

Modelo aditivo de Aalen: uma aplicação para dados de sinusite em pacientes com Aids

Aalen Additive Model: A Application for Data of Sinusite in Patients with Aids

TARCIANA LIBERAL PEREIRA^{1,a}, ENRICO ANTÔNIO COLOSIMO^{2,b},
MARIA CRISTINA RAPOSO^{3,c}

¹DEPARTAMENTO DE ESTATÍSTICA-UFPB, JOÃO PESSOA, BRASIL

²DEPARTAMENTO DE ESTATÍSTICA-UFGM, BELO HORIZONTE, BRASIL

³DEPARTAMENTO DE ESTATÍSTICA-UFPE, RECIFE, BRASIL

Resumo

Em estudos de sobrevivência, as vezes, o interesse não é apenas na distribuição do tempo de falha. É comum a comparação de tempos de sobrevivência de dois ou mais grupos bem como a verificação do efeito de covariáveis na resposta tanto em engenharia quanto, principalmente, em situações clínicas. A maneira mais eficiente de incorporar o efeito desses fatores no estudo é utilizar um modelo estatístico de regressão. Frequentemente em estudos de sobrevivência o efeito das covariáveis de interesse pode variar ao longo do tempo de duração do estudo. Estas covariáveis são denominadas de dependentes do tempo e a inclusão delas na análise pode fornecer resultados mais precisos. Aalen propôs um modelo de risco aditivo que apresenta vantagens práticas quando as covariáveis são acompanhadas ao longo do tempo e os seus valores podem ser modificados durante o estudo. A principal vantagem desse modelo é que através de análise gráfica é possível verificar mudanças no tempo na influência de cada uma das covariáveis. Isto é, análises com este modelo fornecem informações detalhadas a respeito da influência temporal de cada covariável. Dessa forma o modelo aditivo de Aalen é apresentado neste trabalho com o objetivo de mostrar a sua importância na presença de covariáveis dependentes do tempo. Um banco de dados real envolvendo pacientes infectados pelo HIV e o tempo até o desenvolvimento de sinusite é utilizado para ilustrar o ajuste deste modelo.

Palavras chave: covariáveis dependentes do tempo, modelo de Aalen, riscos aditivos.

^aProfessora assistente. E-mail: tarciana@de.ufpb.br

^bProfessor adjunto. E-mail: enricoc@est.ufmg.br

^cProfessora adjunta. E-mail: cristina@de.ufpe.br

Abstract

In survival analysis sometimes the interest is not just on the failure time distribution function. It is common treatment comparisons as well as studying the effect of covariates in the response. This fact happens in engineering and clinical studies. Covariates effects are usually incorporated in the analysis by using a regression model. Moreover, in some real situations, covariates may be monitored and measured along the follow-up period. These covariates are known as time-dependent covariates. Analysis that include these covariates can be more reliable. Aalen proposed an additive risk model that is very attractive. This model has showed some practical advantages especially when the covariates effects varies in time. The main advantage of this model it is that through graphical analysis it is possible to verify changes in the time in the influence of each one of the covariates. That is, analyses with this model supply information detailed regarding the secular influence of each covariate. This model is presented in this paper in terms of time-dependent covariates. A real data set related to HIV patients and time to develop sinusitis is used to illustrate the fit of the additive Aalen model.

Key words: Time-dependent covariate, Aalen model, Additive risk.

1. Introdução

A análise de sobrevivência é uma das áreas da estatística que mais tem crescido nos últimos anos. Este crescimento pode ser observado tanto no desenvolvimento e aprimoramento de métodos estatísticos, quanto no número de aplicações. Estas técnicas foram desenvolvidas primeiramente nas ciências médicas e biológicas, sendo essa uma das áreas de maior aplicação, mas são usadas também extensamente na engenharia, ciências sócias e econômicas.

Freqüentemente em análise de sobrevivência quando covariáveis são incorporadas na análise os seus valores registrados são aqueles medidos no início do estudo. Contudo em muitos estudos que envolvem dados de sobrevivência existem covariáveis, conhecidas como Covariáveis Dependentes do Tempo, que são monitoradas durante o estudo e seus valores mudam neste período. Estas covariáveis têm muita utilidade na análise de dados de sobrevivência pois podem ser utilizadas tanto para acomodar medidas que variam com o tempo durante o estudo como também para modelar o efeito de indivíduos que mudam de grupo durante um tratamento. Análises que consideram estas covariáveis podem fornecer resultados mais precisos e a não inclusão delas pode acarretar em sérios vícios de estimação.

