

## ارزیابی مدل‌های خطرات متناسب کاکس و وایل و کاربرد آنها در شناسایی عوامل مؤثر بر زمان بقای بیماران لوسمی حاد

سعید حسینی تشنیزی<sup>۱</sup> دکتر شهرام زارع<sup>۲</sup> دکتر مهدی تذهیبی<sup>۳</sup>

<sup>۱</sup> مربی آمار زیستی، <sup>۲</sup> دانشیار گروه پزشکی اجتماعی، دانشگاه علوم پزشکی هرمزگان <sup>۳</sup> استادیار آمار زیستی، دانشگاه علوم پزشکی اصفهان

مجله پزشکی هرمزگان سال پانزدهم شماره چهارم زمستان ۹۰ صفحات ۲۷۸-۲۶۹

### چکیده

**مقدمه:** مدل خطرات متناسب از مهم‌ترین مدل‌های مورد استفاده در تحلیل داده‌های بقا می‌باشد. بکارگیری این مدل مستلزم برقراری فرض متناسب بودن خطرات می‌باشد و عدم برقراری این فرضیه منجر به استنباط ناصحیح خواهد شد. در این مطالعه، هدف ارزیابی دو مدل خطرات متناسب کاکس و وایل و کاربرد آنها در شناسایی عوامل مؤثر بر بقای بیماران لوسمی حاد می‌باشد.

**روش کار:** در یک مطالعه گذشته‌نگر با مراجعه و انتخاب کلیه پرونده‌های بیماران مبتلا به لوسمی حاد بیمارستان سیدالشهدا اصفهان، اطلاعات ۱۹۷ بیمار بوسیله یک چکلیست جمع‌آوری و با تماس تلفنی از وضعیت نهایی آنها (مرد یا زنده) مطلع شدیم. پس از جمع‌آوری داده‌ها مدل‌های خطرات متناسب کاکس و وایل برای شناسایی عوامل مؤثر بر زمان بقا بکار گرفته شد. روش نیکویی برازش و روش‌های نموداری شامل باقیمانده‌های شوئنفلد و لگاریتم خطر تجمعی برای بررسی فرض متناسب بودن خطرات استفاده و در پایان برای بررسی درستی مدل‌های برازش داده شده از روش نمودار باقیمانده‌های مارتینگل استفاده شد. در این پژوهش از نرم‌افزار S-plus استفاده شد.

**نتایج:** هر دو روش گرافیکی و مبتنی بر آزمون نشان دادند که مفروضه متناسب بودن برای مدل کاکس و وایل برقرار است و از نظر عوامل مؤثر بر زمان بقای بیماران لوسمی حاد، تقریباً نتایج یکسانی بدست آمد. به طوری که بر اساس این مدل‌ها متغیرهای سن، محل سکونت و تعداد گلبول‌های سفید خون، تأثیر معنی‌داری بر زمان بقا نشان دادند ( $P < 0.05$ ) در نهایت ملاک AIC نشان داد مدل نیمه پارامتری کاکس برازش مناسب‌تری بر روی داده‌های زمان بقای بیماران لوسمی حاد دارد.

**نتیجه‌گیری:** بررسی نمودار باقیمانده‌های کاکس-اسنل و مارتینگل نشان داد که در مجموع مدل خطرات متناسب کاکس دارای کارایی بهتری نسبت به مدل وایل برای زمان بقای بیماران لوسمی حاد دارد. با توجه به این که تعداد گلبول‌های سفید خون فاکتور معنی‌دار بر خطر مرگ در این بیماران می‌باشد، با کنترل این متغیر می‌توان بقای این بیماران را افزایش داد.

**کلیدواژه‌ها:** لوسمی، مدل کاکس و مدل وایل - بقا

نویسنده مسئول:  
سعید حسینی تشنیزی  
دانشگاه پرستاری و مامایی - دانشگاه  
علوم پزشکی هرمزگان  
بندرعباس - ایران  
تلفن: ۰۹۸۷۶۱ ۶۶۶۶۳۶۵  
پست الکترونیکی:  
shosseini@hums.ac.ir

دریافت مقاله: ۸۹/۵/۹ اصلاح نهایی: ۹۰/۲/۲۶ پذیرش مقاله: ۹۰/۵/۱۰

**مقدمه:** لمفوبلاستیک (Acute Lymphoblastic Leukemia) یا به اختصار ALL و اگر از رده سلول‌های غیر لمفوبلاستیک باشد به آن لوسمی حاد غیر لمفوبلاستیک یا میلو بلاستیک (Acute Myeloblastic Leukemia) یا به اختصار AML می‌گویند (۱). وقوع تمام لوسمی‌ها تقریباً ۱۳ در ۱۰۰۰۰۰ در سال و وقوع هر دو شکل حاد و مزمن لوسمی تا حدودی در مردان

سرطان خون طبق تعریف، افزایش خود سرانه یک رده سلول‌ساز در مغز استخوان می‌باشد که اگر این تکثیر توأم با وقفه سلول‌ها در جهت تبدیل شدن به اشکال موجود در خون محیطی باشد، به آن سرطان حاد خون یا لوسمی حاد گویند. اگر ضایعه از رده سلول لنفوئیدی باشد، به آن لوسمی حاد

یعنی نسبت خطر برای یک نفر در زمان مشخصی چون  $t$  به خطر برای فرد دیگر در همان زمان، مقدار ثابتی داشته باشد (۷). تاکنون مطالعات زیادی در رابطه با به‌کارگیری مدل رگرسیون کاکس انجام شده ولی بر اساس یک مطالعه سیستمیک تنها در ۵٪ این مطالعات فرض متناسب بودن خطرات مورد بررسی قرار گرفته شده است (۸). در صورت برقراری مفروضات، مدل‌های پارامتریک (وایبل، نمایی و لگ‌نرمال) تجزیه و تحلیل قوی‌تری نسبت به مدل نیمه پارامتری کاکس دارند (۹).

