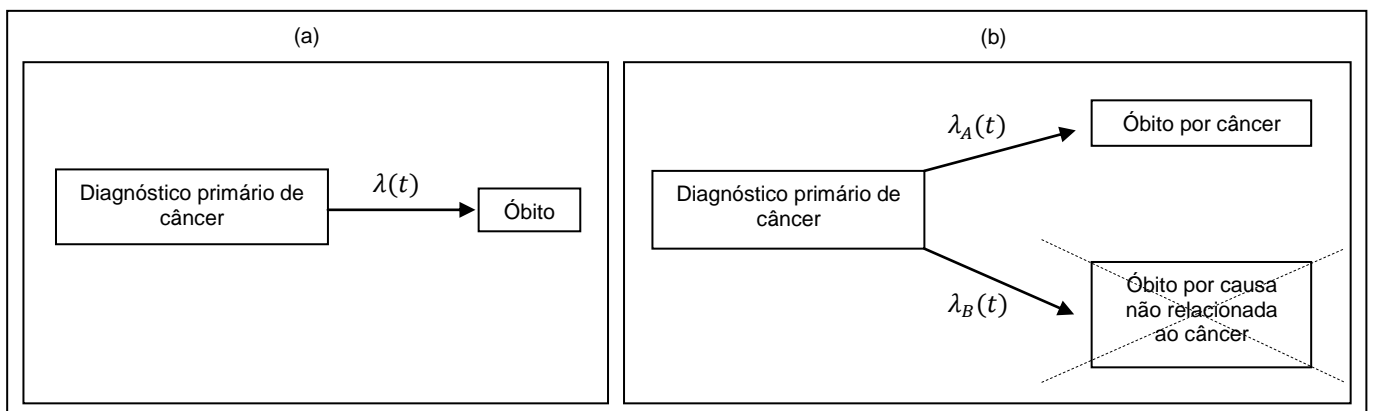


Figura 1: Representação do modelo tradicional de análise de sobrevivência e do modelo de riscos competitivos. Na representação (a), modelo tradicional de sobrevivência, $\lambda(t)$ refere-se ao risco do paciente sofrer o desfecho óbito, desconsiderando sua causa. Na representação (b), modelo de riscos competitivos, $\lambda_A(t)$ refere-se ao risco de o paciente vir a sofrer o desfecho A (óbito por câncer) e $\lambda_B(t)$ é o risco de o paciente vir a sofrer o desfecho B (óbito por causa não relacionada ao câncer).

Utilizar uma das duas primeiras abordagens resultará nos seguintes problemas: (i) haverá uma superestimação da função distribuição acumulada e (ii) os efeitos das covariáveis podem não ser estimados adequadamente. Tais problemas são decorrentes da principal característica dos riscos competitivos: a correspondência direta entre as funções risco e risco acumulado é perdida ^[15]. Isso significa que, ao contrário do modelo tradicional de análise de sobrevivência, o risco acumulado para um desfecho (neste caso, o óbito por câncer) depende do

risco acumulado por causa específica do desfecho competitivo (aqui, o óbito por causa não relacionada ao câncer).

Desta forma, a terceira abordagem com o intuito de corrigir tais problemas propõe a utilização das funções de incidência acumuladas (FIA) ou subdistribuições, que consideram além do desfecho de interesse os demais possíveis desfechos e não supõe, ao contrário das outras abordagens, que os eventos competitivos sejam independentes.

De maneira geral, o modelo de riscos competitivos pode ser expresso por ^[14]:

$$\lambda_k(t|\mathbf{x}) = \lambda_{0k}(t) \exp(\mathbf{x}\boldsymbol{\beta}), \quad (1)$$

onde λ_k diz respeito a função risco da subdistribuição para o evento k , λ_{0k} a função de risco basal do evento k , \mathbf{x} ao vetor de variáveis explicativas e $\boldsymbol{\beta}$ ao vetor de coeficientes.

