

عوامل مؤثر بر بقای بیماران مبتلا به سرطان معده با استفاده از مدل‌های شکنندگی

محمودرضا قدیمی: دانشجوی دکتری، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی تهران
محمود محمودی*: استاد، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت و موسسه تحقیقات ملی سلامت، دانشگاه علوم پزشکی تهران
کاظم محمد: استاد، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی تهران
مصطفی حسینی: دانشیار، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی تهران
محبوبه رسولی: دانشجوی دکتری، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی تهران

فصلنامه پایش

سال دهم شماره چهارم پاییز ۱۳۹۰ صص ۵۱۵-۵۲۴

تاریخ پذیرش مقاله: ۱۳۸۹/۸/۲

انشر الکترونیک پیش از انتشار ۶ شهریور ۱۳۹۰

چکیده

سرطان معده یکی از شایع‌ترین سرطان‌ها در جهان و جزء شایع‌ترین سرطان‌ها در ایران بوده و بیشترین منطقه بروز این نوع سرطان هم از حاشیه دریای خزر گزارش شده است. در تحلیل بقای بیماران به دلیل این که اهمیت بعضی از عوامل مرتبط با بیماری هنوز ناشناخته‌اند، اندازه گیری تمام عوامل مؤثر و مناسب مرتبط با بیماری بسیار دشوار و یا غیرممکن است. لحاظ نکردن این عوامل خطر ناشناخته مشترک باعث ایجاد وابستگی بین زمان‌های بقا می‌گردد که نتایج مدل خطرهای متناسب کاکس و مدل‌های پارامتریک را نامعتبر می‌کند. در این صورت برای مقابله با مشکل فوق از مدل‌های شکنندگی استفاده می‌کنیم.

این پژوهش شامل اطلاعات ۱۱۰ بیمار مبتلا به سرطان معده است که از طریق مرکز ثبت سرطان بابل که زیر نظر دانشکده بهداشت دانشگاه علوم پزشکی تهران قرار دارد، جمع آوری شد. این مطالعه بین سال‌های ۱۳۶۹ تا ۱۳۷۰ انجام شد و بیماران به مدت ۱۵ سال مورد پی‌گیری قرار گرفته‌اند. برای بررسی عوامل مؤثر بر بقای بیماران از مدل کاکس و مدل پارامتریک لگ لجستیک به همراه شکنندگی گاما استفاده شد و معیار مقایسه کارایی مدل‌ها، ملاک آکاییک و بررسی مانده‌ها بود. کلیه محاسبات با نرم‌افزار SAS و STATA انجام و سطح معنی‌داری ۰/۰۵ در نظر گرفته شد.

نتایج حاصل از این مطالعه نشان داد که ۸۳ نفر (۷۵/۴ درصد) از بیماران مرد و ۲۷ نفر (۲۴/۶ درصد) زن بودند. میانگین سنی در بیماران مرد ۶۰/۲ سال و در زنان ۵۷/۵ سال بود. احتمال بقای یک، سه و پنج ساله بیماران به ترتیب ۲۵، ۱۸ و ۱۷ درصد است و میانه زمان بقای بیماران نیز ۸/۶ ماه بود. در مدل نهایی، تنها متغیرهای سن و سابقه خانوادگی سرطان به عنوان مهم‌ترین عوامل پیشگویی کننده بقای بیماران مطرح هستند ($P < 0/05$) و پیشنهاد می‌شود بررسی به موقع افرادی که دارای سابقه خانوادگی سرطان هستند می‌تواند به عنوان یک عامل مهم در کاهش خطر مرگ بیماران مبتلا به سرطان معده مؤثر باشد. طبق ملاک آکاییک و با توجه به ماهیت داده‌ها (غیریکنوا بودن میزان شکست) و نیز رد فرضیه متناسب بودن ($P < 0/05$)، مدل لگ لجستیک به همراه شکنندگی گاما برازش مناسب‌تری برای داده‌ها داشت و می‌تواند به عنوان جایگزین مدل کاکس در تحلیل بقای بیماران مبتلا به سرطان معده به کار رود.