این مطالعه به منظور بررسی فرض متناسب بودن خطرات برای هر کدام از مدل‌های رگرسیون کاکس (ناپارامتری) و وایبل (پارامتری) و سپس برآزش هر کدام از مدل‌های رگرسیون به مشاهدات و مقایسه مدل‌ها و نهایت معرفی مدل رگرسیون مناسب و کارا برای داده‌های زمان بقا بیماران لوسمی حاد طراحی شده است.

#### الف- مدل خطرات متناسب کاکس

این مدل اولین بار توسط D.Cox آمارشناس معروف انگلیسی در سال ۱۹۷۲ به منظور بررسی اثرات متغیرهای توضیحی (مستقل) تأثیرگذار بر زمان بقا ارائه شد و معادله آن به صورت زیر می‌باشد:

$$h_i(t) = h_0(t) \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}) \quad (1)$$

در این مدل، تابع  $h_0(t)$  تابع خطر مبنای زمان  $t$  در این مدل، تابع  $\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik})$  تابع خطر نسبی و  $\beta_i$  ها پارامترهای رگرسیون می‌باشد. مدل رگرسیون کاکس را مدل خطرات متناسب هم می‌نامند و علت آن را این‌گونه بیان می‌کنند که فرض کنید که دو فرد  $i$  و  $j$  دارای مقادیر  $X$  های متفاوت بوده و پیش‌گویی‌های آنها خطی به ترتیب به صورت زیر باشند:

$$\eta_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

و

$$\eta_{i'} = \beta_1 x_{i'1} + \beta_2 x_{i'2} + \dots + \beta_k x_{i'k}$$

بنابر این نسبت خطر برای این دو فرد به شکل زیر است:

$$\frac{h_i(t)}{h_{i'}(t)} = \frac{h_0(t) e^{\eta_i}}{h_0(t) e^{\eta_{i'}}} = \frac{e^{\eta_i}}{e^{\eta_{i'}}} = e^{\eta_i - \eta_{i'}} \quad (2)$$

شایع‌تر از زنان است. ALL اساساً یک بیماری دوران کودکی و بالغین جوان است و تقریباً ۳۰٪ بدخیمی‌های کودکان را تشکیل می‌دهد. در حالی که AML در تمامی سنین دیده می‌شود و میزان شیوع آن در گروه‌های سنی بالاتر بیشتر است، میانگین سن ابتلا ۶۴ سال گزارش شده و شیوع آن ۱/۱ تا ۱/۷ در میلیون نفر می‌باشد (۱).

میزانهای بقاء ۱ تا ۵ ساله بیماران لوسمی حاد با استفاده از مدل مخاطره متناسب کاکس برآورد شد، که مقادیر آن برای ۱ سال، ۲ سال، ۳ سال، ۴ سال و ۵ سال پس از تشخیص به ترتیب ۵۴٪، ۲۷٪، ۲۲٪، ۱۹٪ و ۱۴٪ به دست آمد و سن بیمار و تعداد سلول‌های سفید خون تأثیر معنی داری بر بقای بیماران دارند (۲).

هنگامی که با تحلیل داده‌های بقا سر و کار داریم، دو هدف عمده ما را معطوف می‌سازد. یکی مدل‌بندی برای پیدا کردن ترکیب مناسبی از متغیرهای کمکی که طول بقای اعضای مورد مطالعه را تحت تأثیر قرار می‌دهند و هدف بعدی پیدا کردن برآوردهای مطمئن برای تابع مخاطره در زیرگروه‌های مورد بررسی می‌باشد. به‌طور معمول دو راه کار اساسی در تحلیل داده‌های آماری، روش‌های نیمه پارامتری و پارامتری هستند و هنگامی که در تحلیل بقا با این دیدگاه به بررسی مسئله می‌پردازیم، با توجه به هر کدام از اهداف فوق یکی از راهکارهای ذکر شده از اهمیت بیشتری برخوردار خواهد بود (۳). امروزه به دلیل استفاده روزافزون از تحلیل بقا در مطالعات پزشکی نیاز به مدل‌های کارا و با انعطاف بیشتر برای داده‌های بقا بیش از پیش احساس می‌شود. در برخی از تحقیقات پزشکی با هدف بررسی توزیع بقای بیماران سرطانی از رگرسیون کاکس (Cox) استفاده می‌شود، حال آنکه مدل‌های پارامتری مانند وایبل (Weibull)، نمایی (Exponential) و لگ‌نرمال (Log-Normal) در برخی شرایط می‌توانند جایگزینی مناسب باشند (۶-۴).

به جهت فرضیات کمتر در روش‌های نیمه پارامتری نسبت به روش‌های پارامتری، محققین علوم پزشکی اغلب به استفاده از این روشها راغب‌ترند، ولی پیشنهادهای شده که تحت شرایط خاصی، مدل‌های پارامتری تخمین کاراتری از مدل کاکس دارند (۵). یکی از شرایط بسیار مهم و اساسی به‌کارگیری این مدل‌ها این است که فرض متناسب بودن خطرات برای آنها صادق باشد.

(Hazard) و روش نیکویی برازش استفاده می‌گردد که به اختصار این روشها معرفی می‌گردند (۱۳).