2.2. Delineamento experimental e população em estudo

Este trabalho é baseado na coorte de casos primários de cânceres diagnosticados no HCPA no ano de 2006. No total, são 1.364 casos, acompanhados até 31/12/2011, a fim de garantir, no mínimo, cinco anos de acompanhamento para todos os sujeitos. A base de dados do HCPA, que é encaminhada ao Instituto Nacional de Câncer (INCA), foi criada em 2009 e contém as seguintes informações para cada indivíduo: sexo, data de nascimento, data da primeira consulta, data do diagnóstico, localização primária do tumor e, no caso de morte, data do óbito, causa básica e causa imediata de morte.

As causas de morte, assim como a localização primária do tumor, estão codificadas de acordo com a Classificação Internacional de Doenças e Problemas Relacionados à Saúde, conhecida como CID-10 ^[16]. Assim, para cada paciente, foi atribuída uma categoria única que corresponde a um código do CID-10. Este estudo teve aprovação do comitê de pesquisa e ética em saúde do Hospital de Clínicas de Porto Alegre.

2.3. Identificação dos óbitos

A identificação dos óbitos foi feita através de duas fontes: base de dados do HCPA e Sistema de Informações de Mortalidade (SIM) ^[17]. Os óbitos constantes na base de dados do HCPA foram checados com as informações do SIM, que também foi utilizado para detectar óbitos ocorridos após o registro do caso na base de dados do HCPA. Nesse período de preparação e limpeza do banco de dados alguns problemas foram identificados, sendo eles:

- i. Diferenças ortográficas entre o nome que constava na base de dados do HCPA com o nome localizado no SIM, claramente erros de digitação;

ii. Inversão entre dia e mês nas datas de nascimento e óbito.

A checagem de dados foi feita através da comparação entre o nome e data de nascimento das duas fontes (HCPA e SIM). Alguns procedimentos foram adotados, após a detecção de tais problemas, a fim de padronizar a fonte das informações:

- Para data de nascimento optou-se pela utilização de dados do HCPA, pois, para os casos de não óbito, não teria dupla fonte de informação para checagem de datas;

- Para data de óbito decidiu-se utilizar as informações do SIM. Dos 600 pacientes que vieram a óbito apenas onze não foram localizados na base de dados do SIM. Assim, apenas para esses casos utilizou-se a data de óbito disponibilizada pelo HCPA;

- Informações de óbito de casos que não haviam sido registrados no hospital foram incluídas no banco. No total, foram identificados 263 novos casos;

- A base inicial disponibilizada pelo HCPA possuía quatro pacientes duplicados que foram excluídos das análises.

O detalhamento da população em estudo encontra-se na Figura 2.