Um modelo bastante utilizado para tal situação é o modelo de regressão de Cox que pode ser generalizado para incorporar o efeito destas covariáveis dependentes do tempo. Um modelo de risco aditivo ou linear, alternativo ao de Cox, foi sugerido por Aalen em 1980. Aalen (1980, 1989, 1993) mostrou que seu modelo tem apresentado freqüentemente vantagens práticas especialmente quando as covariáveis têm efeitos variando no tempo pois permite a observação de mudanças no tempo na influência de cada covariável separadamente.

Considerando a grande utilidade dessas técnicas em aplicações na área médica e, a necessidade de analisar modelos contendo covariáveis dependentes no tempo pretende-se com este artigo apresentar o modelo aditivo de Aalen e aplicá-lo a um conjunto de dados envolvente pacientes com AIDS e cuja resposta é o tempo até o desenvolvimento de sinusite.

Na Seção 2 apresenta-se o modelo aditivo de Aalen. Além da descrição do modelo é apresentado um gráfico para verificar o comportamento dos efeitos das covariáveis no tempo. Na Seção 3 é apresentada uma aplicação do modelo descrito a dados incluindo uma covariável dependente do tempo. Os dados utilizados fazem parte de um estudo relacionado a manifestações otorrinolaringológicas em pacientes com AIDS. Pretende-se identificar se a infecção pelo HIV influencia na ocorrência de sinusite. A covariável que indica o grupo de classificação quanto a infecção pelo HIV é dependente do tempo. O artigo termina na seção 4 com alguns comentários finais.

Para realização deste artigo foi utilizado o software R (R Development Core Team 2006). Este software é interpretado como uma linguagem computacional designada para análise de dados estatísticos que se caracteriza pelo compromisso entre a grande flexibilidade oferecida pelas linguagens compiladas, tais como C, C++ e FORTRAN e a conveniência de softwares estatísticos tradicionais. Inclui uma ampla variedade de métodos estatísticos tradicionais e modernos. Para mais detalhes sobre esta linguagem de programação ver Cribari-Neto & Zarkos (1999). A versão utilizada neste trabalho foi a 1.7.0 e está disponível no endereço <http://www.r-project.org>. O programa R foi utilizado neste trabalho para obtenção de todos os resultados que estão apresentados na Seção 3. Para aplicar o modelo de Aalen foi necessário utilizar a função `addreg`, disponível no endereço <http://www.med.uio.no/imb/stat/addreg/>.

2. Modelo aditivo de Aalen

Na teoria clássica de regressão, a esperança das variáveis respostas é o objeto principal de modelagem. Em análise de sobrevivência freqüentemente a função de risco é a base da modelagem de regressão. O risco é uma função natural para descrever a distribuição do tempo de vida. Informalmente a função de risco ou de taxa de falha mede a probabilidade de um evento ocorrer em um dado intervalo de tempo condicional a sobrevivência ao tempo imediatamente anterior e dividido pelo seu comprimento. Existem várias possibilidades de modelos de regressão que têm como base a função de risco. Uma delas é o tão conhecido modelo de riscos proporcionais de Cox com a sua verossimilhança parcial, que tem as vantagens de uma simples interpretação dos resultados e de estar disponível em vários softwares computacionais. Entretanto Aalen (1989) citou algumas limitações do modelo de Cox. A primeira delas é que as suposições do modelo podem não valer, as vezes o modelo de Cox é usado na literatura sem que suas propriedades sejam cheçadas e também não é claro se satisfazendo as propriedades usuais de proporcionalidade garantem a adequação do modelo de Cox. Em segundo lugar mudanças ao longo do tempo na influência das covariáveis não são facilmente identificadas e o modelo

de Cox não é adaptado para uma descrição detalhada de efeitos de covariáveis ao longo do tempo. Por último a suposição de proporcionalidade do risco é vulnerável á mudanças no número de covariáveis modeladas. Se algumas covariáveis são retiradas de um modelo ou medidas com um diferente nível de precisão, a proporcionalidade é geralmente afetada. Portanto Aalen verificou uma falta de consistência do modelo de Cox a este respeito.