#### ۱- روش باقیمانده‌های شوئنفلد

فرض کنیم  $p$  متغیر توضیحی و  $\Pi$  مشاهده مستقل از زمان و  $\delta_i$  نماد سانسور شدن باشد به طوری که برای داده‌های سانسور نشده مقدار یک و برای داده‌های سانسور شده مقدار صفر را به خود بگیرد.

$i$  امین باقیمانده شوئنفلد برای  $j$  امین متغیر توضیحی یعنی  $x_j$  که در آن  $j = 1, 2, \dots, p$  می‌باشد، به صورت زیر به دست می‌آید:

$$r_{S_{ji}} = \delta_i \{x_{ji} - a_{ji}\} \quad (5)$$

که در آن  $x_{ji}$  مقدار  $j$  امین متغیر توضیحی برای  $i$  امین فرد مورد مطالعه می‌باشد.

$$a_{ji} = \frac{\sum_{l \in R(t_i)} x_j (\exp(\beta' x_l))}{\sum_{l \in R(t_i)} \exp(\beta' x_l)} \quad (6)$$

در رابطه بالا  $R(t_i)$  مجموعه افراد در معرض خطر در زمان  $t_i$  می‌باشد. باید توجه داشت که مقادیر غیر صفر این باقیمانده فقط برای مشاهدات سانسور نشده آورده می‌شود.

باقیمانده‌های شوئنفلد یک برآوردی از  $i$  امین مؤلفه رتبه اثر برای  $j$  امین پارامتر مدل ارائه می‌دهد.

موقعی که مدل خطرات متناسب برقرار باشد نمودار باقیمانده‌های شوئنفلد به صورت تصادفی در طول زمان مطالعه پراکنده می‌باشند و این نمودار در مقابل مقادیر متغیر پیشگوکننده (متغیر توضیحی) مورد نظر دارای شکل و روند مشخصی نمی‌باشد (۱۳، ۱۴).

#### ۲- روش گرافیکی لگاریتم خطر تجمعی

منحنی بقا (log-log) یک منحنی برآوردی است که در نتیجه دو بار لگاریتم‌گیری از تابع بقا برآوردی  $\hat{S}$  حاصل می‌شود و در واقع انتقالی از  $\hat{S}$  می‌باشد و آن را به طور خلاصه به صورت  $Ln(-Ln(s))$  نشان می‌دهند. در

این نسبت خطرات در همه زمانها مقداری ثابت است، به عبارت دیگر به زمان بستگی ندارد و این یکی از مفروضات مهم مدل کاکس است (۱۰).

#### ب- مدل خطرات متناسب وایبل

این مدل علیرغم شباهت بسیار زیادش با مدل خطرات متناسب کاکس، یک تفاوت اساسی با آن دارد. این تفاوت در توزیع تابع خطر پایه  $h_o(t)$  می‌باشد. بدین معنا که در مدل رگرسیون کاکس فرم  $h_o(t)$  نامشخص می‌باشد. یعنی توزیع خاصی برای آن در نظر گرفته نمی‌شود ولی در مدل خطرات متناسب وایبل فرض شده است که زمانهای بقا دارای توزیع وایبل هستند و این فرضیه شکل تابع خطر  $h_o(t)$  را به سمت مدل پارامتری سوق می‌دهد (۱۱).

توزیع وایبل دارای دو پارامتر به نامهای پارامتر شکل  $\alpha$  و پارامتر مقیاس  $\gamma$  می‌باشد. تابع چگالی توزیع وایبل با پارامترهای  $(\alpha, \gamma)$  به صورت زیر است:

$$f_T(t) = \alpha \gamma^{\alpha-1} \exp(-\gamma t^\alpha) \quad (3)$$

$$\alpha > 0, \gamma > 0, t > 0$$

همچنین تابع خطر این توزیع به شکل زیر می‌باشد:

$$h(t) = \alpha \gamma^{\alpha-1} \quad (4)$$

در معادله بالا اگر  $\alpha > 1$  باشد، خطر افزایشی،  $\alpha < 1$  باشد خطر کاهششی و اگر  $\alpha = 1$  باشد خطر ثابت است. بنابراین توزیع وایبل در کلیه موقعیت‌هایی که تابع خطر یکنوا یا ثابت باشد، به کار می‌رود. از طرفی با تغییر پارامترهای این توزیع، چگالی وایبل می‌تواند برای زمانهای چوله به راست و زمانهای چوله به چپ استفاده شود. با توجه به این خواص، توزیع وایبل کاربرد زیادی در تحلیل بقا دارد و در بسیاری از کاربردها، پارامتر شکل این توزیع بستگی زیادی با متغیرهای کمکی ندارد (۱۲).

#### ج- روشهای بررسی متناسب بودن خطرات

برای ارزیابی متناسب بودن خطرات از روشهای باقیمانده‌های شوئنفلد (Schoenfeld residuals method)، روش گرافیکی لگاریتم خطر تجمعی (Log Cumulative

خون، میزان هموگلوبین، تاریخ تشخیص سرطان، وضعیت نهایی (مرده یا زنده) در چک لیست وارد شد. پیگیری بقا یا فوت بیماران تا تاریخ ۸۴/۱۲/۲۹ به‌عنوان زمان شکست (Failure Time) از طریق مراجعه به پرونده یا تماس تلفنی صورت گرفت و زمان بقا بیماران با تقریق تاریخ تشخیص بیماری از تاریخ فوت یا سانسور شده بر حسب هفته محاسبه شد. موارد سانسور شده به افراد زنده تا پایان مطالعه یا افرادی که امکان تماس با آنها نبوده و یا کسانی که به هر علتی غیر از لوسمی حاد فوت کرده بودند، اطلاق شد. پس از ورود اطلاعات در نرم‌افزار آماری S-plus2000 با استفاده از شاخص‌های فراوانی و درصد، میانگین و انحراف معیار، متغیرهای مطالعه توصیف و برای بررسی متناسب بودن خطرات در دو مدل کاکس و وایبل روش نیکویی برازش، روش‌های نموداری شامل باقیمانده‌های شوئنفلد و نمودار لگاریتم خطر جمعی به‌کار گرفته شد و سپس مدل‌های کاکس و وایبل به مشاهدات برازش و در پایان برای بررسی درستی مدل‌های برازش داده شده از روش نمودار باقیمانده‌های مارتینگل استفاده شده است. برای تمام آزمون‌ها  $P < 0.05$  به‌عنوان سطح معنی‌داری در نظر گرفته شد.