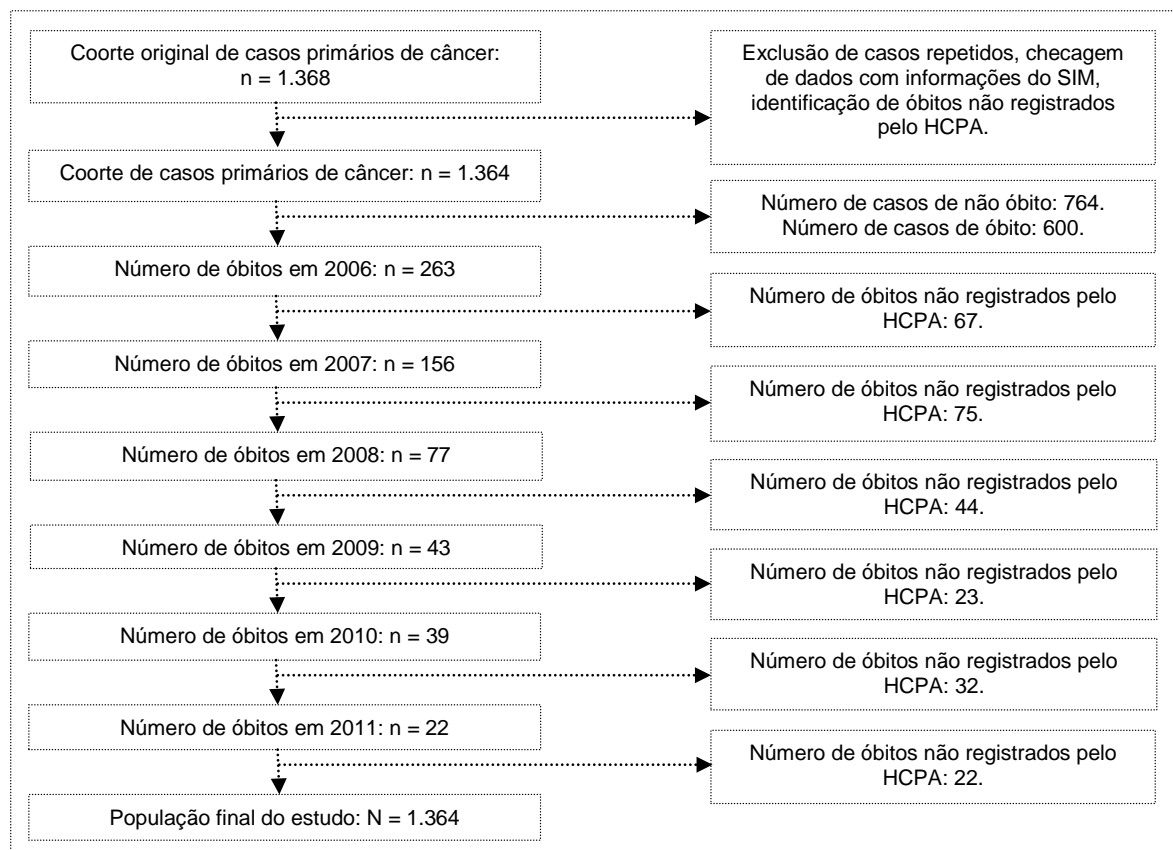


Figura 2: Detalhamento da população em estudo

2.4. Aplicando Modelo de Riscos Competitivos no software R

Esta seção tem por objetivo explicar o processo da modelagem de riscos competitivos utilizando o software R^[18].

2.4.1 Organização dos dados

Para utilização da abordagem de riscos competitivos, o banco de dados deve ser estruturado de acordo com a Figura 3. Na aplicação deste artigo o tempo de acompanhamento de cada paciente foi calculado a partir da data de diagnóstico de câncer, terminando no momento em que um dos eventos competitivos ocorria. O banco de dados contém duas linhas para cada paciente (pois estamos considerando apenas dois eventos competitivos), sendo que cada uma deve conter um valor diferente para a variável **ENPOINT**, que codifica os possíveis eventos: 1 = óbito por câncer, 2 = óbito por causa não relacionada ao câncer. Na linha em que a variável **ENPOINT** possui o evento que realmente ocorreu a variável **STATUS** deve receber o valor 1 e, na outra linha, o valor 0.

Por exemplo, o paciente 1 é do sexo masculino, sua idade ao diagnóstico de câncer era de 59 anos, faleceu por causa não relacionada a doença e teve apenas 2 meses de acompanhamento a partir da data de diagnóstico.

ID	SEXO	IDADE_DIAG	STATUS	MOTIVO	TEMPO	ENPOINT
1	M	59	0	ncancer	0,18	1
1	M	59	1	ncancer	0,18	2
2	M	31	0	ncancer	0,17	1
2	M	31	1	ncancer	0,17	2
3	M	79	1	mcancer	0,07	1
3	M	79	0	mcancer	0,07	2
4	M	69	0	censura	5,25	1
4	M	69	0	censura	5,25	2

Figura 3: Exemplo de estruturação do banco de dados.

2.4.2 Códigos R

O software R é um programa gratuito e de código aberto disponibilizado em <http://www.r-project.org>. Após instalação do software deve-se instalar e executar a biblioteca `survival`^[19], necessária para proceder às análises. Os comandos necessários para essa etapa, que devem ser copiados e colados no console do R, são:

```
install.packages("survival")
require(survival)
```

Os comandos a seguir foram utilizados para gerar as figuras 4 e 5. Para se produzir a Figura 4(a), que mostra a função distribuição acumulada pelo método de Kaplan-Meier, considerou-se primeiramente como desfecho de interesse o *óbito por câncer*, sendo os demais eventos tratados como censura. Em seguida realizou-se o processo inverso, ou seja, considerou-se como desfecho de interesse o *óbito por causa não relacionada ao câncer* e os demais eventos foram tratados como censura.

```
obito_cancer = survfit(formula = Surv(TEMPO, KM_CANCER)~ 1, type = "kaplan-meier",
data = dados)
obito_naocancer = survfit(formula = Surv(TEMPO, KM_NAOCANCER)~ 1, type = "kaplan-
meier", data = dados)
```

