

کومش - جلد ۲۰، شماره ۱ (پیاپی ۶۹)، زمستان ۱۳۹۶

## تحلیل بقای پارامتری و نیمه پارامتری بیماران مبتلا به سرطان ریه

نگار قاسمی<sup>۱</sup> (M.Sc.)، نامعلی آزادی<sup>۱</sup> (Ph.D.)، مژگان پورمختار<sup>۲</sup> (Ph.D.)، جمیله ابوالقاسمی<sup>۱</sup> (Ph.D.)، آسیه عشوری<sup>۳</sup> (Ph.D.)، توفیق مبادری<sup>۱</sup> (M.Sc.)، محبوبه رسولی<sup>۱\*</sup> (Ph.D.)

۱- گروه آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی ایران، تهران، ایران

۲- مرکز تحقیقات انتقال خون، موسسه عالی آموزشی و پژوهشی طب انتقال خون، تهران، ایران

۳- گروه آموزش بهداشت و ارتقاء سلامت، دانشکده بهداشت، دانشگاه علوم پزشکی گیلان، رشت، ایران

### چکیده

هدف: سرطان ریه یکی از شایع ترین سرطان‌ها در جهان و جز پنج سرطان شایع و سومین عامل مرگ در ایران است. در این مطالعه جهت بررسی عوامل موثر بر بقای بیماران سرطان ریه از مدل‌های نیمه پارامتری کاکس و پارامتری وایبل استفاده و کارایی آن‌ها مقایسه شد.

مواد و روش‌ها: در این مطالعه توصیفی-تحلیلی که از نوع طولی بود، ۲۲۸ بیمار مبتلا به سرطان ریه طی سال‌های ۱۳۷۰ تا ۱۳۸۵ پیگیری و اطلاعات لازم از طریق مرکز ثبت سرطان بابل جمع‌آوری شد. سن هنگام تشخیص بیماری، جنسیت، محل سکونت، تحصیلات، وضعیت اقامت، استعمال دخانیات، سابقه خانوادگی و استان محل زندگی به عنوان عوامل دموگرافیک به مدل‌های مخاطرات متناسب کاکس و وایبل اضافه شد و سپس کارایی دو مدل با استفاده از معیار اطلاع آکائیکه مقایسه گردید.

یافته‌ها: در این مطالعه که ۷۵ درصد از نمونه‌های آن مرد بودند، میزان بقای یک، سه و پنج ساله به ترتیب ۱۳، ۸ و ۸ درصد بود و از عوامل بررسی شده در مدل وایبل، اثر استعمال دخانیات بر زمان بقا معنی‌دار بود ( $P < 0.05$ ) ولی در مدل کاکس اثر هیچ یک از عوامل معنی‌دار نبود ( $P > 0.05$ ). بر اساس ملاک آکائیکه، کارایی مدل وایبل بهتر از مدل کاکس بود.

نتیجه‌گیری: علی‌رغم تمایل بسیاری از محققین به استفاده از مدل کاکس در تحلیل داده‌های بقا، نتایج این مطالعه نشان داد مدل وایبل در تحلیل داده‌های بقای سرطان ریه کارا تر است. با توجه به نتایج این تحقیق، پیشگیری از استعمال دخانیات جهت افزایش طول عمر بیماران سرطان ریه ضروری به نظر می‌رسد.

واژه‌های کلیدی: تحلیل بقا، مدل‌های آماری، سرطان‌های ریه، استعمال دخانیات

### مقدمه

سراسر دنیا، سرطان ریه از جمله مهم‌ترین سرطان‌های تشخیص داده شده می‌باشد و به عنوان یکی از محتمل‌ترین علت مرگ و میرهای ناشی از سرطان به شمار می‌رود [۴]. سرطان ریه در ایران جز یکی از پنج سرطان شایع است و میزان بروز آن روندی افزایشی دارد [۵]. دو نوع اصلی سرطان ریه وجود دارد: سرطان ریه سلول کوچک و سرطان ریه

سرطان پس از بیماری‌های قلبی-عروقی اولین عامل مرگ و میر در سراسر جهان می‌باشد [۱] و به عنوان یک مشکل سلامت عمومی در تمام جوامع مطرح می‌باشد [۲]. در ایران نیز به عنوان سومین عامل مرگ و میر بعد از بیماری‌های قلبی-عروقی و حوادث جاده‌ای محسوب می‌شود [۳]. در

سلول غیر کوچک. در صورتی که سرطانی ویژگی‌های هر دو نوع را داشته باشد به عنوان یک کارسینوم سلول کوچک و بزرگ (آمیخته) شناخته می‌شود. حدود ۹۰-۸۵ درصد سرطان‌های ریه تشخیص داده شده از نوع سرطان سلول غیر کوچک هستند [۶]. حدود ۱۲/۴ درصد موارد جدید سرطان تشخیص داده شده و ۱۷/۶ درصد از مرگ و میرهای ناشی از سرطان‌ها در جهان مربوط به سرطان ریه می‌باشد [۷]. سرطان ریه عمدتاً در افراد سالمند رخ می‌دهد به طوری که تقریباً ۷۰ درصد از موارد تشخیص داده شده سرطان ریه در سن ۶۵ سال یا بیش‌تر و کم‌تر از ۳ درصد مبتلایان کم‌تر از ۴۵ سال سن دارند. شانس متوسط ابتلای مردان و زنان به سرطان ریه در طول زندگی آن‌ها به ترتیب برابر با ۶/۲۵ و ۷/۱۴ درصد می‌باشد [۶]. برخی عوامل محیطی و رفتاری در ابتلا به سرطان ریه نقش دارند اما مطالعات نشان می‌دهند که تاکنون سیگار به عنوان یکی از ریسک فاکتورهای بسیار مهم در ابتلا به سرطان ریه بوده است [۵]. علاوه بر این مواجهه با برخی ترکیبات صنعتی از جمله رادون، آزبستوز، امواج رادیو اکتیو مثل اورانیوم، مواد شیمیایی و معدنی مثل آرسنیک، ترکیبات حاوی نیکل و کروم، رژیم غذایی و ویتامین‌ها، سابقه خانوادگی سرطان ریه، آلودگی هوا یا آلاینده‌های محیطی، مصرف ماری جوانا، تغییرات DNA و ژن در افزایش خطر ابتلا به کارسینوم اولیه ریه موثر هستند [۶]. علائم سرطان ریه پس از گذشت چندین سال بروز می‌کنند و همین امر باعث می‌شود این بیماری در مراحل پیشرفته تشخیص داده شود [۸]. یکی از مهم‌ترین راه‌کارها جهت کنترل بیماری و ارزیابی روش‌های درمانی، تعیین میزان بقای بیماران سرطان ریه پس از تشخیص و بررسی عوامل موثر بر میزان بقای این بیماران می‌باشد [۹]. روش‌های آماری جهت بررسی توزیع بقای بیماران سرطانی و اختلاف‌های آن شامل مدل‌های پارامتری و ناپارامتری می‌باشد. مدل خطرات متناسب کاکس به عنوان یک مدل نیمه پارامتری و مدل‌های زمان شکست شتابیده به عنوان مدل‌های پارامتری دو مدل رگرسیونی برای بررسی داده‌های سانسور راست می‌باشند [۱۰].