Estas limitações conduziram a uma ampla variedade de modelos que generalizam o modelo de Cox. Uma alternativa, baseada em processos de contagem (Aalen 1978), foi sugerida originalmente por Aalen (1980). Este modelo, apresentado de forma mais simples em 1989 (Aalen 1989), é de riscos aditivos para análise de regressão de dados censurados e permite que tanto os parâmetros quanto os vetores de covariáveis variem com o tempo. Dessa forma o modelo de Aalen é capaz de fornecer informações detalhadas a respeito da influência temporal de cada covariável. Uma outra vantagem é que o modelo de Aalen é completamente não-paramétrico no sentido de que funções são ajustadas e não parâmetros. Ou seja, na estimação das quantidades desconhecidas, o modelo de Aalen usa apenas informação local o que faz este modelo bastante flexível. Os estimadores propostos por Aalen generalizam o tão conhecido estimador de Nelson-Aalen que é o estimador natural no caso de população homogênea. Aplicações foram apresentadas por Mau(1986, 1988) e Andersen & Vaeth (1989) e resultados teóricos foram feitos por McKeague (1986), McKeague & Utikal (1988) e Huffer & McKeague (1987) indicando que o modelo pode ser útil e é sem dúvida razoável para explorar vantagens da linearidade analogamente a teoria clássica de modelo linear.

Em um estudo típico, um número de indivíduos é observado ao longo do tempo para verificar a ocorrência de um determinado evento. O acontecimento deste evento é assumido independente entre os indivíduos e tem-se um tempo até a ocorrência do evento para cada indivíduo, cuja distribuição depende de um vetor dado por $x_i(t) = (1, x_{1i}(t), x_{2i}(t), \dots, x_{pi}(t))'$ onde $x_{ij}(t)$, com $j = 1, \dots, p$, são os valores observados, para o i -ésimo indivíduo, das covariáveis que podem variar no tempo. Seja n o número de indivíduos, p o número de covariáveis na análise e $h_i(t)$ a função de risco para o tempo de sobrevivência t de um indivíduo i .

O modelo de risco aditivo de Aalen é dado por

$$h_i(t) = \alpha_0(t) + \sum_{j=1}^p \alpha_j(t)x_{ij}(t) \quad (1)$$

Considerando a forma matricial

$$h(t) = Y(t)\alpha(t)$$

em que $\alpha(t) = (\alpha_0(t), \alpha_1(t), \dots, \alpha_p(t))'$ é um vetor de funções do tempo desconhecidas, cujo primeiro elemento $\alpha_0(t)$ é interpretado como uma função de parâmetro básica, enquanto que os $\alpha_j(t)$'s, $j = 1, \dots, p$, chamados de funções de regressão, medem a influência das respectivas covariáveis. A matriz $Y(t)$ de ordem $n \times (p+1)$ é construída da seguinte maneira: se o evento considerado ainda não ocorreu para o i -ésimo indivíduo e ele não é censurado então a i -ésima linha de $Y(t)$ é o vetor

$x_i(t) = (1, x_{i_1}(t), x_{i_2}(t), \dots, x_{i_p}(t))'$. Caso contrário, se o indivíduo não está sob risco no tempo t , então a linha correspondente de $Y(t)$ contém apenas zeros.

No modelo de Aalen assume-se que as covariáveis agem de maneira aditiva na função de risco e os coeficientes de riscos desconhecidos podem ser funções do tempo, ou seja, os efeitos das covariáveis podem variar durante o estudo revelando mudanças na influência das covariáveis, sendo esta uma das vantagens do modelo (1).

Os estimadores dos parâmetros são baseados nas técnicas de mínimos quadrados e a obtenção dessas quantidades é similar aos do estimador de Nelson-Aalen para a função de risco acumulada (Klein & Moeschberger 1997). A aproximação para estimação depende das suposições sobre a forma funcional das funções de regressão que, neste caso, são não-paramétricas. Considerando que é mais fácil estimar uma função de distribuição acumulada do que uma função de densidade de probabilidade estima-se a função de risco acumulada dado que a estimação direta das funções de regressão é difícil na prática. Considera-se então a estimação do vetor coluna $A(t)$ com elementos $A_j(t)$ dados por

$$A_j(t) = \int_0^t \alpha_j(s) ds$$

$j = 1, \dots, p$. Sejam $t_1 < t_2 < \dots < t_k$ os tempos distintos de falhas ordenados. Aalen considerou um estimador razoável de $A(t)$, denominado estimador de mínimos quadrados de Aalen, que é dado por

$$A^*(t) = \sum_{t_k \leq t} Z(t_k) I_k \quad (2)$$

em que I_k é um vetor de zeros que assume o valor 1 somente para os indivíduos cujo evento ocorre no tempo t_k . Enquanto que $Z(t)$ é a inversa generalizada de $Y(t)$. Em princípio, $Z(t)$ pode ser qualquer inversa generalizada de $Y(t)$. Uma escolha simples pode ser baseada no princípio de mínimos quadrados local, ou seja

$$Z(t) = [Y(t)'Y(t)]^{-1}Y(t)'$$

É importante notar que o estimador de $A(t)$ é definido apenas sobre um intervalo de tempo em que $Y(t)$ tem posto completo. Ou seja, a estimação termina assim que $Y(t)$ perde o posto completo, que é uma consequência do princípio não paramétrico.