A curva de sobrevida estimada pelo método de Kaplan-Meier é gerada através do comando `survfit(formula = , data =)`. O primeiro argumento desta função é a fórmula, que possui o formato: `formula = Surv(tempo, status) ~ 1`. O tempo de sobrevida é construído utilizando o comando `Surv()`. O segundo argumento refere-se ao banco de dados onde estão as variáveis tempo e status. Para gerar o gráfico da função distribuição acumulada é necessário usar o comando `plot()` do objeto em que foi armazenada a função `survfit()` acrescentando o argumento `fun = "event"`:

```
plot(obito_cancer, conf.int = F, fun = "event", xlab = "Tempo de acompanhamento
(anos)", lty = 2, ylim = c(0,1), ylab = "Função distribuição acumulada")
lines(obito_naocancer, conf.int = F, fun = " event ", lty = 1, ylim = c(0,1))
legend("topleft", c("Óbito por câncer", "Óbito por causa não relacionada ao
câncer"), lty = c(2,1), bty = "n")
```

Optou-se por utilizar o estimador de Kaplan-Meier, mas as mesmas análises poderiam ser feitas utilizando o modelo de Cox. Para tal é necessário substituir o comando `survfit()` por `coxph()`.

Para gerar a Figura 4(b) utilizou-se o modelo de riscos competitivos, na abordagem da subdistribuição dos riscos ou função de incidência acumulada, considerando todos os tipos de cânceres e os seguintes eventos: *óbito por câncer* e *óbito por causa não relacionada ao câncer*. A fim de construir a função de incidência acumulada utilizando o pacote `survival` é necessário acrescentar o argumento `etype` ao vetor onde estão classificados os diferentes eventos. Considerando que a variável `status` pode ter diversos valores, dependendo da quantidade de eventos competitivos, é necessário indicar no argumento `event`, da função

Surv() os valores indicativos do evento. Vale ressaltar que os eventos censurados devem estar classificados sempre com o valor zero.

```
fia_cancer = survfit(Surv(TEMPO, event = STATUS > 0) ~ 1, etype = MOTIVO, data =
dados)
plot(fia_cancer, xlab = "Tempo de acompanhamento (anos)", ylab = "FIA", fun =
"event", lty = c(2,1), ylim = c(0,1))
lines(obito_cancer, fun = "event", lty = 1, col = "gray")
lines(obito_naocancer, fun = "event", lty = 1, col = "gray")
legend("topleft", c("Óbito por câncer", "Óbito por causa não relacionada ao
câncer", "Estimativas por Kaplan-Meier"), lty = c(2,1,1), bty = "n", col =
c("black", "black", "gray"))
```

3. Resultados

A coorte de pacientes estudada é composta por 576 (42,2%) mulheres e 788 (57,8%) homens. A idade média ao diagnóstico primário de câncer é de 58,3 anos em relação a todos os pacientes. Entre as mulheres a idade média ao diagnóstico é de 57,3 anos, já entre os homens é de 59,0 anos. Em relação ao tempo de acompanhamento, o tempo médio de seguimento foi de 3,7 anos (3,9 anos entre as mulheres e 3,5 anos entre os homens).

Os tipos de neoplasias malignas mais incidentes em 2006 na população estudada são: pele (232 casos – 17,0%), próstata (142 casos – 10,4%), brônquios e pulmões (98 casos – 7,2%), mama (92 casos – 6,7%), cólon (77 casos – 5,6%) e sistema hematopoiético e retículo endotelial (67 casos – 4,9%). Quanto à mortalidade, do total de 600 óbitos, 498 têm como causa básica de morte o câncer e em 102 óbitos a causa básica não é relacionada ao câncer. Os tipos de cânceres com mais óbitos são: brônquios e pulmões (89 óbitos – 14,8%), esôfago (44 óbitos – 7,3%), cólon (29 óbitos – 4,8%), estômago (25 óbitos – 4,2%) e próstata (22 óbitos – 3,7%).