اصطلاح سانسورشدگی در داده‌های بقا بسیار رایج می‌باشد در برخی مواقع، قبل از پایان مطالعه، پیشامد مورد نظر برای برخی از بیماران رخ نمی‌دهد و یا بنا به دلایل دیگری غیر از بیماری مورد نظر، از مطالعه خارج می‌شوند لذا زمان دقیق رخداد پیشامد این افراد مشخص نیست. سانسورشدگی به سه دسته‌ی عمده تقسیم می‌شود: سانسور شده از راست، سانسور شده از چپ، سانسور فاصله‌ای. سانسور شده از راست زمانی رخ می‌دهد که زمان دقیق رخداد پیشامد مورد نظر بزرگ‌تر یا مساوی با زمان مشاهده شده بقا می‌باشد. در حالتی که زمان بقای افراد در سمت چپ دوره پیگیری ناتمام باشد سانسور شدگی چپ به وقوع می‌پیوندد. سانسور فاصله‌ای از انواع پیچیده سانسورشدگی می‌باشد و در حالتی که زمان دقیق رخداد پیشامد مشخص نیست اتفاق می‌افتد [۱۱]. مدل خطرات متناسب کاکس متداول‌ترین روش برای تجزیه و تحلیل داده‌های بقا در تحقیقات پزشکی جهت بررسی توزیع بقای بیماران سرطانی بر اساس گروه‌بندی‌های دموگرافیک و کلینیکی می‌باشد [۹، ۱۰]. اکثر محققین در حیطه علوم پزشکی از مدل نیمه پارامتری کاکس به دلیل وجود فرضیات کم‌تر در این مدل نسبت به مدل‌های پارامتری استفاده می‌کنند. این مدل کاربردی‌ترین ابزار مدل‌سازی در مطالعات بقا محسوب می‌شود [۱۰، ۱۲]. که بر فرض متناسب بودن خطرات مبتنی است و ممکن است در بیش‌تر داده‌ها این فرض برقرار نباشد و در برخی شرایط مدل‌های پارامتری به جهت این‌که یک توزیع احتمال فرضی را برای زمان‌های بقا در نظر می‌گیرند تجزیه و تحلیل قوی‌تر و استنباط آماری دقیق‌تری نسبت به زمانی که هیچ توزیعی برای زمان‌های بقا نداریم ارائه می‌دهند [۱۳]. در مطالعه‌ی ناردی نیکویی برآزش مدل خطرات متناسب کاکس و مدل‌های پارامتری بررسی شد که نتایج نشان داد مدل پارامتری وایبل از سایر مدل‌ها بهتر بود [۱۲]. در مطالعه قدیمی و همکاران مدل لگ لجستیک با شکنندگی و بدون شکنندگی از سایر مدل‌ها بهتر بود [۱۴]. علیزاده و همکاران نیکویی برآزش مدل خطرات متناسب کاکس و مدل‌های پارامتری را بررسی کردند در این مطالعه

بیماران و سایر منابع به دست آمد و محرمانگی اطلاعات در این مطالعه حفظ شد.

برای بررسی اثر متغیرهای مستقل بر روی زمان بقای بیماران از مدل خطرات متناسب کاکس و مدل پارامتری وایبل استفاده شد. با فرض اینکه  $x$  متغیر مستقل است تابع بقای متناسب با مدل کاکس به صورت زیر است:

$$S(t;x) = \exp\left[-\exp(X\beta) \int_0^t \lambda_0(u) du\right]$$

مدل کاکس علی‌رغم محدودیت‌هایی که دارد، کاربردی‌ترین مدل برای یافتن ارتباط متغیرهای مستقل با متغیر پاسخ بقا که از راست سانسور شده‌اند. مدل پایه بقای وایبل به صورت زیر است:

$$S(t;x, \beta) = \exp\left[\frac{-t^p}{\exp(\beta_0 + \beta_1 x)}\right]$$

مدل‌های پارامتری مانند مدل وایبل به پیش‌فرض‌های بیشتری نیاز دارند که اگر این پیش‌فرض‌ها رعایت شوند، نسبت به مدل کاکس کارا تر خواهند بود.