Os componentes de $A^*(t)$ convergem assintoticamente, sob condições apropriadas, para um processo gaussiano. Então um estimador da matriz de covariância de $A^*(t)$ é dado por

$$\Omega^*(t) = \sum_{t_k \leq t} Z(t_k) I_k^D Z(t_k)'$$

em que I_k^D é uma matriz diagonal com I_k como diagonal.

Como consequência dos resultados obtidos anteriormente, pode-se estimar o risco acumulado e a função de sobrevivência correspondentes dados os valores

das covariáveis. Seja $x(t) = (1, x_1(t), x_2(t), \dots, x_p(t))'$ o conjunto de valores das covariáveis no tempo t . O estimador do risco acumulado $H^*(t)$ é dado por

$$H^*(t) = A^*(t)'x(t)$$

A função de sobrevivência pode então ser estimada por

$$S^*(t) = \exp(-H^*(t)) \quad (3)$$

É possível testar se uma covariável específica tem algum efeito na função de risco total. A hipótese nula para algum j , com $j = 1, \dots, p$, é estabelecida como

$$H_j : \alpha_j(t) = 0, \quad t \in [0, T]$$

em que T é o maior tempo observado.

Aalen (1980) desenvolveu, para todo tempo de falha, uma estatística de teste para H_j dada pelo j -ésimo elemento U_j do vetor

$$U = \sum_{t_k} K(t_k)Z(t_k)I_k \quad (4)$$

em que $K(t)$, uma função peso não negativa, é uma matriz diagonal $(p+1) \times (p+1)$. A estatística de teste da Equação (4) surge como uma combinação ponderada da soma do estimador de $A_j(t)$. Os elementos diagonais de $K(t)$ são funções pesos e suas escolhas podem depender das alternativas para a hipótese nula de interesse.

Uma escolha ótima da função peso necessitará do conhecimento das verdadeiras variâncias dos estimadores, entretanto isto dependerá de funções de parâmetros desconhecidas.

Uma escolha, considerada por Aalen (1989), é tomar

$$K(t) = \{ \text{diag} [(Y(t)'Y(t))^{-1}] \}^{-1}$$

em que $K(t)$ é dada como a inversa de uma matriz diagonal tendo a mesma diagonal principal da matriz $(Y(t)'Y(t))^{-1}$. Este peso é escolhido por analogia ao problema da regressão de mínimos quadrados em que as variâncias dos estimadores são proporcionais aos elementos diagonais da matriz $(Y'Y)^{-1}$. Um estimador para a matriz de covariância de U é dado por

$$V = \sum_{t_k} K(t_k)Z(t_k)I_k^D Z(t_k)'K(t_k)' \quad (5)$$

Suponha que se queira testar simultaneamente todos H_j para j em algum subconjunto A de $\{1, \dots, p\}$ consistindo de s elementos. Seja U_A definido como o subvetor correspondente de U e V_A a submatriz correspondente de V , isto é, V_A é a matriz de covariâncias estimadas de U_A . A estatística de teste normalizada $U_A'V_A^{-1}U_A$ é assintoticamente distribuída como uma distribuição qui-quadrado com s graus de liberdade quando H_j vale para todo j em A . Se o interesse é

testar apenas uma das hipóteses H_j , então é usada a estatística de teste $U_j V_{jj}^{-1/2}$. Esta estatística tem uma distribuição assintótica normal padrão sob a hipótese nula.

Em alguns estudos, a significância de uma covariável pode mudar durante o período de acompanhamento. Através do modelo de Aalen é possível estimar a contribuição das covariáveis para a função de risco em cada tempo de falha. O resumo desta contribuição sobre o tempo produz uma função de regressão para cada covariável que pode ser apresentada em um gráfico contra o tempo. Ou seja, $A_j^*(t)$ pode ser considerada como uma função empírica descrevendo a influência da j -ésima covariável.

Esse gráfico das funções de regressão do modelo linear de Aalen pode ser usado para detectar efeitos de covariáveis dependentes do tempo ou como uma técnica de diagnóstico que pode extrair informações adicionais úteis. A inclinação do gráfico da função de regressão acumulada contra o tempo fornece informação sobre a influência de cada covariável, sendo possível verificar se uma covariável particular tem um efeito constante ou varia com o tempo ao longo do período de estudo. Por exemplo, se $\alpha_j(t)$ é constante, então o gráfico deve aproximar-se de uma linha reta paralela ao eixo do tempo e isso ocorre em períodos em que as covariáveis não influenciam a função risco.