### نتایج:

از تعداد ۱۹۷ نفر بیمار لوسمی حاد این مطالعه ۱۰۲ نفر (۵۱/۸٪) مبتلا به ALL و ۹۵ نفر (۴۸/۲٪) AML، از نظر توزیع جنسی تعداد ۱۱۶ نفر (۵۹٪) مرد و بقیه زن، محل سکونت ۱۱۲ نفر (۵۷٪) شهر و ۸۵ نفر (۴۳٪) روستا، از نظر گروه خون ۸۹ نفر (۴۵/۲٪) گروه خون O، ۵۵ نفر (۲۸٪) گروه A، ۴۵ نفر (۲۲/۸٪) گروه خون B و بقیه گروه خون AB بود. میانگین سن در شروع مطالعه برای بیماران گروه AML برابر با  $42.1 \pm 20.83$  سال و برای گروه ALL  $23.1 \pm 20$  سال بدست آمد.

پس از پایان پیگیری مشخص شده تعداد ۱۰۴ نفر (۵۳٪) زنده (سانسور شده) و برای تعداد ۹۳ نفر (۴۷٪) آنها مرگ رخ داده است. میانه مدت زمان بقا برای ۱۹۷ بیمار برابر با ۴۰/۶ هفته که کمترین مدت بقا ۰/۳ هفته و بیشترین آن ۱۵۶/۱ هفته، همچنین میانه مدت زمان بقا برای بیماران گروه All برابر با ۵۸

حالت کلی برای بررسی فرض متناسب بودن خطرات برای دو یا چند گروه زمان بقا می‌توان منحنی لگاریتم خطر جمعی را برای گروه‌های تحت بررسی ترسیم نمود، چنانچه منحنی‌ها با هم موازی باشند، نشان‌دهنده متناسب بودن خطرات است و در غیر این صورت در متناسب بودن خطرات باید شک نمود (۱۵، ۱۶).

### ۳- روش نیکویی برازش

این آزمون در سال ۱۹۹۴ توسط گرامبچ و ترینو (Grambsch and Therneau) در آزمونی بر پایه باقیمانده‌های شوئنفلد پیشنهاد شد (۱۳) که در این مقاله به دلیل پیچیدگی مطالب مربوطه از آوردن عبارات ریاضی آن صرفنظر می‌شود.

### د- باقیمانده‌های مارتینگل

باقیمانده‌های مارتینگل (Martingel Residuals) را به‌صورت زیر تعریف می‌کنند:

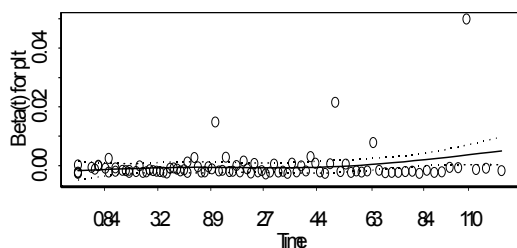
$$\hat{M}_i(t) = N_i(t) - \hat{E}_i(t) \quad (7)$$

که در آن  $N_i(t)$  تعداد پیشامدهائی است که تا زمان  $t$  برای فرد  $i$  رخ داده است و  $\hat{E}_i(t)$  هم خطر جمعی برآورد شده برای فرد  $i$  تا زمان  $t$  می‌باشد. اختلاف بین  $N_i(t)$  و  $\hat{E}_i(t)$  در حقیقت اختلاف واقعی بین تعداد مشاهده شده و تعداد مورد انتظار پیشامدها می‌باشد.

اگر مدلی که به داده‌های زمان بقا برازنده شده است درست باشد، شکل نمودار باقیمانده‌ها یک الگو و روند مشخص را نمایش نمی‌دهد و نقاط روی نمودار کاملاً به‌صورت تصادفی پراکنده خواهند بود (۱۹-۱۷).

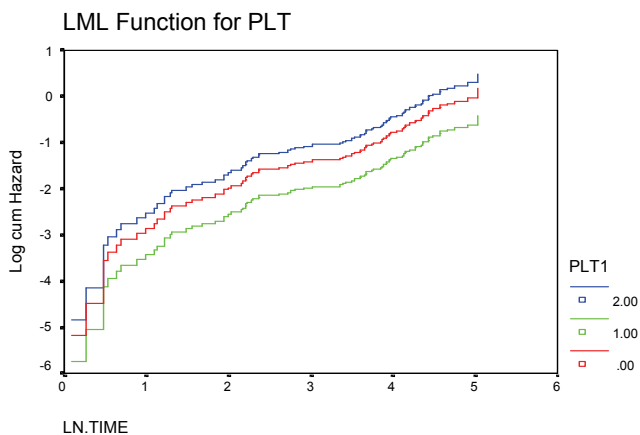
### روش کار:

در یک مطالعه همگروهی گذشته‌نگر که جامعه آماری آن کلیه پرونده‌های بیماران مبتلا به ALL و AML بیمارستان سیدالشهداء (امید) اصفهان از تاریخ ۱۳۸۰/۱/۱ تا ۱۳۸۴/۱۲/۲۹ بود. بصورت تمام شماری، اطلاعات ۱۹۷ بیمار به‌وسیله یک چک‌لیست محقق ساخته شامل متغیرهای: سن، محل سکونت، جنسیت، شغل، گروه خون، نوع لوسمی، تعداد سلول‌های سفید



نمودار شماره ۱- باقیمانده‌های شوئفولد برای متغیر تعداد پلاکت‌های خون

برای بررسی مفروضه متناسب بودن خطرات مدل وایبل روش مبتنی بر آزمون ارئه نشده به همین دلیل در این مطالعه از روش ترسیمی لگاریتم خطر تجمعی (log-log) استفاده شد. پس از ترسیم و تحلیل نمودارها مشخص شد برای همه متغیرها، خطوط مربوط به رده‌ها موازی هم می‌باشند به عبارتی فرض متناسب بودن برقرار است (شکل ۲).



نمودار شماره ۲- لگاریتم خطر تجمعی متغیر تعداد پلاکت‌های خون

به جز متغیر پلاکت در مدل رگرسیون کاکس و برای مدل وایبل بر روی تمامی متغیرها برازش داده شد که نتایج نشان داد در هر دو مدل سه متغیر سن ( $P=0/0352$ )، محل سکونت ( $P=0/0460$ ) و تعداد گلبول‌های سفید خون ( $P=0/0024$ ) بر مدت زمان بقا بیماران تأثیر معنی‌داری دارند و در مدل باقی مانده‌اند.

هفته و برای گروه AML برابر با ۲۸ بدست آمده و از نظر آماری با هم تفاوت معنی‌داری داشتند ( $P=0/000011$ ). به منظور بررسی مفروضه متناسب بودن خطرات مدل کاکس، همزمان از دو شیوه مبتنی آزمون (روش نیکویی برازش) و شیوه‌های گرافیکی باقیمانده‌های شوئفولد استفاده شده که نتایج آزمون نیکویی برازش (جدول شماره ۱) نشان داد فرض متناسب بودن خطرات برای متغیرهای سن، محل سکونت، جنسیت، شغل، گروه خون، نوع لوسمی، تعداد گلبول‌های سفید خون، میزان هموگلوبین برقرار و تنها برای متغیر تعداد پلاکت‌های خون بیماران برقرار نمی‌باشد ( $P=0/019$ ).

جدول شماره ۱- نتایج آزمون نیکویی برازش برای بررسی مفروضه متناسب بودن متغیرهای کمی

متغیر	Rho ( $\rho$ )	Chi-Square	P-value
سن	۰/۰۵۷۲	۰/۰۲۲	۰/۸۱۱
محل سکونت	-۰/۱۲۷	۱/۶۵۸	۰/۱۹۷
جنسیت	-۰/۰۵۵	۰/۳۰۲	۰/۵۸۲
شغل	-۰/۰۷۸	۰/۷	۰/۴۰۲
گروه خون	-۰/۱۷۵	۳/۳۹۹	۰/۰۶۹
نوع لوسمی	۰/۰۱	۰/۰۱۳۸	۰/۹۰۶
Log (wbc)	-۰/۱۴۳	۲/۱۲	۰/۱۴۵
PLT	۰/۲۱۴	۵/۵۰۴	*۰/۰۱۹
HGB	۰/۱۳۸	۱/۵۱۸	۰/۲۱۷

نمودار باقیمانده‌های شوئفولد نیز برای تمامی متغیرهای مذکور ترسیم شد که پس از بررسی دقیق شکل آنها مشخص شد در نمودار متغیر تعداد پلاکت‌ها علاوه بر وجود چند مشاهدت پرت، درصد زیادی از مشاهدات در انتها و قسمت پایین خط پررنگ عبور داده شده از میان آنها قرار دارند و این دلالت بر این دارد که مفروضه متناسب خطرات برای این متغیر برقرار نمی‌باشد (نمودار شماره ۱).

جدول شماره ۲- نتایج آزمون‌های مدل‌های خطرات متناسب کاکس و وایبل

نتایج آزمون رگرسیون وایبل			نتایج آزمون رگرسیون کاکس طبقه‌بندی شده			متغیر
P-value	S.E (β) انحراف استاندارد ضرایب	(β)	P-value	S.E (β) انحراف استاندارد ضرایب	(β)	
۰/۰۰۰۰۰۰۶	۰/۷۳۰	۷/۴۹	-	-	-	مقدار ثابت (α)
۰/۰۲۵۲	۰/۰۰۸	-۰/۰۱۷	۰/۰۳۷	۰/۰۰۵	-۰/۰۱۱	سن
۰/۰۴۶۰	۰/۳۴۶	۰/۶۹۱	۰/۰۴۳	۰/۳۲۸	۰/۴۴	محل سکونت
۰/۰۰۲۴	۰/۱۱۵	-۰/۳۴۸	۰/۰۰۴۲	۰/۰۷۴	-۰/۲۱۴	(WBC)
	۹۱۵			۱۰۷		AIC

- برای متغیر محل سکونت که در آن ۱=روستایی و ۰=شهری می‌باشد، مقدار  $e^{-0/44} = 0/642$  بیانگر این است که خطر مرگ برای افراد روستایی ۰/۶۴۲ برابر افراد شهری می‌باشد. با توجه به اینکه میزان خطر نسبی کمتر از یک می‌باشد این نتیجه را می‌گیریم که خطر مرگ بیماران روستایی کمتر از بیماران شهری است.