جهت بررسی فرضیه متناسب بودن خطرات، برای مدل کاکس از آزمون نیکویی برازش گرامبچ و ترینسو و برای بررسی این فرضیات در مدل پارامتری وایبل چون تا به حال آزمونی معرفی نشده است، از روش‌های گرافیکی استفاده شد. برای مقایسه کارایی مدل کاکس و وایبل از ملاک آکائیکه (Akaike Information Criteria) استفاده شده است. AIC معیاری است که به وسیله آکائیکه (۱۹۷۴) پیشنهاد شد و هدف آن اندازه‌گیری نیکویی برازش مدل است [۱۸، ۱۷]. AIC معیاری است که میزان تعامل بین پیچیدگی مدل و برازش مناسب مدل را اندازه‌گیری می‌کند و اندازه آن هر چقدر کم‌تر باشد بهتر است. برای مدل‌های مورد استفاده در این پژوهش AIC از فرمول زیر محاسبه شده است:

$$AIC = -2 * \log(\text{likelihood}) + 2 * (p+k)$$

که  $P$  تعداد پارامترهای موجود در مدل است و  $K$  ضریب ثابتی است که وابسته به نوع مدل به کار رفته می‌باشد. که برای مدل وایبل  $K=2$  است [۱۸، ۱۱]. هر چقدر مقدار آکائیکه کم‌تر باشد کارایی مدل بهتر است. تمامی تحلیل‌ها با نرم‌افزار stata انجام شد و سطح معنی‌داری ۰/۰۵ در نظر گرفته شد.

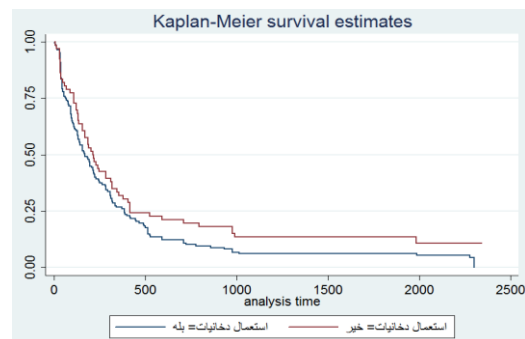
مدل پارامتری در حالت چندمتغیره و تک‌متغیره کارایی بالاتری نسبت به مدل کاکس داشتند [۱۵]. مدل وایبل به عنوان یک مدل پارامتری بقا به دلیل داشتن تابع بقای نسبتاً ساده و همچنین سادگی در محاسبات ریاضی در مقالات عمومیت بیشتر دارد. این مدل دارای این ویژگی است که اگر فرض مدل شکست شتابنده (AFT) برقرار باشد در این صورت فرض مدل خطرات متناسب (PH) نیز برقرار می‌شود و بالعکس. این ویژگی در مدل وایبل منحصر به فرد است [۱۶، ۹]. مدل وایبل همچنین انعطاف‌پذیری بیشتری نسبت به مدل نیمه پارامتری کاکس دارد [۱۷]. در مدل وایبل مخاطره‌ها می‌توانند افزایش، کاهش و یا ثابت باشند و این بستگی به مقدار پارامتر شکل دارد [۹]. هدف از این مطالعه بررسی عوامل موثر بر بقای بیماران سرطان ریه در استان مازندران با استفاده از مدل‌های نیمه پارامتری کاکس و پارامتری وایبل و مقایسه کارایی آن‌ها می‌باشد.

## مواد و روش‌ها

در این مطالعه توصیفی-تحلیلی که از نوع طولی بود، نمونه‌گیری به صورت در دسترس شامل اطلاعات ۲۲۸ بیمار مبتلا به سرطان ریه است که بین سال‌های ۱۳۶۹ تا ۱۳۷۰ به مرکز ثبت سرطان بابل مراجعه کردند و مشخصات فردی آن‌ها ثبت گردید. بیماران به مدت ۱۵ سال از سال ۱۳۷۰ تا ۱۳۸۵ مورد پیگیری قرار گرفتند. مشخصات فردی، بالینی و زمان بقا از طریق پرونده‌های بیماران موجود در مرکز تحقیقات استخراج و جمع‌آوری شد. بیمارانی وارد مطالعه شدند که اطلاعات مندرج در پرونده‌ی آن‌ها خوانا بود. مشخصه‌های فردی که در این مطالعه مورد بررسی قرار گرفته شامل جنسیت (زن، مرد)، سن در زمان تشخیص، محل اقامت (شهر، روستا)، تحصیلات (باسواد، بی‌سواد)، استعمال دخانیات (بلی، خیر)، سابقه خانوادگی سرطان (بلی، خیر) و استان محل زندگی (گلستان، مازندران) بودند. اطلاعات مربوط به بقای بیماران نیز به صورت تلفنی، پرسش‌نامه، مراجعه به منازل

## نتایج

نمونه مورد مطالعه شامل ۲۲۸ بیمار مبتلا به سرطان ریه می‌باشد که از این تعداد ۵۶ نفر (۲۴/۶ درصد) را زنان و ۱۷۲ نفر (۷۵/۴ درصد) را مردان تشکیل می‌دهند. جدول ۱ مشخصات جمعیت‌شناختی نمونه‌ها را به تفکیک جنسیت نشان می‌دهد. در این جدول ارتباط متغیرهای مستقل با جنسیت نمونه بررسی شد که نتایج نشان داد بین سطح تحصیلات و جنسیت ارتباط معنی‌داری وجود دارد ( $P=0/001$ ) ( $value <$  به طوری که ۶۴/۳ درصد از زنان مورد مطالعه باسواد بودند و این در حالی است که فقط ۳۳/۷ درصد از مردان باسواد بودند. میانگین (انحراف معیار) سنی بیماران مبتلا به سرطان ریه در زمان تشخیص ۵۵/۵۱ (۱۳/۵۱) سال بود که در مردان ۵۵/۴۴ (۱۲/۹۲) سال و برای زنان ۵۵/۷۱ (۱۵/۳۲) سال بوده است. همچنین ۵۶/۶ درصد از نمونه‌ها شهری و بقیه روستایی، ۷۸/۵ درصد مازندرانی و ۲۱/۵ درصد اهل گلستان، ۶۵/۸ درصد سابقه خانوادگی سرطان ریه، ۷۱/۱ درصد مصرف‌کننده دخانیات و ۸۶/۸ درصد از آن‌ها بومی بودند. شکل ۱ نمودار بقا را به تفکیک استعمال و عدم استعمال دخانیات نشان می‌دهد، همان‌گونه که از این نمودار مشاهده می‌شود میزان بقای گروه عدم استعمال دخانیات، اندکی بیشتر از گروه استعمال‌کنندگان است ولی با توجه به نتیجه‌ی آزمون لگ-رتبه این اختلاف معنی‌دار نبود ( $P-value = 0/091$ ).