Através de um estudo, Mau (1986) mostrou que as funções de regressão podem fornecer informações importantes que podem ser perdidas quando outro modelo é utilizado.

3. Ilustração numérica: dados de sinusite em pacientes com Aids

Estudos na área médica realizados em pacientes com AIDS (“Acquired Immnodeficiency Syndrome”) indicam que a infecção pelo HIV (“Human Immunodeficiency”) é fator de risco para o desenvolvimento de doenças otorrinolaringológicas (ORL). Os primeiros estudos não mencionaram a sinusite ou citaram apenas casos isolados de sinusite crônica. Sample et al. (1989) realizaram um estudo sobre manifestações ORL em pacientes com AIDS. Nos resultados obtidos 30% dos pacientes com AIDS apresentaram sinusite, sendo assim demonstrado que a sinusite era uma manifestação ORL mais freqüente do que citado em estudos anteriores. A sinusite se apresenta na AIDS com elevada freqüência e precária resposta aos tratamentos administrados. Entretanto a literatura mundial ainda não determinou a importância da sinusite no contexto geral da síndrome e outros fatores de riscos relacionados a estas manifestações em pacientes com AIDS têm sido pouco estudados.

Os dados utilizados nesta aplicação fazem parte de um estudo para avaliar a incidência de manifestações ORL em pacientes infectados pelo HIV realizado no Hospital das Clínicas da Universidade Federal de Minas Gerais (UFMG) Gonçalves (1995). Os pacientes que fizeram parte do estudo foram acompanhados no período de março de 1993 a fevereiro de 1995, considerando apenas os pacientes que entraram no estudo até julho de 1994. Desta forma o mecanismo de censura

utilizado foi a censura do tipo I, que ocorre quando o estudo é terminado após um período pré-estabelecido de tempo.

Para entrar no estudo os pacientes tinham que ter idade superior a 15 anos, ter um exame prévio HIV positivo ou pertencer a grupos de comportamento de risco para adquirir o HIV, tais como, indivíduos que têm relações sexuais sem o uso de preservativo com parceiro desconhecido ou que possa ser portador do vírus, indivíduos que usam drogas endovenosas compartilhando agulhas e indivíduos que sofrem transfusões de sangue.

As doenças ORL avaliadas nos estudos de prevalência destas manifestações na literatura em pacientes infectados pelo HIV foram: afta, candidíase oral, herpes labial, oral ou nasal, leucoplasia pilosa, sinusite, sarcoma de kaposi, otite aguda e serosa, linfadenopatia cervical no cavum e na glândula parótida. Entre as enfermidades citadas foi utilizada neste trabalho apenas a sinusite. A classificação dos pacientes com relação a infecção pelo HIV foi de acordo com os critérios do CDC ("Centers of Disease Control", 1987) onde os pacientes foram classificados como HIV soronegativo (grupo 1), HIV soropositivo assintomático (grupo 2), pacientes com ARC ("AIDS Related Complex") (grupo 3) ou pacientes com AIDS (grupo 4). Pacientes HIV soronegativo são aqueles que não possuem o HIV, constituindo o grupo controle do estudo. Pacientes HIV soropositivo assintomático são aqueles que possuem o vírus mas não desenvolveram o quadro clínico de AIDS e que apresentam um perfil imunológico estável. Os pacientes com ARC apresentam baixa imunidade e outros indicadores clínicos que antecedem o quadro clínico de AIDS. E, por último, os pacientes com AIDS são aqueles que já desenvolveram infecções oportunistas que segundo o critério do CDC de 1987 definem o quadro clínico de AIDS. O acompanhamento foi feito através de consultas trimestrais e o número mediano de consultas para cada paciente foi igual a quatro. A cada consulta a classificação do paciente foi reavaliada. Deste modo esta covariável indicadora do grupo depende do tempo, dado que os pacientes podem mudar de grupo ao longo do estudo.

Fizeram parte do estudo 112 pacientes, sendo 91 pacientes HIV positivo e 21 HIV negativo, dos quais aproximadamente 75% foram censurados. A variável de interesse foi o tempo desde a primeira consulta até a ocorrência de sinusite.

A tabela 1 apresenta as covariáveis utilizadas neste estudo que podem ou não ser consideradas como fator de risco para a ocorrência de sinusite.

É importante citar que para pacientes soronegativos em um determinado tempo, as três variáveis indicadoras que indicam o grupo de classificação assumirão valor zero neste tempo.