- ضریب متغیر (WBC) برابر با  $\hat{\beta} = 0/214$  می‌باشد. لذا مخاطره نسبی فردی که دارای  $WBC = 5000$  است در مقابل فردی دیگری با  $WBC = 1000$  برابر با ۱/۰۷ می‌باشد.

### بحث و نتیجه‌گیری:

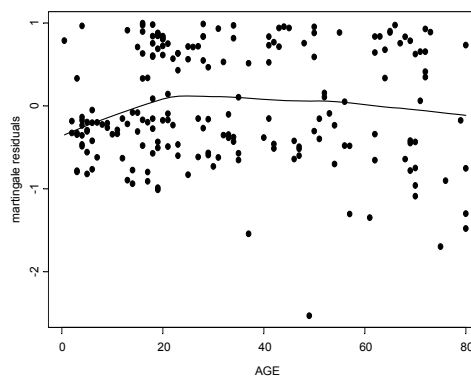
اگر پیش فرض خطرات متناسب برقرار نباشد، نتایجی که از مدل کاکس بدست می‌آید، قابل اعتماد نخواهند بود. در چنین وضعیتی مدل‌های پارامتریک مانند وایبل، نمایی و لگ نرمال انتخاب‌های مناسب‌تری می‌توانند باشند. این مدل‌ها با فرض وجود یک توزیع خاص برای متغیر زمان بقا و بدون نیاز به فرض خطرات متناسب، مدل را برازش می‌دهند (۷). در این مطالعه ابتدا فرض متناسب بودن خطرات برای مدل نیمه پارامتریک کاکس و مدل پارامتریک وایبل بررسی و سپس کارایی این دو مدل بر روی داده‌های زمان بقای بیماران مبتلا به لوسمی حاد مورد بررسی و مقایسه قرار گرفته شد. مفروضه متناسب بودن خطرات کاکس با استفاده از هر دو روش نیکویی برازش و نمودار باقیمانده‌های شونفلد نشان داد که این فرضیه برای همه متغیرها، بجز متغیر تعداد پلاکت‌های خون برقرار است. در ادامه متغیرهایی که فرض متناسب بودن برای آنها برقرار بود به طور مجزا وارد مدل‌های رگرسیون کاکس و وایبل شدند و ملاک AIC نشان داد مدل کاکس از کارایی بهتری نسبت مدل وایبل برخوردار است و همچنین درستی مدل نهایی

هر چند در مدل کاکس برای متغیر سن ضریب  $\beta$  کوچکتر است ولی برای دو متغیر دیگر محل سکونت و WBC بزرگتر و میزان  $S.E(\beta)$  نیز کوچکتر از مدل وایبل است. همچنین مقدار ملاک AIC برای مدل کاکس برابر با ۱۰۷ و برای مدل وایبل برابر با ۹۱۵ به دست آمد که بیانگر کاراتر بودن مدل کاکس نسبت به مدل پارامتریک وایبل می‌باشد (جدول شماره ۲).

درستی مدل رگرسیون کاکس برآزش داده شده به مشاهدات با استفاده از نمودار باقیمانده‌های مارتینگل مورد بررسی قرار گرفته شد (شکل ۳) و مشخص گردید مشاهدات بدون هیچ‌گونه روند و مشاهده پرتی در اطراف خط عبور داده شده از میان آنها قرار گرفته‌اند که از آن به درستی مدل نهایی کاکس که رابطه آن به صورت زیر است، پی می‌بریم:

$$h(t|x) = h_0(t) \exp[-0/01 \text{Age} + 0/44(\text{Residen}) + 0/214(WBC)]$$

- برای متغیر سن، با فرض ثابت نمودن اثر سایر متغیرها، مخاطره نسبی برای بیمار ۳۰ ساله در مقابل بیمار ۳۱ ساله برابر با:  $e^{0/011} = 1/011$  می‌باشد و در حالت کلی با افزایش یکسال به سن بیماران، خطر مرگ آنها به اندازه ۱/۰۱۱ افزایش می‌یابد.



نمودار شماره ۳- باقیمانده‌های مارتینگل برای دو متغیر سن و تعداد گلبول‌های سفید خون

دادند، به این نتیجه رسیدند که سن و WBC عوامل مؤثر بر بقای این بیماران بوده و مدل نهایی دو متغیر مذکور را نیز در بر دارد.

نتایج این مطالعه نشان داد، تأثیر WBC بر مدت زمان بقای بیماران بیشتر از دو متغیر محل سکونت و سن می‌باشد. در مطالعه چن (Chen) و همکاران (۲۴) در تایوان که بر روی ۷۸ بیمار لوسمی حاد انجام گرفت نیز فاکتور  $WBC > 5000$  به صورت معنی‌داری بر مدت زمان بقای بیماران تأثیر دارد. همچنین در مطالعه هم‌گروهی آینده‌نگر هاشمی و همکاران (۲۵) تأثیر فاکتورهای سن، جنسیت، هموگلوبین، نوع لوسمی حاد، LDH، WBC بر روی مدت زمان بقا ۵۶ بیمار لوسمی حاد بررسی شد که از بین تمام فاکتورهای مذکور تنها عامل WBC بر پیش‌بینی زمان بقا تأثیر معنی‌دار داشت.