شکل ۱. نمودار بقای کاپلان-مایر بر اساس استعمال دخانیات

ریه ۶/۲ ماه به دست آمد. در طول دوره مطالعه ۲۱۰ (۹۲ درصد) مرگ و میر رخ داده است که این میزان برای مردان ۹۲ درصد و برای زنان ۹۱ درصد بوده است. ۱۸ نفر (۸ درصد) زنده مانده‌اند و به عنوان مشاهدات سانسور شده از راست در نظر گرفته شدند. نتایج آزمون نیکویی برازش گرامیچ و ترینو در جدول شماره ۲ نشان می‌دهد که مقدار احتمال برای هیچ یک از متغیرها معنی‌دار نشد، لذا فرضیه خطرات متناسب برای مدل کاکس برقرار بود. جهت بررسی فرضیه خطرات متناسب برای مدل وایبل از ترسیم لگاریتم خطر تجمعی در مقابل لگاریتم زمان بقا استفاده شد و نتایج نشان داد برای همه متغیرهای مستقل، خطوط مربوط به رده‌ها موازی هم می‌باشند، به عبارتی فرض متناسب بودن خطرات برقرار بود. جداول ۳ و ۴ نتایج حاصل از برازش مدل‌های تک متغیری و چند متغیری را نشان می‌دهد. این جداول نشان می‌دهد که مقدار ملاک آکائیکه مدل کاکس هم در آنالیز تک متغیری و هم در آنالیز چند متغیری بیش از ۲ برابر مقدار آن در مدل وایبل است که نشان از کارآتر بودن مدل وایبل نسبت به کاکس می‌باشد. هر چند میزان مخاطره در مدل وایبل تقریباً مشابه نتایج به دست آمده از مدل کاکس بود. نتایج حاصل از مدل وایبل نشان می‌دهد هم در آنالیز تک متغیری و هم در آنالیز چند متغیری یک اختلاف معنی‌داری را در بین افرادی که دخانیات مصرف می‌کنند وجود دارد ( $P-value < 0/05$ ). در این مدل مقدار نسبت خطر برای استعمال دخانیات در آنالیز تک متغیری ۰/۷۲۵ (چند متغیری ۰/۷۱۳) محاسبه شد که نشان می‌دهد زمان بقا برای استعمال‌کنندگان نسبت به کسانی که دخانیات مصرف نمی‌کنند ۰/۷۲۵ برابر کم‌تر است. اما نتایج حاصل از برازش مدل کاکس نه در آنالیز تک متغیری نه در آنالیز چند متغیری اثر معنی‌داری برای استعمال دخانیات نشان نداد. همچنین هیچ یک از مدل‌ها در آنالیز تک متغیری و چند متغیری اثر معنی‌داری برای سن، محل سکونت، میزان تحصیلات، نوع استان، سابقه فامیلی سرطان، وضعیت اقامت نشان ندادند.

میزان بقای یک، سه و پنج ساله بیماران نیز به ترتیب ۱۳، ۸ و ۸ درصد می‌باشد. میانه بقا برای بیماران مبتلا به سرطان

جدول ۱. توزیع فراوانی مشخصه های بیماران مبتلا به سرطان ریه به تفکیک جنسیت

P-value	کل بیماران (n=۲۲۸)	زن (n=۵۶)	مرد (n = ۱۷۲)*	
۰/۸۹۶	۵۵/۵۱(۱۳/۵۱)	۵۵/۷۱(۱۵/۳۲)	۵۵/۴۴(۱۲/۹۱)*	سن در زمان تشخیص(سال)
۰/۱۸۰				محل سکونت
	۱۲۹(۵۶/۶)	۳۶(۶۴/۳)	۹۳(۵۴/۱)	شهری
	۹۹(۴۳/۴)	۲۰(۳۵/۷)	۷۹(۴۵/۹)	روستایی
۰/۴۶۲				استان
	۱۷۹(۷۸/۵)	۴۲(۷۵/۰)	۱۳۷(۷۹/۷)	مازندران
	۴۹(۲۱/۵)	۱۴(۲۵/۰)	۳۵(۲۰/۳)	گلستان
۰/۷۸۵	۱۵۰(۶۵/۸)	۳۶(۶۴/۳)	۱۱۴(۶۶/۳)	وجود سابقه فامیلی سرطان
<۰/۰۰۱				میزان تحصیلات
	۹۴(۴۱/۲)	۳۶(۶۴/۳)	۵۸(۳۳/۷)	باسواد
	۱۳۴(۵۸/۸)	۲۰(۳۵/۷)	۱۱۴(۶۶/۳)	بی سواد
				استعمال دخانیات
۰/۷۸۹	۱۶۲(۷۱/۱)	۳۹(۶۹/۶)	۱۲۳(۷۱/۵)	بله
	۶۶(۲۸/۹)	۱۷(۳۰/۴)	۴۹(۲۸/۵)	خیر
۰/۲۳۱				وضعیت اقامت
	۱۹۸(۸۶/۸)	۴۶(۸۲/۱)	۱۵۲(۸۸/۴)	بومی
	۳۰(۱۳/۲)	۱۰(۱۷/۹)	۲۰(۱۱/۶)	غیر بومی

\*برای متغیر کمی سن میانگین (انحراف معیار) و برای بقیه متغیرهای کیفی فراوانی (درصد) گزارش شد.