Foi verificado que dos pacientes incluídos no estudo 19, aproximadamente 14%, mudaram de grupo de classificação durante o período de acompanhamento. Assim, considerando que o grupo de classificação quanto a infecção pelo HIV pode ser um fator de risco para ocorrência de sinusite foram construídas inicialmente curvas de sobrevivência para os quatro grupos de classificação. Os pacientes foram alocados aos grupos de acordo com a sua classificação no início do acompanhamento. As curvas mostradas na figura 1 referem-se aos pacientes do grupo 1, 2, 3 e 4, respectivamente. Pode-se constatar que os pacientes com AIDS (grupo = 4) tem a menor

TABELA 1: Covariáveis Utilizadas.

Variáveis	Descrição
SEXO	Indica o sexo do paciente, toma o valor 0 se o paciente é do sexo masculino e 1 se é do sexo feminino.
IDADE	Informa a idade do paciente em anos.
ASSINTOMÁTICO	Variável indicadora, toma o valor 1 se o paciente no tempo t é HIV soropositivo assintomático.
ARC	Variável indicadora, toma o valor 1 para pacientes com ARC no tempo t .
AIDS	Variável indicadora, toma o valor 1 para pacientes com AIDS no tempo t .

Fonte dos dados básicos: arquivos da Pesquisa.

sobrevivência, o que significa que estes pacientes têm uma maior probabilidade de adquirir sinusite do que os pacientes dos demais grupos. Os pacientes que fazem parte do grupo HIV soronegativo e HIV soropositivos assintomáticos, grupos 1 e 2 respectivamente, não diferem muito e os que fazem parte do grupo de pacientes com ARC (grupo 3) apresentam uma sobrevivência final aproximada de 0.71. O que significa que a probabilidade de um paciente do grupo 3 sobreviver até o final do estudo sem sinusite é de aproximadamente 71%.

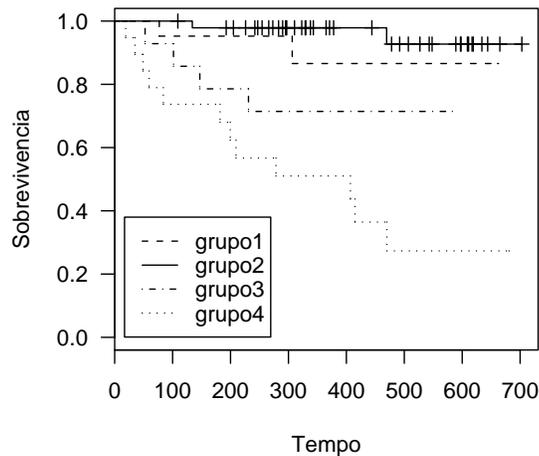


FIGURA 1: Curvas de sobrevivência para os diferentes níveis de infecção pelo HIV.

Para identificar quais covariáveis dentre as pesquisadas influenciam no tempo de desenvolvimento de sinusite foi utilizado inicialmente o modelo de regressão de Aalen. A Tabela 2 apresenta os resultados deste ajuste de onde se pode observar que a covariável “sexo” é não significativa, o que implica que será feito um novo ajuste sem a presença dessa covariável.

As covariáveis idade do paciente e os grupos de riscos foram consideradas como fatores influentes na ocorrência da sinusite (ver tabela 3). A covariável que indica o grupo HIV soropositivo assintomático não apresentou coeficiente significativo mas permaneceu no modelo por representar um dos grupos de classificação quanto

TABELA 2: Resultados do primeiro ajuste do modelo aditivo de Aalen.

Covariável	Coefficiente	Erro Padrão	P-valor	I.C (95%)
constante	1.051	0.415	0.005	(0.238, 1.865)
sexo	0.063	0.195	0.889	(-0.319, 0.445)
idade	-0.033	0.015	0.013	(-0.062, -0.004)
assintomático	0.004	0.140	0.463	(-0.270, 0.278)
arc	0.917	0.371	0.020	(0.189, 1.644)
aids	1.566	0.545	0.000	(0.497, 2.635)

a infecção pelo HIV, portanto o risco de desenvolver sinusite em pacientes HIV soropositivo assintomáticos não difere significativamente ($p = 0.498$) do grupo HIV soronegativo. Pacientes que fazem parte do grupo com AIDS têm um risco maior de desenvolver a sinusite do que os pacientes dos demais grupos de classificação. Comparando-se com o grupo HIV soronegativo este risco é de aproximadamente 1.6 vezes. Verifica-se também, por exemplo, que um aumento de 20 anos na idade do paciente diminui em 0.6 vezes, aproximadamente, o risco de ocorrência da sinusite, o que confirma que quanto maior a idade do paciente menor o risco de desenvolvimento desta doença.