ناقص بودن پرونده‌های بیماران لوسمی حاد بیمارستان، عدم ثبت برخی از اطلاعات مهم بیماران مانند مرحله پیشرفت بیماری، نوع درمان در پرونده بیماران توسط پزشکان و موارد دیگری چون تغییر شماره تلفن و عدم پاسخ‌گویی برخی از بیماران جهت دریافت وضعیت نهایی بیمار و عدم دسترسی به مطالعات مشابه از جمله محدودیت‌های این مطالعه بودند. همچنین در مطالعات بقا بهتر است که بیش از ۴۰٪ افراد مورد مطالعه سانسور نباشند (۲۶) که در این مطالعه تعداد ۱۰۴ نفر (۵۳٪) موارد سانسور شده (سانسور راست) بودند و این از دیگر محدودیت‌های این مطالعه می‌باشد.

در این مطالعه مشخص شد فرض متناسب بودن هم برای مدل نیمه پارامتریک کاکس و هم مدل پارامتریک وایبل برقرار است که هر چند مدل کاکس از کارایی بهتری نسبت به مدل وایبل برخوردار است ولی در مجموع هر دو مدل‌های خوبی برای برآزش به داده‌های زمان بقا بیماران لوسمی حاد می‌باشند. همچنین با توجه به این که سن بیماران، محل سکونت و تعداد گلبول‌های سفید خون عوامل مؤثر و معنی‌داری بر خطر مرگ در این بیماران بوده‌اند، لذا پیشنهاد می‌شود فاکتور قابل‌کنترلی مانند تعداد گلبول‌های سفید را برای این بیماران تحت کنترل قرار گیرد تا شاید شاهد کاهش مرگ و میر بیماران لوسمی حاد باشیم.

کاکس که شامل سه متغیر تعداد گلبول‌های سفید خون، سن و محل سکونت بود، با استفاده از روش باقیمانده‌های مارتینگل مورد تأیید قرار گرفت. در مدل نهایی سه متغیر مذکور از جمله متغیرهای مهم و تأثیرگذار بر مدت زمان بیماران مبتلا به لوسمی حاد بودند.

در مطالعه‌ای که ناردی و اسکمپر (۲۰) با استفاده از روش‌های باقیمانده‌ها مدل کاکس و مدل‌های پارامتریک جایگزین را در سه مطالعه بالینی مقایسه کردند، به این نتیجه رسیدند که مدل پارامتریک وایبل بر اساس معیار ضریب تغییرات پارامترهای برآورد شده از سایر مدل‌ها بهتر بود. همچنین پورحسینقلی و همکاران (۲۱) در مطالعه‌ای به منظور مقایسه مدل کاکس و مدل‌های پارامتریک بر روی داده‌های ۷۴۶ بیمار مبتلا به سرطان معده با بکارگیری ملاک آکائیکه نشان دادند، رگرسیون نمایی (در آنالیز چند متغیری) و رگرسیون لگ نرمال (در آنالیز تک متغیری) از سایر مدل‌ها کارتر می‌باشند و نتایجی هم که در این مطالعه بدست آمد، نشان داد فرض متناسب بودن هم برای مدل کاکس و هم برای مدل وایبل برقرار بوده که نه تنها با نتایج مطالعه پورحسینقلی (۲۱) و با نتایج مطالعه ناردی و اسکمپر نیز همسو می‌باشد.

اربه، فریرا و نونز آنتون (۲۲) در یک مطالعه شبیه‌سازی شده، رگرسیون کاکس و مدل زمان شکست شتابیده را مقایسه کردند. نتایج این مطالعه نشان داد چه در وضعیتی که فرض خطرات متناسب برقرار است و چه در حالتی که این فرض برقرار نیست، مدل‌های پارامتریک لگ لجستیک، لگ نرمال و مدل استوت از کاکس کارتر هستند و در مطالعه رجایی‌فرد و همکاران (۸) به این نتیجه رسیدند که اگرچه نسبت مخاطره در مدل کاکس و مدل‌های پارامتری وایبل و نمایی تقریباً یکسان است ولی شاخص AIC برای دو مدل وایبل و نمایی نشان از این دارد که این دو مدل برآزش بهتری بر روی داده‌های زمان بقا بیماران سرطان معده دارند.

در مطالعه فعلی، سه فاکتور سن، زمان تشخیص سرطان، محل سکونت و WBC عوامل مؤثر بر مدت زمان بقا بیماران لوسمی حاد می‌باشد و بنابراین مدل رگرسیون کاکس نهایی شامل این سه متغیر می‌باشد. در مطالعه آینده‌نگری هم که چسلس و همکاران (۲۳) بر روی ۳۷۸ بیمار لوسمی حاد انجام

## سپاسگزاری:

در خاتمه لازم می‌دانیم از کارکنان قسمت بایگانی بیمارستان سیدالشهداء اصفهان که در گردآوری اطلاعات این مطالعه ما را یاری نمودند، تشکر و قدردانی نماییم.