جدول ۲. نتایج آزمون نیکویی برازش برای بررسی فرضیه مخاطرات متناسب متغیرهای مستقل

P-value	آماره ی کای دو	ضریب همبستگی	
۰/۳۷	۰/۷۹	-۰/۰۶۱	سن در زمان تشخیص
۰/۲۱	۱/۷۱	-۰/۰۸۲	محل سکونت
۰/۱۶	۱/۹۳	-۰/۰۹۳	استان
۰/۱۴	۲/۲۰	-۰/۱۰۲	سابقه فامیلی
۰/۴۷	۰/۵۱	۰/۰۵۱	سطح تحصیلات
۰/۴۳	۰/۶۱	-۰/۰۵۱	استعمال دخانیات
۰/۹۵۱	۰/۰۰	۰/۰۰۴	وضعیت اقامت

\*معنی دار در سطح ۰/۰۵

جدول ۳. نتایج مدل کاکس و مدل پارامتریک وایبل در آنالیز تک متغیری بقای بیماران مبتلا به سرطان ریه

وایبل			کاکس			
AIC	نسبت خطر	فاصله اطمینان ۹۵٪	AIC	نسبت خطر	فاصله اطمینان ۹۵٪	
۸۲۲/۱۶۸	۱	(۰/۹۹۰-۱/۰۱۰)	۱۹۳۹/۲۱۸	۱/۰۰۳	(۰/۹۹۳-۱/۰۱۳)	سن در زمان تشخیص
۸۱۸/۴۲۹	۱/۳۱۲	(۰/۹۹۸-۱/۷۲۵)	۱۹۳۸/۰۸۸	۱/۱۹۰	(۰/۹۰۵-۱/۵۶۵)	محل سکونت
۸۱۹/۹۹۲	۱/۲۸۰	(۰/۹۱۵-۱/۷۹۰)	۱۹۳۸/۷۷۵	۱/۲۵۷	(۰/۸۹۷-۱/۷۶۰)	استان
۸۲۱/۰۲۹	۰/۸۵۶	(۰/۶۴۵-۱/۱۳۶)	۱۹۳۹/۰۳۸	۰/۸۹۵	(۰/۶۷۴-۱/۱۸۸)	سابقه فامیلی
۸۲۲/۱۶۳	۱/۰۱۳	(۰/۷۷۰-۱/۳۳۲)	۱۹۳۹/۲۹۷	۰/۹۲۳	(۰/۷۰۱-۱/۲۱۵)	سطح تحصیلات
۸۱۷/۶۷۷	۰/۷۲۵*	(۰/۵۳۵-۰/۹۸۲)	۱۹۳۶/۷۱۱	۰/۷۷۱	(۰/۵۶۹-۱/۰۴۵)	استعمال دخانیات
۸۲۱/۵۷۹	۱/۱۶۸	(۰/۷۸۰-۱/۷۵۰)	۱۹۳۹/۱۸۹	۱/۱۴۳	(۰/۷۶۲-۱/۷۱۳)	وضعیت اقامت

\*معنی دار در سطح ۰/۰۵

جدول ۴. نتایج مدل کاکس و مدل پارامتری و اویل در آنالیز چند متغیری بقای بیماران مبتلا به سرطان ریه

واویل			کاکس			
AIC	نسبت خطر	فاصله اطمینان ۹۵٪	AIC	نسبت خطر	فاصله اطمینان ۹۵٪	
	۱/۰۰۲	(۰/۹۸۹-۱/۰۱۴)		۱/۰۰۴	(۰/۹۹۱-۱/۰۱۶)	سن در زمان تشخیص
	۱/۲۷۲	(۰/۹۲۷-۱/۷۴۴)		۱/۱۲۹	(۰/۸۲۳-۱/۵۴۷)	محل سکونت
	۰/۹۶۸	(۰/۶۴۳-۱/۴۵۵)		۱/۰۶۴	(۰/۷۱۱-۱/۵۹۱)	استان
۸۲۴/۹۸۴	۰/۸۶۶	(۰/۶۴۹-۱/۱۵۴)	۱۹۴۵/۸۵۴	۰/۸۹۹	(۰/۶۷۴-۱/۱۹۹)	سابقه فامیلی
	۱/۰۶۳	(۰/۷۶۳-۱/۴۸۰)		۰/۹۸۰	(۰/۷۰۲-۱/۳۶۸)	سطح تحصیلات
	۰/۷۱۳*	(۰/۵۱۵-۰/۹۸۳)		۰/۷۷۳	(۰/۵۶۲-۱/۰۶۴)	استعمال دخانیات
	۱/۱۱۰	(۰/۷۲۰-۱/۷۱۲)		۱/۰۹۴	(۰/۷۰۸-۱/۶۹۰)	وضعیت اقامت

\*معنی دار در سطح ۰/۰۵

### بحث و نتیجه گیری

سرطان ریه بیشترین میزان مرگ و میر را در مقایسه با سرطان سینه، پروستات و سایر سرطان‌های چندگانه دارد، و به علت شانس پایین زنده ماندن از مهم‌ترین و کشنده‌ترین نوع سرطان‌ها محسوب می‌شود. به طوری که تنها ۱۳ درصد از مبتلایان به سرطان ریه پنج سال پس از تشخیص زنده می‌مانند [۲۰، ۱۹]. متأسفانه چندین سال طول می‌کشد تا نشانه‌های سرطان ریه خود را بروز دهند. لذا در بیش‌تر مواقع این بیماری در مراحل پیشرفته تشخیص داده می‌شود، یعنی زمانی که درمان تا حد زیادی بی‌اثر شده است [۸]. بدیهی است شناخت و بررسی عوامل موثر بر افزایش طول عمر بیماران پس از تشخیص و شناخت اولیه بیماری‌شان، ضروری به نظر می‌رسد. مدل خطرات متناسب کاکس متداول‌ترین و کاربردی‌ترین ابزار جهت مدل‌سازی در مطالعات بقاء می‌باشد [۱۲]. پژوهشگران حوزه‌ی پزشکی اغلب تمایل دارند جهت تبیین و تحلیل اثر متغیرهای توضیحی بر بقای بیماران از مدل خطرات متناسب کاکس به جای مدل‌های پارامتری استفاده کنند [۲۱]. هنگام به‌کارگیری مدل کاکس بایستی مفروضات اساسی مدل بررسی گردد. مدل کاکس فرض می‌کند تغییرات تابع خطر در سطوح متغیرهای مستقل متناسب و مستقل از زمان است [۱۷]. بر اساس مطالعه سیستماتیک که توسط آلمن و همکاران روی ژورنال‌های سرطان انجام شد تنها ۵ درصد از همه مطالعاتی که از مدل کاکس استفاده کرده‌اند