کلیدواژه‌ها: سرطان معده، آنالیز بقا، توزیع لگ نرمال، مدل شکنندگی

* نویسنده پاسخگو: تهران، خیابان انقلاب اسلامی، خیابان قدس، خیابان پورسینا، دانشگاه علوم پزشکی تهران، ساختمان سبز، دانشکده بهداشت و انستیتو تحقیقات

بهداشتی، گروه اپیدمیولوژی و آمار زیستی

تلفن: ۸۸۹۸۹۱۲۳

E-mail: mahmoodim@tums.ac.ir

مقدمه

سرطان معده بیماری است که در آن سلول‌های سرطانی در معده افزایش پیدا می‌کنند و باعث ایجاد توده یا زخم در معده می‌شوند. منشأ این سلول‌های سرطانی معمولاً سلول‌های مخاط معده است. سرطان معده با بیش از ۸۷۰/۰۰۰ مورد جدید در سال یکی از بیماری‌های شایع در جهان است [۱]. این بیماری اغلب هنگامی تشخیص داده می‌شود که در مرحله پیشرفته است و میزان بقای پنج ساله آن کم است (معمولاً کمتر از ۳۰ درصد است) [۱]. به طور کلی، بروز سرطان معده و مرگ و میر آن به طور چشمگیری در ۷۰ سال گذشته کاهش یافته است [۲-۴]. با این وجود، سرطان معده هنوز چهارمین سرطان شایع و دومین علت مرگ در بین سرطان‌ها در جهان است [۵، ۶]. سرطان معده به دلیل پیش آگهی ضعیف و نیز به خاطر بالاترین علت مرگ در سراسر جهان به نام «کاپیتان مرگ» لقب گرفته است [۷، ۸]. افراد سیگاری، افراد بالای ۵۰ سال و مردان در معرض خطر بالاتری برای سرطان معده قرار دارند. رژیم‌های غذایی نیز نقش اساسی در بروز سرطان معده دارند، افرادی که غذاهای سرشار از نمک و میوه‌ها و سبزیجات کمی مصرف می‌کنند در معرض خطر بالاتری هستند. عوامل زیست محیطی نیز مانند مواجهه شغلی با شوینده‌ها، می‌تواند نقش اساسی در بروز سرطان معده داشته باشند [۹-۱۴]. تقریباً ۶۰ درصد از همه موارد سرطان معده در کشورهای در حال توسعه رخ می‌دهد و بالاترین میزان آن نیز در شرق آسیا است [۱]. در ایران بروز این سرطان حدود ۷۳۰۰ مورد در سال است که در بین مردان شایع‌تر است [۱۵-۲۱]. مرگ از سرطان معده نیز اولین علت مرگ در هر دو جنسیت در ایران است [۱۹، ۱۵]. در سال‌های ۱۳۸۴-۱۳۷۹ استان‌های شمالی کشور از قبیل گیلان، مازندران و اردبیل بالاترین میزان بروز را داشتند [۲۲، ۱۸]. کمتر مطالعه‌ای در زمینه بررسی الگوی بقای بیماران مبتلا به سرطان معده در نواحی مختلف ایران انجام شده است [۲۰، ۱۶، ۹].