A idade no diagnóstico foi calculada através da diferença entre a data de diagnóstico e a data de nascimento; o tempo de acompanhamento para os casos de não óbito foi obtido através da diferença entre o final do período de observação da coorte (31/12/2011) e a data do diagnóstico. Já para os casos de óbito, foi obtido através da diferença entre a data de óbito e a data do diagnóstico. Para as análises descritivas das causas de morte e localização primária do tumor utilizaram-se os dois primeiros dígitos do código CID-10 a fim de agrupar neoplasias em locais iguais, independente da especificação do local.

A Figura 4 mostra a diferença entre o resultado da análise de sobrevivência para os desfechos *óbito por câncer* e *óbito por causa não relacionada ao câncer* utilizando o método tradicional de Kaplan-Meier, Figura 4(a) – que considera um desfecho como principal e os demais como censura – e o modelo de riscos competitivos, Figura 4(b). O resultado do método de Kaplan-Meier indica que, considerando todos os tipos de câncer, a taxa de letalidade do câncer, independente do tipo, é de cerca de 40% em cinco anos. Já considerando a abordagem de riscos competitivos, essa taxa reduz para aproximadamente 36%. Além disso, como esperado, esse resultado sugere que, dentre os pacientes com câncer, o risco de morrer pela doença é maior do que o de morrer por causa não relacionada a ela.

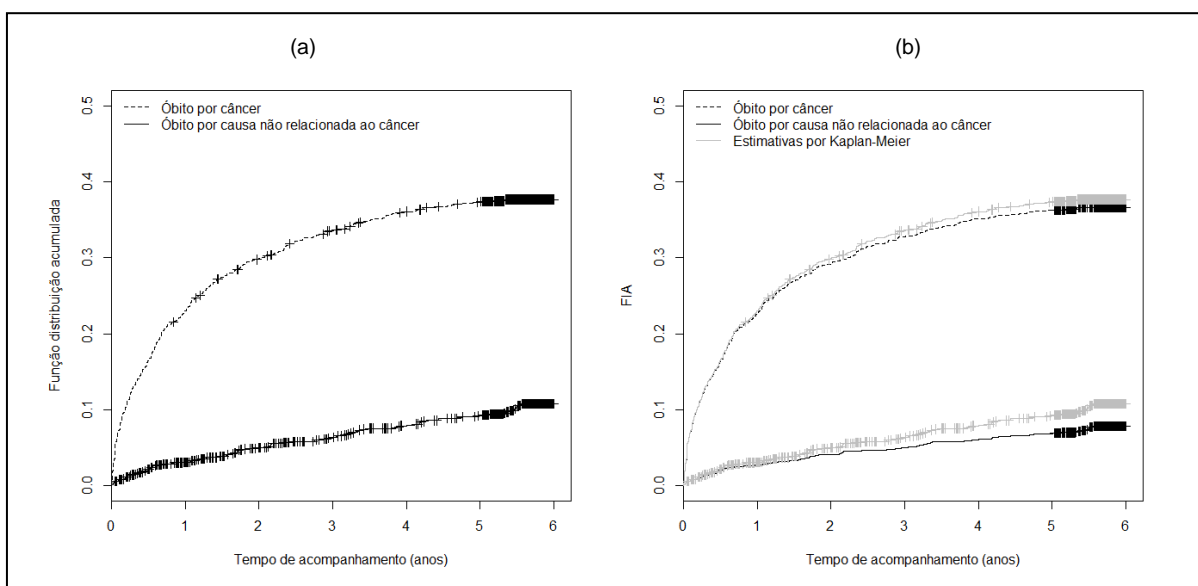


Figura 4: Análise de sobrevida utilizando o método de Kaplan-Meier e o modelo de riscos competitivos. A figura (a), método de Kaplan-Meier, mostra a função distribuição acumulada para os desfechos *óbito por câncer*, considerando os demais eventos como censura, e *óbito por causa não relacionada ao câncer*, considerando os demais eventos como censura. Na figura (b), modelo de riscos competitivos, tem-se a função de incidência acumulada para cada um dos desfechos competitivos: *óbito por câncer* e *óbito por causa não relacionada ao câncer* (a estimativa obtida por Kaplan-Meier é mostrada em cinza).