## References

## منابع

1. Alavi S, Arzanian MT, Moradmand M, Ashraftalesh H. Prevalence of prognostic factors in children with Acute Lymphoblastic Leukemia admitted to Mofid Children's. *Iranian Journal of Pediatrics*. 2005;15:237-242. [Persian]
2. Bahrami M, Moshkani MR, Alam Samimi M. Effective Factors on Survival Time of the leukemic Patients and Estimating the Mean of Survival Time. *Journal of Isfahan Medical School*. 2007;24:49-57. [Persian]
3. Klein JP, Moeschberger ML. Survival analysis: Techniques for censored and truncated data. 3<sup>rd</sup> ed. New York: Springer-Verlag Press; 1997.
4. Efron B. The efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association*. 1972;72:557-565.
5. Oakes D. Comparison of models for survival data. *Statistics in Medicine*. 1983;2:305-311.
6. Ngandu NH. An empirical comparison of statistical tests for assessing the proportional Hazards Assumption of Cox's Model. *Statistics in Medicine*. 1997;16:611-626.
7. Hogaard P. Analysis of multivariate survival data. Second ed. New York: Springer-verlag Press; 2000.
8. Rajaefard A, Moghimi dehkordi B, Tabatabaee HR, Zeighami B, Safaee A, Pourhoseingholi M, et al. Application of parametric models in survival analysis of gastric cancer. *Feyze Journal*. 2010;13:83-88. [Persian]
9. Saneie H. Analysis of survival data. First ed. Tehran: Andishmand Press; 2001. [Persian]
10. Cox DR, Oakes D. Analysis of survival data. Second ed. London: Chapman and Hall Press; 1984.
11. Hogaard p. Fundamentals of survival data, Biometrics. First ed. New York: Springer-verlag Press; 1999.
12. Fleming Tr, Lin Dy. Survival analysis in clinical trial: Past development and future direction. *Biometrics*. 2000;56:971-983.
13. Kimber AC. A weibull-based score test for heterogeneity lifetime data analysis. *Biometrics*. 1996;2:63-71.
14. Sarker SJ. Test for weibull based proportional hazard frailty models. PhD thesis, Surrey University: 2002:31-38.
15. Kleinbum DG. Survival analysis: A Self-Learning Text, pringer rerlag. First ed. New York: Wiley Press; 1996.
16. Hosmer DW, Lemeshow S. Applied Survival analysis second ed. New York: Wiley Press; 1999.
17. Fox J. Cox Proportional-Hazards Regression for Survival Data. 2010 Jan. Available form: URL:<http://socserv.mcmaster.ca/jfox/courses/soc761/survival-analysis.pdf>.
18. Lawless JF. Parametric models in survival analysis. 4<sup>th</sup> ed. In Encyclopedia of Biostatistics. New York: Wiley Press; 1970.
19. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. first ed. New York Wiley Press; 1980.
20. Nardi A, Schemper M. Comparing Cox and parametric models in clinical studies. *Stat Med*. 2003;22:3597-3610.



21. Purhosseinigholi M. Comparison of Cox regression and parametric models in survival analysis of patients Gastric cancer. *Iranian Journal of Epidemiology*. 2007;3:25-29.
22. HRbe J, Ferreira E, Nunez-Anton V. Comparing proportional hazards and accelerated failure time models for survival analysis. *Statist Med*. 2002;21:493-510.
23. Chessells JM, Hardisty RM, Richards S. Long survival in childhood lymphoblastic leukemia. *Br J Cancer*. 1987;55:315-319.
24. Chen BW, Lin DT, Lin KH, Chuu WM, Su S, Lin KS. An analysis of risk factor and survival in childhood acute lymphoblastic leukemia. *Zhonghua Min Guo-xiao Er Ke Yi Xue Hui Za Zhi*. 1989;30:299-308.
25. Hashemi AS, Manochehri Naieni Z, Eslami A, Bahrami Ahmadi M, Kheirandish M, Rafieyan M. Effective factors on survival of children with acute lymphoblastic leukemia referred Shahid Sadughi Hospital of Yazd. *Journal of Shahid Sadoughi University of Medical Sciences & Health Services*. 2008;16:14-19. [Persian]
26. Bolstad W, Manda OS. Investigating child mortality in Mallawi using family and community random effects: Abayesian analysis. *Journal of the American Statistical Association*. 2001;96:12-19.

## The evaluation of Cox and Weibull proportional hazards models and their applications to identify factors influencing survival time in acute leukemia

S. Hosseini Teshnizi, MSc<sup>1</sup>   S. Zare, PhD<sup>2</sup>   M. Tazhibi, PhD<sup>3</sup>

Instructor Of Biostatistics of Premedical<sup>1</sup>, Associate Professor Department of Community Medicine<sup>2</sup>, Hormozgan University of Medical Sciences, Bandar Abbas, Iran. Assistant Professor Department of Biostatistics & Epidemiology<sup>3</sup>, Isfahan University of Medical Sciences, Isfahan, Iran.

(Received 31 Jul, 2010   Accepted 1 Aug, 2011)

### ABSTRACT

**Introduction:** The most important models used in analysis of survival data is proportional hazards models. Applying this model requires establishment of the relevance proportional hazards assumption, otherwise it would lead to incorrect inference. This study aims to evaluate Cox and Weibull models which are used in identification of effective factors on survival time in acute leukemia.

**Methods:** In a retrospective study the information on 197 patients with acute leukemia were collected. The proportional hazards models were used to identify effective factors on survival time. Goodness-of-fit test and two graphical procedures based on the scaled Schoenfeld residuals were used to evaluate the proportional hazards assumption of the Cox and Weibull models. Finally, to assess the accuracy of fitted models, Martingel residuals graphic method was applied for goodness of fit investigation. S-plus statistical package was used for analysis.

**Results:** The results showed that the assumptions of proportional hazards for Cox and Weibull models are correct. Based on the models, variables such as, age, residence place, and WBC, showed significant effect on survival time ( $P < 0.05$ ). Finally, AIC criteria showed that Cox semi- parametric model had better fit on the survival data of the patients with acute leukemia.

**Conclusion:** Assessing the Cox-snell and Martingel residuals plots showed that Cox model is more efficient than Weibull model for the survival time of the patients with acute leukemia. Significance of WBC factor in the model, indicate that control of this factor is effective in increasing patients' survival.

*Correspondence:*

S. Hosseini Teshnizi, MSc.  
Nursing Midwifery School.  
Hormozgan University of  
Medical Sciences.  
Bandar Abbas, Iran  
Tel: +98 761 6666365  
Email:  
shosseini@hums.ac.ir

**Key words:** Leukemia - Proportional Hazards Models - Survival