پیش‌فرض‌های مدل کاکس را بررسی کرده‌اند [۲۲]. برقرار نبودن فرضیات مدل کاکس نتایج قابل اعتمادی نخواهند داشت. در چنین وضعیتی مدل‌های پارامتری مختلفی چون وایل و لگ نرمال می‌توانند مدل‌های جایگزین مناسب‌تری باشند. این مدل‌ها بدون نیاز به فرضیات خطرات متناسب و مدل را بر اساس فرض یک توزیع خاص برای متغیر زمان برازش می‌دهند. هدف از این مطالعه بررسی مقایسه‌ای مدل کاکس و مدل وایل به عنوان یک مدل پارامتری در تحلیل بقای بیماران مبتلا به سرطان ریه بود و برای مقایسه مدل‌ها از ملاک آکائیکه استفاده شد. نتایج مطالعه حاضر نشان داد مصرف دخانیات به عنوان یک عامل خطر تاثیرگذار (یا عامل پیشگویی‌کننده‌ی) مرگ در بیماران مبتلا به سرطان ریه می‌باشد به این مفهوم که بیمارانی که دخانیات مصرف می‌کنند در معرض خطر مرگ بیش‌تری که ناشی از بیماری سرطان ریه است، می‌باشند. این یافته‌ها با نتایج بسیاری از مطالعات در این زمینه همخوانی دارد [۲۳-۲۷]. با وجود این‌که متغیر جنسیت یکی از فاکتورهای تاثیرگذار در بقا می‌باشد اما مطالعه ما تاثیر این فاکتور را نشان نداد. نتایج مطالعه ما با نتایج زهیر و همکاران همخوانی داشت [۲۸]. اما برخی مطالعات بقا برای زنان را بیش‌تر از مردان گزارش کردند [۲۹، ۳۰]. هم‌چنین نتایج این تحقیق نشان داد با وجود این‌که فرضیه خطرات متناسب در مورد داده‌های مورد مطالعه برقرار بود، مدل وایل به عنوان یک مدل پارامتری با توجه به معیار آکائیکه و بررسی نمودار مانده‌ها، کاراتر از مدل کاکس

پارامترهای برآورد شده از مقدار صفر دور باشند این مدل‌ها کارایی بالاتری نسبت به مدل کاکس خواهند داشت [۳۵]. بردبرن و همکاران به بررسی کفایت انواع مدل‌های پارامتری و مدل خطرات متناسب کاکس از طریق مانده‌ها و ملاک آکائیکه پرداختند. در این مطالعه که بر روی داده‌های سرطان ریه و تخمدان انجام شد مدل گامای تعمیم‌یافته به عنوان کاراترین مدل معرفی شد [۳۶]. اربه، فریرا و نونزانتون در یک مطالعه شبیه‌سازی مدل رگرسیون کاکس و مدل زمان شکست شتابیده را مقایسه کردند آن‌ها از روشی که توسط استوت پیشنهاد شده بود و برای برازش مدل‌های رگرسیون خطی با سانسور راست به کار می‌رود استفاده کردند. نتایج مطالعه نشان داد چه در حالتی که فرض خطرات متناسب برقرار است و چه در وضعیتی که این فرض برقرار نیست مدل‌های لگ نرمال و لگ لجستیک و مدل استوت از مدل رگرسیون کاکس کاراترند [۳۷]. ناردی و همکاران نیکویی برازش مدل خطرات متناسب کاکس و مدل‌های پارامتری را بر روی چهار مطالعه بالینی بررسی کردند. این مطالعه که بر اساس باقی‌مانده‌های مارتینگل، باقی‌مانده‌های موزون اسونفیلد و باقی‌مانده‌های انحراف از نرمال انجام شد، مدل وایبل بر اساس معیار میزان تغییرات پارامتر برآورد شده از سایر مدل‌ها برازش بهتری داشت [۳۸، ۱۲-۴۱]. یکی از محدودیت‌های این مطالعه عدم ثبت اطلاعات مربوط به متغیرهای بالینی از جمله نوع سرطان ریه، مرحله بیماری و ... توسط مرکز ثبت سرطان بابل و عدم دسترسی به پرونده‌های پزشکی بیماران بود. نظر به این‌که اهمیت بعضی از فاکتورهای تاثیرگذار بر بیماری هنوز ناشناخته است، شناسایی و اندازه‌گیری این عوامل نیز گاهی غیر ممکن است، از طرفی لحاظ نکردن این عوامل ریسک ناشناخته بین زمان‌های بقا وابستگی ایجاد می‌کند لذا توصیه می‌شود جهت تبیین اثرات این عوامل ناشناخته از مدل‌های شکنندگی در تحلیل بقا استفاده شود.

زمانی که فرضیه خطرات متناسب برقرار است بدون شک محققین تمایل دارند از مدل خطرات متناسب کاکس برای تحلیل داده‌های بقا استفاده کنند اما در بعضی شرایط از جمله