امروزه به دلیل استفاده روز افزون از تحلیل بقا در مطالعات پزشکی، نیاز به مدل‌های کارا و با انعطاف بیشتر برای داده‌های بقا بیش از پیش احساس می‌شود. روش‌های استاندارد بقا در تحقیقات بالینی بر این فرض استوارند که جامعه مورد بررسی یک جامعه همگن است. بنابراین مدل‌های بقا فرض می‌کنند داده‌های بقای بیماران مختلف مستقل از یکدیگرند و توزیع زمان بقای افراد بیمار یکسان است. یعنی توزیع زمان تا پیشامد مورد نظر برای بیماران

مختلف مستقل و هم توزیع هستند. حال آن که در بسیاری از مواقع جامعه مورد مطالعه را نمی‌توان یک جامعه همگن فرض نمود، بلکه باید جامعه‌ای ناهمگن که در آن ترکیبی از افراد با مخاطرات متفاوت وجود دارند را در نظر گرفت. چرا که در بسیاری از مواقع، به دلایل اقتصادی و یا به دلیل این که اهمیت بعضی از عوامل مرتبط با بیماری هنوز ناشناخته‌اند، اندازه گیری تمام عوامل مؤثر و مناسب مرتبط با بیماری مورد نظر بسیار دشوار و یا غیرممکن است [۲۳-۲۵].

لحاظ نکردن این عوامل خطر ناشناخته مشترک باعث ایجاد وابستگی بین زمان‌های بقا می‌گردد که بکار گیری مدل خطرهای متناسب کاکس و مدل‌های پارامتریک را غیرممکن می‌گرداند. در این صورت برای مقابله با پدیده فوق از مدل‌های شکنندگی (Frailty) استفاده می‌کنیم. مدل‌های شکنندگی، عمده‌ترین روش برای وارد کردن وابستگی درون گروهی (فردی) در آنالیز بقا هستند. این مدل‌ها تغییر پذیری زمان‌های بقا را در دو منبع فرموله می‌کنند. اولین منبع، تغییر پذیری طبیعی است که توسط تابع خطر بیان می‌شود و دومین منبع تغییر پذیری مشترک بین افراد یک گروه (فرد) است که توسط متغیر شکنندگی بیان می‌گردد. از آنجایی که تابع خطر همواره مثبت است، پس باید توزیع مثبتی برای اثر شکنندگی در نظر گرفته شود. در مطالعات بقا برای متغیر شکنندگی معمولاً از توزیع‌های گاما، گاوسی و وارون، لگ نرمال، پایدار مثبت (Positive Stable)، پواسون مرکب و توزیع خانواده واریانس توانی با میانگین متناهی استفاده می‌شود. همچنین به خاطر امکان پذیر شدن برآورد پارامترها و ساده سازی محاسبات و از آنجایی که متغیر شکنندگی نماینده خطر نسبی عوامل ناشناخته است، در بیشتر موارد میانگین توزیع متغیر شکنندگی را برابر یک فرض می‌کنند [۲۶، ۲۷، ۲۳]. در این مطالعه از توزیع گاما برای برآورد شکنندگی استفاده شده است. در تحلیل بقا هنگامی که مرگ و میر به یک حداکثر می‌رسد و سپس بعد از یک دوره متناهی به تدریج کاهش می‌یابد، بهتر است از مدلی استفاده کنیم که میزان شکست غیریکتو داشته باشد، مدل‌های لگ لجستیک و لگ نرمال دارای این خصوصیت هستند. از طرفی توزیع لگ لجستیک تقریب خوبی برای توزیع لگ نرمال دارد و ارجحیت بیشتری برای استفاده نسبت به توزیع لگ نرمال دارد، چون عبارات ساده‌ای برای توابع مخاطره و بقا داشته و نیز استفاده از این توزیع در کار کردن با داده‌هایی که شامل مشاهدات سانسور شده است،

که a تعداد پارامترهای موجود در مدل است و c ضریب ثابتی است که وابسته به نوع مدل به کار رفته است، که $c=2$ برای مدل لگ لجستیک است. هر چه قدر مقدار AIC کمتر باشد کارایی مدل بهتر است [۲۹].