Na Figura 5 tem-se o resultado do modelo de riscos competitivos em pacientes diagnosticados com câncer de pulmão – Figura 5(a) -, considerando os desfechos: *óbito por câncer de pulmão* e *óbito por outros tipos de câncer ou causas não relacionadas ao câncer*. De maneira análoga, a Figura 5(b) mostra a FIA estimada para pacientes diagnosticados com câncer de esôfago, sendo os possíveis desfechos *óbito por câncer de esôfago* e *óbito por outros tipos de câncer ou causas não relacionadas ao câncer*. A fim de comparação, em cada um dos gráficos acrescentou-se o resultado obtido pelo método de Kaplan-Meier, que não considera os diferentes tipos de óbito. Especificamente para o câncer de pulmão, a taxa de letalidade pelo método de Kaplan-Meier é de aproximadamente 78% em cinco anos, enquanto pela abordagem de riscos competitivos essa taxa reduz para cerca de 70%. Para o câncer de esôfago essas taxas variam entre 70%, na abordagem de Kaplan-Meier e 60% quando se considera os riscos competitivos.

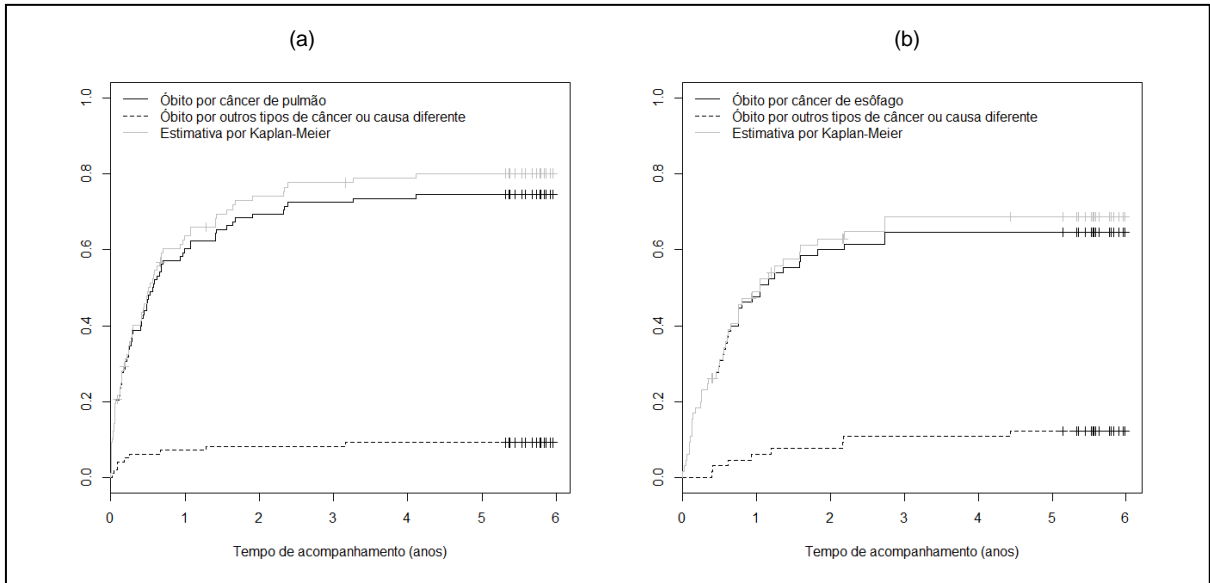


Figura 5: Análise de sobrevivência utilizando o modelo de riscos competitivos para os casos de câncer de pulmão e esôfago, respectivamente. Em cada uma das figuras está incluída, a fim de comparação, a estimativa obtida por Kaplan-Meier. A figura (a) mostra a função de incidência acumulada para os desfechos *óbito por câncer de pulmão* e *óbito por outros tipos de câncer ou causa não relacionada ao câncer*. Na figura (b) temos a função de incidência acumulada para os desfechos *óbito por câncer de esôfago* e *óbito por outros tipos de câncer ou causa não relacionada ao câncer*.