TABELA 3: Resultados do ajuste final do modelo aditivo de Aalen.

Covariável	Coefficiente	Erro Padrão	P-valor	I.C (95%)
constante	1.038	0.405	0.004	(0.244, 1.831)
idade	-0.031	0.013	0.011	(-0.057, -0.005)
assintomático	0.004	0.136	0.498	(-0.263, 0.271)
arc	0.833	0.336	0.002	(0.175, 1.491)
aids	1.544	0.536	0.000	(0.493, 2.595)

Uma vantagem do modelo aditivo de Aalen, conforme já referido, é que ele é capaz de fornecer informações detalhadas a respeito da influência temporal de cada covariável. A figura 2 mostra as funções de regressão acumuladas e os respectivos intervalos, com 95%, de confiança para o intercepto e para as covariáveis significativas no modelo de Aalen. A função de regressão acumulada para a idade tem uma inclinação consistentemente negativa e seu efeito no risco da ocorrência da sinusite diminui razoavelmente com o tempo. Isto indica que crescimentos nos valores da idade, neste período, estão associados com decréscimos na função de risco. A covariável que indica o grupo com ARC parece ter uma influência clara e crescente por cerca de 300 dias com uma influência menor que parece desaparecer depois desse período.

4. Conclusões

O modelo aditivo de Aalen é uma alternativa bastante útil na presença de covariáveis dependentes do tempo pois permite que tanto as covariáveis como os parâmetros variem no tempo fornecendo assim informações detalhadas a respeito

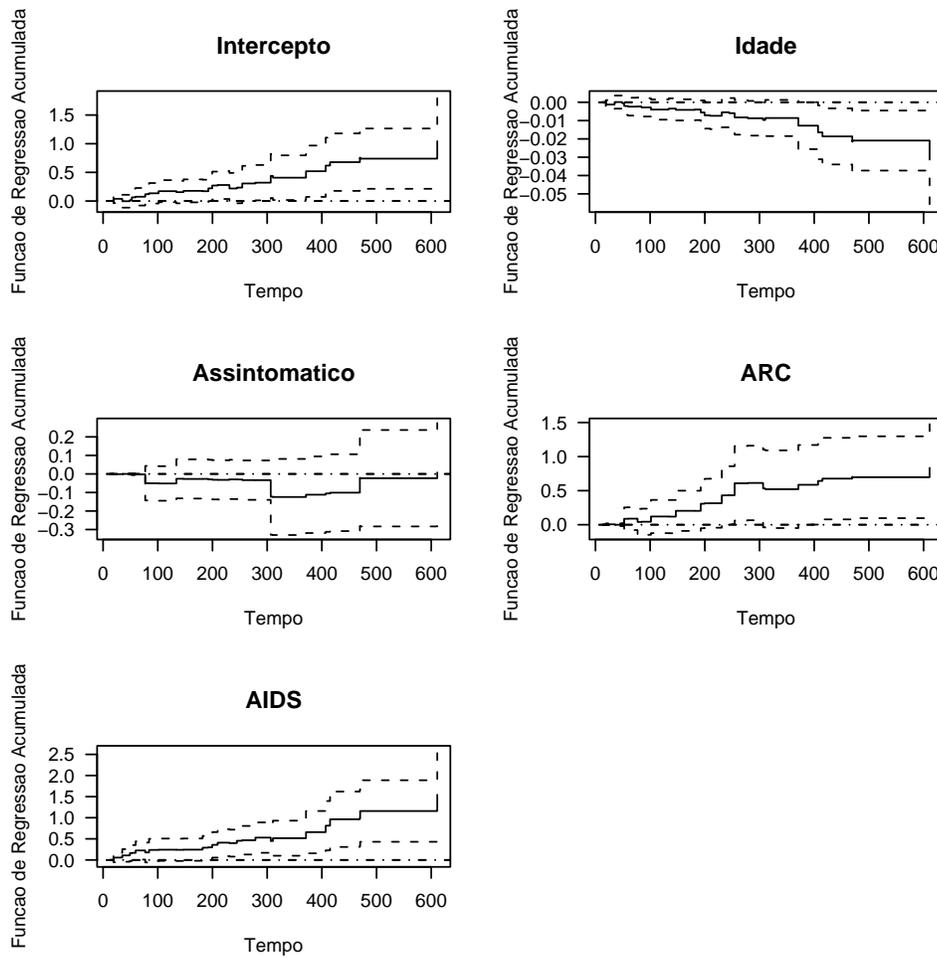


FIGURA 2: Estimativas das funções de regressão acumuladas com intervalo de 95% de confiança.

da influência temporal de cada covariável. Neste trabalho foi apresentado o modelo aditivo de Aalen e uma aplicação envolvendo dados censurados na presença de covariáveis dependentes do tempo.