تابع خطر در مدل‌های شکنندگی به صورت زیر نمایش داده می‌شود:

$$h_i(t) = h_0(t)u_i \exp(\beta^T X)$$

که در آن $h_0(t)$ تابع خطر پایه (که برای مدل لگ لجستیک به صورت $h_0(t) = \lambda p t^{p-1} / (1 + \lambda t^p)$ است، که $\lambda > 0$ و $p > 0$ است)، β ها بردار $p \times 1$ بعدی از پارامترهای رگرسیون، X ها بردار $1 \times p$ بعدی از متغیرهای توضیحی و u_i متغیر شکنندگی برای فرد i ام است [۲۶]. با توجه به این که تابع خطر مثبت است، لازم است برای متغیر تصادفی u توزیع مثبتی در نظر گرفت، که در عمل بیشتر از توزیع گامای یک پارامتری با میانگین یک و واریانس θ استفاده می‌شود [۲۶]، که تابع چگالی آن به شکل زیر است:

$$f_U(u) = \frac{u^{1/\theta-1} \exp(-u/\theta)}{\theta^{1/\theta} \Gamma(1/\theta)}$$

تمامی محاسبات با نرم‌افزار SAS و STATA انجام و سطح معنی‌داری ۰/۰۵ در نظر گرفته شد.

یافته‌ها

در این پژوهش ۱۱۰ بیمار مبتلا به سرطان معده مورد بررسی قرار گرفتند که میانگین سن بیماران ۵۹/۵ با انحراف معیار ۱۰/۳ بود که از این تعداد ۸۳ نفر (۷۵/۴ درصد) مرد و ۲۷ نفر (۲۴/۶ درصد) زن بودند. میانه بقای این بیماران ۸/۶ ماه و میزان بقای یک ساله، سه ساله و پنج ساله این بیماران به ترتیب ۲۵، ۱۸ و ۱۷ درصد به دست آمد (جدول شماره ۱). از تعداد ۱۱۰ بیمار مبتلا به سرطان معده ۹۱ نفر در طول این مدت فوت نمودند و ۱۷/۳ درصد همچنان زنده بودند یا اطلاعات دقیقی از وضعیت بقای آنها موجود نبود و به عنوان مشاهدات سانسور شده از راست در نظر گرفته شدند.

از آنجایی که فرض متناسب بودن در مورد داده‌های مورد مطالعه رد می‌شود، بررسی نمودار مانده‌های مدل کاکس و مدل پارامتریک نیز نمایانگر برازش بهتر مدل پارامتریک لگ لجستیک به همراه شکنندگی گاما نسبت به مدل کاکس بود (شکل شماره ۱ و ۲)، که معیار آکاییک نیز این نتایج را تأیید می‌کند. جدول‌های شماره ۲ و

نسبت به توزیع لگ نرمال ساده‌تر است و در عین حال به جز در مواردی که داده پرت داشته باشیم، تقریب خوبی برای توزیع لگ نرمال می‌دهد. مدل لگ لجستیک تنها مدل پارامتریک با دو خاصیت متناسب بودن بخت‌ها (Proportional odds) و زمان شکست شتابیده است [۲۴، ۲۵]. در مطالعه ما نیز یک چنین الگویی برای تابع خطر وجود دارد، یعنی تابع مخاطره ابتدا افزایش و سپس بعد از یک دوره زمانی کاهش می‌یابد، به همین خاطر استفاده از مدل‌های کاکس، وایبل و نمایی مناسب نیست و همانطور که در بالا ذکر شد مدل لگ لجستیک مناسب به نظر می‌رسد [۲۵، ۲۸-۳۰]. هدف از این مطالعه بررسی عوامل مؤثر بر بقای بیماران مبتلا به سرطان معده با استفاده از مدل پارامتریک لگ لجستیک به همراه شکنندگی گاما و مقایسه آن با مدل کاکس است.