4. Discussão

Uma vez que o objetivo principal desse artigo consiste na divulgação do modelo de riscos competitivos, os aspectos epidemiológicos dos resultados obtidos não serão discutidos. Sendo assim, a discussão terá como foco apenas a análise estatística dos dados. Acrescenta-se que todas as análises realizadas foram procedidas no pacote *survival* do software R versão 2.15.0 e não consideram nenhuma covariável. A vantagem de utilizar o R é que se trata de um programa livre, de código aberto e de fácil utilização.

Através desse artigo, mostrou-se que desprezar a existência de eventos competitivos pode afetar o resultado da distribuição acumulada. Na Figura 4, foram comparados os resultados do método tradicional de Kaplan-Meier com o modelo de riscos competitivos considerando todos os tipos de câncer. Fica evidente, pela Figura 4(a), que o método de Kaplan-Meier superestima a função distribuição acumulada para ambos os desfechos, uma vez que cada um deles foi obtido considerando o outro como censura. Ou seja, os valores da distribuição acumulada para cada desfecho individual é maior que os valores da função de incidência acumulada que leva em conta ambos os desfechos ao mesmo tempo, resultando no fato de que a estimativa das funções em separado é viesada ^[14].

A correção desse problema é dada pela utilização da subdistribuição (FIA), apresentada na Figura 4(b). Considerou-se como desfechos competitivos o *óbito por câncer* e o *óbito por causa não relacionada ao câncer*. A ideia é que para se estimar de forma adequada a probabilidade de óbito por câncer, o óbito por causa não relacionada ao câncer deve ser considerado um evento competitivo e vice-versa. Dessa forma, a cada momento consideramos um dos desfechos como de interesse e o outro como competitivo ^[14]. Vale reforçar que a distribuição acumulada da Figura 4(a) está superestimada, como demonstra a curva em cinza, resultado da análise por Kaplan-Meier.

Para a avaliação das taxas de letalidade dos cânceres mais frequentes na população em estudo procedeu-se a análise de sobrevivência para os desfechos competitivos *óbito por câncer de pulmão* e *óbito por outros cânceres ou causa não relacionada ao câncer* – Figura 5(a) – e *óbito por câncer de esôfago* e *óbito por outros cânceres ou causa não relacionada ao câncer* – Figura 5(b). As categorias *morte por outros cânceres* e *causa não relacionada ao câncer* foram agrupadas em razão do pequeno número de casos observados em cada uma. A fim de comparar o resultado do modelo de riscos competitivos com a abordagem tradicional, acrescentou-se a cada um dos gráficos da Figura 5, em cor cinza, o resultado obtido pelo método de Kaplan-Meier considerando os óbitos por câncer dentre os pacientes com câncer de

pulmão – censurando as demais causas – e os óbitos por câncer dentre os pacientes com câncer de esôfago, respectivamente.

Novamente, é nítido que o método de Kaplan-Meier superestima a função distribuição acumulada para ambos os tipos de câncer. Assim, como conclusão, recomenda-se aos pesquisadores que não ignorem em seus estudos os desfechos competitivos. Não considerá-los resulta em estimativas superestimadas da distribuição acumulada dos desfechos de interesse.

Finalizando, sugere-se continuar o estudo considerando covariáveis nas análises a fim de verificar que seus efeitos podem não ser estimados adequadamente se eventos competitivos forem desprezados. Além disso, propõe-se utilizar outros bancos de dados e considerar outros desfechos. Mais estudos com foco na abordagem metodológica devem ser produzidos objetivando auxiliar os pesquisadores da área médica a lidar com eventos dessa natureza, ressaltando suas vantagens sobre as abordagens tradicionais.