Foi feita uma aplicação clínica para verificar quais os fatores de risco para o desenvolvimento da sinusite, entre eles a infecção pelo HIV. Foi observado que pacientes com AIDS têm uma maior probabilidade de desenvolver sinusite pois os mesmos apresentaram a menor sobrevivência. Por exemplo pacientes com AIDS tem 27% de risco de não desenvolver a sinusite ao final do estudo enquanto que pacientes HIV soronegativos têm aproximadamente 87%. Com relação ao modelo de Aalen foi verificado que tanto a idade quanto a infecção pelo HIV são fatores de risco para o desenvolvimento da sinusite e que a progressão da imunodeficiência aumenta de forma significativa este risco, ou seja a contribuição no risco para

pacientes com AIDS é maior do que a contribuição dos pacientes HIV soropositivos assintomáticos.

Uma característica importante do modelo de Aalen é que como os parâmetros também variam no tempo é possível construir um gráfico das funções de regressão acumuladas onde é possível verificar o comportamento da influência de cada covariável no tempo. Através do gráfico das funções de regressão acumuladas do modelo de Aalen foi constatado que a influência da idade no risco de desenvolvimento da sinusite diminui com o tempo ao passo que o efeito dos grupos quanto a infecção pelo HIV aumenta com o tempo. Este gráfico representa uma ferramenta importante na análise dos dados pois covariáveis medidas no início do período de observação podem frequentemente perder sua influência no tempo sendo útil ter um método que revela isto.

Agradecimentos

A pesquisa do Enrico A. Colosimo foi parcialmente financiada pelo CNPq (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*) do Ministério de Ciência e Tecnologia do Brasil.

Recibido: noviembre de 2006

Aceptado: abril de 2007

Referências

- Aalen, O. O. (1978), 'Nonparametric Inference for a Family of Counting Processes', *Annals of Statistics* **6**, 701–726.
- Aalen, O. O. (1980), *A Model for Non-parametric Regression Analysis of Counting Processes*, Lecture Notes in Statistics 2, Springer-Verlag, New York. pp. 1-25.
- Aalen, O. O. (1989), 'A Linear Regression Model for the Analysis of Life Times', *Statistics in Medicine* **8**, 907–925.
- Aalen, O. O. (1993), 'Further Results on the Non-Parametric Linear Regression Model in Survival Analysis', *Statistics in Medicine* **12**, 1569–1588.
- Andersen, P. K. & Vaeth, M. (1989), 'Simple Parametric and Non-Parametric Models for Excess and relative Mortality', *Biometrics* **45**, 523–535.
- Cribari-Neto, F. & Zarkos, S. G. (1999), 'R Yet Another Econometric Programming Environment', *Journal Appl. Econ.* **14**, 319–329.
- Gonçalves, D. U. (1995), Incidência, Marcadores de Prognóstico e Fatores de Risco Relacionados às Manifestações Otorrinolaringológicas em Pacientes Infectados pelo HIV, Dissertação de mestrado, UFMG, Faculdade de Medicina, Belo Horizonte.

- Huffer, F. W. & McKeague, I. W. (1987), Survival Analysis Using Additive Risk Models, Technical Report 396, Departamento de Statistics, Stanford University.
- Klein, J. P. & Moeschberger, M. L. (1997), *Survival analysis: Techniques for Censored and Truncated Data*, Springer-Verlag, New York.
- Mau, J. (1986), 'On a Graphical Method for the Detection of Time-Dependent Effects of Covariates in Survival Data', *Applied Statistics* **35**, 245–255.
- Mau, J. (1988), 'A Comparison of Counting Process Models for Complicated Life Histories', *Applied Stochastic Models and Data Analysis* **4**, 283–298.
- McKeague, I. W. (1986), 'Estimation for a Semimartingale Regression Model Using the Method of Sieves', *Annals of Statistics* **14**, 579–589.
- McKeague, I. W. & Utikal, K. J. (1988), Goodness-of-fit Tests for Additive Hazards and Proportional Hazards Models, Technical Report M-793, Departamento de Statistics, Florida State University.
- R Development Core Team (2006), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*<http://www.R-project.org>
- Sample, S., Lenahan, G. A. & Serwonska, M. H. (1989), 'Allergic Diseases and Sinusitis in Acquired Immunodeficiency Syndrome', *Journal Allergy Clin. Immunol.* **83**, 190.