مواد و روش کار

این پژوهش شامل اطلاعات ۱۱۰ بیمار مبتلا به سرطان معده است که از طریق مرکز ثبت سرطان بابل که زیر نظر دانشکده بهداشت دانشگاه علوم پزشکی تهران قرار دارد، جمع‌آوری شد. این مطالعه بین سال‌های ۱۳۶۹ تا ۱۳۷۰ انجام شد و بیماران به مدت ۱۵ سال مورد پی‌گیری قرار گرفتند. مشخصه‌های فردی که در این مطالعه مورد بررسی قرار گرفته، شامل سن، جنس (مرد، زن)، سابقه خانوادگی، وضعیت سیگار کشیدن (سیگاری، غیرسیگاری)، شغل (کشاورز، کارمند، سایر)، نژاد (آریایی، گیلکی، ترکمن، سایر)، وضعیت سواد (باسواد، بی‌سواد)، محل سکونت (شهری، روستایی)، وضعیت اقامت (بومی، غیر بومی) هستند. برای بررسی عوامل مؤثر بر بقای بیماران از مدل کاکس و مدل پارامتریک لگ لجستیک به همراه شکنندگی گاما استفاده شد. در این مطالعه برای مقایسه کارایی مدل‌های پارامتریک و مدل کاکس از معیار اطلاع آکاییک (Akaike Information Criteria) و میزان تغییرات استاندارد شده (Standardized Variability) پارامترها که به صورت $(sv = \hat{\sigma}_\beta / |\hat{\beta}|)$ است، استفاده شده است. ملاک اطلاع آکاییک معیاری است که به وسیله آکاییکه [۳۱] پیشنهاد شد و هدف آن اندازه‌گیری نیکویی برازش مدل است. برای مدل‌های مورد استفاده در این تحقیق ملاک اطلاع آکاییک از رابطه زیر محاسبه شده است:

$$AIC = -2 * \log(\text{likelihood}) + 2 * (a + c)$$

سن رابطه معنی‌داری با بقا بیماران نشان نداد. در همه مدل‌ها وجود سابقه خانوادگی سرطان به عنوان یک عامل مهم پیش‌گویی‌کننده بقا در بیماران مبتلا به سرطان معده شناخته شد ($P < 0.001$) (جدول شماره ۳ الف و ب).

در آنالیز تک‌متغیری، نتایج حاصل از مدل لگ لجستیک و مدل کاکس یک اختلاف معنی‌دار را در بین بیمارانی که سابقه مثبت خانوادگی دارند گزارش می‌کنند ($P < 0.001$) (جدول شماره ۲ الف و ب). هیچ‌یک از کدام مدل‌ها نه در آنالیز تک‌متغیری و نه در آنالیز چندمتغیری اثر معنی‌داری برای متغیرهای جنسیت، محل سکونت، استان، میزان تحصیلات، شغل، نژاد و وضعیت اقامت نشان ندادند (جدول‌های شماره ۲ و ۳).

۳ (الف و ب) نتایج حاصل از تحلیل چندمتغیری و تک‌متغیری، مدل کاکس و مدل پارامتریک لگ لجستیک (با و بدون شکنندگی) را بر اساس میزان تغییرات استاندارد شده و نسبت خطر (Relative Risk-RR, Hazard Ratio-HR) برای تک‌تک متغیرها را نشان می‌دهند. در این تحلیل‌ها مدل لگ لجستیک به همراه شکنندگی گاما بر اساس ملاک آکاییک کاراترین مدل شناخته شد. یافته‌ها حاکی از آن است که معیار آکاییک در مدل کاکس بیش از ۱/۵ برابر مدل لگ لجستیک است. نتایج حاصل از مدل لگ لجستیک به همراه شکنندگی گاما در آنالیز چندمتغیری یک اختلاف معنی‌دار بین گروه‌های سنی را گزارش می‌کند ($P = 0.017$). اما در مدل کاکس (با و بدون شکنندگی) متغیر