Referências bibliográficas

- [1] Wunsch VF, Moncau JE. Mortalidade por câncer no Brasil 1980-1975: padrões regionais e tendências temporais. Revista da Associação Médica Brasileira. 2002; 48(3): 250-57.
- [2] Alwan A et al. Global status report on noncommunicable diseases 2010. World Health Organization, 2011. 176 p.
- [3] Instituto Nacional de Câncer. Atlas de Mortalidade por Câncer [Internet]. Rio de Janeiro (RJ): Ministério da Saúde;[atualizado 08 de junho de 2012 ; citado 07 de setembro de 2012]. Disponível em: <http://mortalidade.inca.gov.br/Mortalidade/>.
- [4] Instituto Nacional de Câncer. Estimativas de incidência e mortalidade por câncer no Brasil [Internet]. Rio de Janeiro (RJ): Ministério da Saúde; 2003 [citado 07 de setembro de 2012]. Disponível em: http://bvsmms.saude.gov.br/bvs/publicacoes/23estimativas_incidencia.pdf.
- [5] Instituto Nacional de Câncer. Tipos de câncer – Câncer de mama [Internet]. Rio de Janeiro (RJ): Ministério da Saúde; 2012 [citado 07 de setembro de 2012]. Disponível em: <http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama>.
- [6] Instituto Nacional de Câncer. Estimativa 2012 – Incidência de câncer no Brasil [Internet]. Rio de Janeiro (RJ): Ministério da Saúde; 2012 [citado 15 de setembro de 2012]. Disponível em: <http://www.inca.gov.br/estimativa/2012/index.asp?ID=5..>
- [7] Instituto Nacional de Câncer. Fazer a mamografia a cada dois anos é uma das novas recomendações do Inca para combater o câncer de mama [Internet]. Rio de Janeiro (RJ): Ministério da Saúde; 2010 [citado 15 de setembro de 2012]. Disponível em: http://www2.inca.gov.br/wps/wcm/connect/comunicacaoinformacao/site/home/namidia/fazer_mamografia_cada_dois_anos_novas_recomendacoes_inca_combater_cancer_mama.
- [8] Schmidt MI et al. Chronic non-communicable diseases in Brazil: burden and current challenges. The Lancet. 2011; 377(9781): 1949-61.

[9] Moraes AB et al. Estudo da sobrevida de pacientes com câncer de mama atendidas no hospital da Universidade Federal de Santa Maria, Rio Grande do Sul, Brasil. Cadernos de Saúde Pública. 2006; 22(10): 2219-28.

[10] Instituto Nacional de Câncer. Sobrevida, estadiamento – relação incidência/mortalidade [Internet]. Rio de Janeiro (RJ): Ministério da Saúde; [citado 01 de setembro de 2012]. Disponível em: http://www1.inca.gov.br/situacao/arquivos/ocorrencia_sobrevida.pdf.

[11] Instituto Nacional de Câncer. Tipos de câncer – Câncer de pulmão [Internet]. Rio de Janeiro (RJ): Ministério da Saúde; 2012 [citado 15 de setembro de 2012]. Disponível em: <http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/pulmao>.

[12] PUBMED. Search [Internet]. Estados Unidos (USA): Centro Nacional para Informação Biotecnológica; [citado em 17 de dezembro de 2012]. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed>.

[13] Colosimo EA, Giolo SR. Análise de sobrevivência aplicada. 1. ed. São Paulo: Edgard Blücher, 2006. 370 p.

[14] Andreozzi VL, Barbosa MTS, Carvalho MS, Codeço CT, Shimakura SE. Análise de sobrevida – Teoria e Aplicações em Saúde. 2. ed. revista e ampliada. Rio de Janeiro: Fiocruz, 2011. 434 p.

[15] Andersen PK, Geskus RB, Witte TD, Putter H. Competing risks in epidemiology: possibilities and pitfalls. International Journal of Epidemiology. 2012; 41(3):861–70.

[16] Medicinanet. Lista CID 10 [Internet]. Porto Alegre (RS): Medicinanet Informações de Medicina S/A; [citado em 02 de novembro de 2012]. Disponível em: <http://medicinonet.com.br/cid10.htm>.

[17] Departamento de Análise de Situação em Saúde, Secretaria de Vigilância em Saúde, Ministério da Saúde. Saúde Brasil 2006: uma análise da situação de saúde no Brasil. Brasília: Ministério da Saúde; 2006. (Série G. Estatística e Informação em Saúde).

[18] R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Disponível em: <http://www.r-project.org>.

[19] Therneau, T. (2012). Survival analysis, including penalised likelihood – Package ‘survival’. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Disponível em: <http://cran.r-project.org/web/packages/survival/survival.pdf>.