می‌باشد. نتایج حاصل با نتایج سایر مطالعات که مدل وایبل را کاراتر از مدل کاکس برآورد کرده‌اند همسو می‌باشد. در مطالعه روشنی و همکاران که به بررسی رویکردهای پارامتری، نیمه پارامتری و ناپارامتری پرداخته شد نتایج نشان داد بر اساس ملاک آکائیکه مدل وایبل برازش دقیق‌تری نسبت به مدل نیمه پارامتری کاکس و رویکرد ناپارامتری کاپلان-مایر نشان داد. اگرچه در این بررسی نسبت مخاطره در مدل کاکس و مدل‌های پارامتری تقریباً مشابه بود و نتایج حاصل از رویکرد ناپارامتری کاپلان-مایر تا حد زیادی شبیه به رویکردهای پارامتری و نیمه پارامتری بود. همچنین مطالعه روشنی نشان داد مدل‌های پارامتری این قابلیت را دارند که در بررسی‌هایی که موارد سانسور شده نسبتاً کم باشد، نتایج بهتری نسبت به مدل کاکس ارائه دهند [۳۱]. در مطالعه حاضر نیز با توجه به درصد نسبتاً پایین موارد سانسور نتایج با مطالعه روشنی همخوانی دارد. در مطالعه پورحسینقلی و همکاران که به مقایسه رگرسیون کاکس و مدل‌های پارامتری پرداخته شد نتایج حاصل نشان داد مدل کاکس و مدل نمایی در آنالیز چند متغیره نتایج مشابه داشتند. اگرچه یک مدل مشخص به عنوان کاراترین مدل در آنالیز چند متغیره به دست نیامد اما نتایج نشان داد در میان مدل‌های پارامتری برازش داده شده به داده‌های مورد مطالعه مدل لگ نرمال می‌تواند به عنوان مدل جایگزین رگرسیون کاکس در تحلیل بقای بیماران مبتلا به سرطان معده به کار رود [۳۲]. در مطالعه قدیمی و همکاران با توجه به رد شدن فرض خطرات متناسب و نیز مقایسات انجام شده در مورد داده‌های مورد مطالعه مدل لگ لجستیک به عنوان مدل جایگزین برای مدل‌های کاکس یا وایبل در تعیین عوامل موثر بر پیش‌آگهی بیماران مبتلا به سرطان مری معرفی شد [۳۳]. یزدانی و همکاران به مقایسه مدل‌های پارامتری با استفاده از معیار آکائیکه پرداختند. در این مطالعه که بر روی بیماران مبتلا به سرطان معده انجام شد مدل پارامتری وایبل از کارایی بالاتری نسبت به مدل‌های نمایی و گاما توسعه‌یافته برخوردار بود [۳۴]. اکس در مقاله خود بیان می‌کند در صورتی که در مدل‌های پارامتری ضرایب

survival of patients with gastric cancer using frailty model. *BMC Gastroenterol* 2011; 11: 104.

[15] Alizadeh A, Mohammadpour RA, Barzegar MR. Comparing cox model and parametric models in estimating the survival rate of patients with prostate cancer on radiation therapy. *J Mazandaran Univ Med Sci* 2013; 23: 21-29. (Persian).

[16] Cox DR. *Regression models and life-tables*. Breakthroughs in statistics: Springer; 1992. p. 527-41.

[17] Zhu HP, Xia X, Chuan HY, Adnan A, Liu SF, Du YK. Application of Weibull model for survival of patients with gastric cancer. *BMC Gastroenterol* 2011; 11: 1.

[18] Smith PJ. *Analysis of failure and survival data*: CRC Press; 2002.

[19] Özlü T, Bülbül Y. Smoking and lung cancer. *Tuberk Toraks* 2005; 53: 200-209.

[20] Haraguchi S, Koizumi K, Mikami I, Junichi O, Iijima Y, Ibi T, Shimizu K. Clinicopathological characteristics and prognosis of non-small cell lung cancer patients associated with a family history of lung cancer. *Int J Med Sci* 2012; 9: 68.

[21] Royston P. Flexible parametric alternatives to the Cox model: update. *Stata J* 2004; 4: 98-101..

[22] Altman D, De Stavola B, Love S, Stepniwska K. Review of survival analyses published in cancer journals. *Br J Cancer* 1995; 72: 511.

[23] Peto R, Lopez A, Boreham J, Thun M, Heath C. Mortality from smoking in developed countries 1950-2000. Indirect estimates from national statistics. Oxford, New York, Tokyo: Oxford University Press; 1994.

[24] Biesalski HK, De Mesquita BB, Chesson A, Chytil F, Grimble R, Hermus R, et al. European consensus statement on lung cancer: risk factors and prevention. lung cancer panel. *CA Cancer J Clin* 1998; 48: 167-176.

[25] Sopori M. Effects of cigarette smoke on the immune system. *Nat Rev Immunol* 2002; 2: 372-377.

[26] Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat Rev Cancer* 2003; 3: 733-744.

[27] Horn L, Pao W, Johnson D. In Longo. DL, Kasper, DL, Jameson, JL, Fauci, AS, Hauser, SL, Loscalzo, J(eds) *Harrison's Principles of Internal Medicine* (18th ed) Chapter-89 McGraw-Hill 2012.

[28] Zahir ST, Mirtalebi M. Survival of patients with lung cancer, Yazd, Iran. *Asian Pac J Cancer Prev* 2012; 13: 4387-4391.

[29] Battafarano RJ, Piccirillo JF, Meyers BF, Hsu H-S, Guthrie TJ, Cooper JD, Patterson GA. Impact of comorbidity on survival after surgical resection in patients with stage I non-small cell lung cancer. *J Thorac Cardiovasc Surg* 2002; 123: 280-287.

[30] Puri V, Garg N, Engelhardt EE, Kreisel D, Crabtree TD, Meyers BF, et al. Tumor location is not an independent prognostic factor in early stage non-small cell lung cancer. *Ann Thorac Surg* 2010; 89: 1053-1059.

[31] Roshany D, Azadi N, ESMailNASAB N, Yaghoobi M. Application Of Parametric, Semiparametric And Nonparametric Approaches In Survival Analysis Of Patients With Acute Myocardial 2011. (Persian).

[32] Pourhoseingholi MA, Hajizadeh E, Moghimi Dehkordi B, Safaee A, Abadi A, Zali MR. Comparing Cox regression and parametric models for survival of patients with gastric carcinoma. *Asian Pac J Cancer Prev* 2007; 8: 412-416.

[33] Ghadimi M, Mahmoodi M, Mohammad K, Zeraati H, Hosseini M, Fotouhi A. To determine the prognostic factors in esophageal cancer using log-logistic regression model in Iran. *Iran J Epidemiol* 2011; 7: 1-7.

[34] Yazdani J, Sadeghi S, Janbabaei G, Haghighi F. Applying survival analysis to estimate survival time in

در بررسی‌هایی که موارد سانسور شده نسبتاً کم باشد، مدل پارامتری وایبل این قابلیت را دارد که نتایج بهتری نسبت به مدل کاکس ارائه دهد. با توجه به نتایجی که از این مطالعه حاصل می‌شود می‌توان گفت استفاده از مدل پارامتری وایبل سبب شناسایی ریسک فاکتورهای مهم و تاثیرگذار بر بقای بیماران شده که شناسایی این عوامل در بیماران سرطانی، این مزیت را خواهد داشت که بتوانیم از پیشرفت بیماری جلوگیری کرده و زمان بقای بیماران را افزایش دهیم.

## تشکر و قدردانی

از تمام کسانی که در این مطالعه با ما همکاری کردند

تشکر و قدردانی می‌شود.

## منابع

[1] Pak F, Barati M, Shokrolahi M, Kokhaei P. Tumor immunology and tumor escape mechanisms from immune response. *Koomesh* 2014; 412-430. (Persian).

[2] Jamshidi-Naeini Y, Esmaili S, Akbari ME, Davoodi H. Iron deficiency and overload and its association with cancer risk in humans: a review. *Koomesh* 2015; 16. (Persian).

[3] Mousavi SM, Gouya MM, Ramazani R, Davanlou M, Hajsadeghi N, Seddighi Z. Cancer incidence and mortality in Iran. *Ann Oncol* 2009; 20: 556-563.

[4] Smith RA, Glynn TJ. Epidemiology of lung cancer. *Radiol Clin North Am* 2000; 38: 453-470.

[5] Hosseini M, Naghan PA, Karimi S, SeyedAlinaghi S, Bahadori M, Khodadad K, Mohammadi F, Keynama K, Masjedi MR. Environmental risk factors for lung cancer in Iran: a case-control study. *Int J Epidemiol* 2009; 38: 989-996.

[6] American Cancer Society. *Cancer Facts & Figures*. Atlanta: American Cancer Society; 2007..

[7] Cruz CSD, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. *Clin Chest Med* 2011; 32: 605-644.

[8] Silvestri GA, Alberg AJ, Ravenel J. The changing epidemiology of lung cancer with a focus on screening. *BMJ* 2009; 339: 30-53.

[9] Kleinbaum DG, Klein M. *Survival analysis: a self-learning text*: Springer Science & Business Media; 2006.

[10] Cox DR. *Regression models and life-tables*. *Methodological* 1972; 34: 187-220.

[11] Klein JP, Moeschberger ML. *Survival analysis: techniques for censored and truncated data*: Springer Science & Business Media; 2005.

[12] Nardi A, Schemper M. Comparing Cox and parametric models in clinical studies. *Stat Med* 2003; 22: 3597-3610.

[13] Oakes D. Comparison of models for survival data. *Stat Med* 1983; 2: 305-311.

[14] Ghadimi M, Mahmoodi M, Mohammad K, Zeraati H, Rasouli M, Sheikhfathollahi M. Affecting factors on



- [38] Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika* 1982; 239-241.
- [39] Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994; 515-526.
- [40] Grambsch PM, Therneau TM, Fleming TR. Diagnostic plots to reveal functional form for covariates in multiplicative intensity models. *Biometrics* 1995; 51: 1469-1482.
- [41] Nardi A, Schemper M. New residuals for Cox regression and their application to outlier screening. *Biometrics* 1999; 55: 523-529.
- gastric cancer patients. *J Mazandaran Univ Med Sci* 2011; 21: 28-36. (Persian).
- [35] Oakes D. The asymptotic information in censored survival data. *Biometrika* 1977; 441-448.
- [36] Bradburn M, Clark T, Love S, Altman D. Survival analysis Part III: multivariate data analysis-choosing a model and assessing its adequacy and fit. *Br J Cancer* 2003; 89: 605-611.
- [37] Stute W. Consistent estimation under random censorship when covariables are present. *J Mul Anal* 1993; 45: 89-103.

## Parametric and semi parametric survival analysis of patients with lung cancer

Negar Ghasemi (M.Sc)<sup>1</sup>, Namam-Ali Azadi (Ph.D)<sup>1</sup>, Mojhgah Pourmokhtar (Ph.D)<sup>2</sup>, Jamileh Abolghasemi (Ph.D)<sup>1</sup>, Asieh Ashouri (Ph.D)<sup>3</sup>, Tofiqh Mobaderi (M.Sc)<sup>1</sup>, Mahboobeh Rasouli (Ph.D)<sup>\*1</sup>

1. Dept. of Biostatistics, School of Public Health, Iran University of Medical Sciences, Tehran, Iran

2. Blood Transfusion Research Center, High Institute for Research and Education in Transfusion Medicine, Tehran, Iran

3. School of public Health, Guilan University of Medical Sciences, Rasht, Iran

(Received:15 Apr 2017; Accepted: 30 Oct 2017)

**Introduction:** The lung cancer is the third cause of death and also one of the five common cancers in Iran. In this study, the efficiency of semi-parametric Cox models and Weibull parametric models in order to evaluate the effective factors of survival time of patients with lung cancer were investigated.

**Materials and Methods:** This study was a prospective-cohort study in which the total number of 228 patients with lung cancer followed up from 1991-2007 and necessary information such as age at the time of diagnosis, gender, place of residence, education, residence status smoking, family history of cancer, province were collected from Babol cancer registry center. Then, these added to Cox and Weibull models as demographic factors. Akaike information criterion (AIC) was used to compare the efficiency of competing models.

**Results:** In this study, 75% of cases were men and only 8% of patients survived until the end of the study. The 1, 3, and 5- year survival rates were 13, 8 and 8% respectively. Among the factors studied in the Weibull model, the effect of smoking on survival time was significant ( $p < 0.05$ ) but none of them were significant in the Cox model ( $p > 0.05$ ). However, AIC suggested higher efficiency for parametric Weibull model.

**Conclusion:** In spite of the importance of the Cox model is as a more common method by researchers, this study showed that the Weibull model is more efficient in survival data analysis. According to the results of this study smoking prevention is necessary to increase the longevity of patients with lung cancer.

**Keywords:** Survival Analysis, Statistical Models, Lung Neoplasms, Smoking.

---

\* Corresponding author. Tel: +98 9125763631

mahrasouli@gmail.com