جدول شماره ۱- مشخصه‌های بیماران مبتلا به سرطان معده به تفکیک جنسیت

متغیر	مرد ($n=83$)	زن ($n=27$)	کل بیماران ($n=110$)	مقدار احتمال
سن در زمان تشخیص (سال)	60.19 ± 9.84	57.52 ± 11.43	59.53 ± 10.26	۰/۲۴۱
محل سکونت				۰/۴۴۲
شهری	۳۶ (۴۵/۴)	۱۴ (۵۱/۸)	۵۰ (۴۵/۴)	
روستایی	۴۷ (۵۶/۶)	۱۳ (۴۸/۲)	۶۰ (۵۴/۶)	
استان				۰/۷۴۰
مازندران	۶۷ (۸۰/۷)	۲۱ (۷۷/۸)	۸۸ (۸۰/۰)	
گلستان	۱۶ (۱۹/۳)	۶ (۲۲/۲)	۲۲ (۲۰/۰)	
وجود سابقه فامیلی سرطان	۴۷ (۷۴/۶)	۱۶ (۲۵/۴)	۶۳ (۵۷/۳)	۰/۸۱۰
وضعیت تحصیلات				۰/۰۲۹
باسواد	۱۳ (۱۵/۷)	۰ (۰/۰)	۱۳ (۱۱/۸)	
بی سواد	۷۰ (۸۴/۳)	۲۷ (۱۰۰)	۹۷ (۸۸/۲)	
شغل				<۰/۰۰۱
کشاورز	۵۴ (۶۵/۱)	۴ (۱۴/۸)	۵۸ (۵۲/۷)	
کارمند	۲ (۲/۴)	۱ (۳/۷)	۳ (۲/۷)	
سایر	۲۷ (۳۲/۵)	۲۲ (۸۱/۵)	۴۹ (۴۴/۶)	
استعمال دخانیات	۵۱ (۶۱/۴)	۶ (۲۲/۲)	۵۷ (۵۱/۸)	<۰/۰۰۱
نژاد				۰/۶۲۵
آریایی	۷۳ (۸۷/۹)	۲۲ (۸۱/۵)	۹۵ (۸۶/۳)	
گیلکی	۱ (۱/۲)	۰ (۰/۰)	۱ (۰/۹)	
ترکمن	۵ (۶/۰)	۲ (۷/۴)	۷ (۶/۴)	
سایر	۴ (۴/۹)	۳ (۱۱/۱)	۷ (۶/۴)	
وضعیت اقامت				۰/۰۹۵
بومی	۶۵ (۷۸/۳)	۲۵ (۹۲/۶)	۹۰ (۸۱/۸)	
غیر بومی	۱۸ (۲۱/۷)	۲ (۷/۴)	۲۰ (۱۸/۲)	

جدول شماره ۲ (الف) - نتایج مدل کاکس در آنالیز تک متغیری بقای بیماران مبتلا به سرطان معده

AIC	مقدار احتمال	Standardized variability (Gamma frailty)	AIC	مقدار احتمال	Standardized variability (with out frailty)	
۷۴۵	۰/۲۵۰	۰/۸۶ (HR=۱/۰۱)	۷۴۳	۰/۲۴۸	۰/۸۶ (HR=۱/۰۱)	سن در زمان تشخیص (سال)
۷۴۶	۰/۶۲۰	۱/۹۲ (HR=۱/۱۳)	۷۴۴	۰/۶۰۰	۱/۹۲ (HR=۱/۱۳)	جنسیت
۷۴۶	۰/۸۷۰	۶/۲۵ (HR=۰/۹۶)	۷۴۴	۰/۸۷۰	۶/۲۵ (HR=۰/۹۷)	محل سکونت
۷۴۴	۰/۱۷۸	۰/۷۴ (HR=۰/۷۱)	۷۴۳	۰/۱۷۷	۰/۷۴ (HR=۰/۷۱)	استان
۷۳۱	<۰/۰۰۱	۰/۲۷ (HR=۲/۳۱) *	۷۳۰	<۰/۰۰۱	۰/۲۷ (HR=۲/۳۱) *	وجود سابقه فامیلی سرطان
۷۴۵	۰/۲۹۶	۰/۹۴ (HR=۰/۶۹)	۷۴۳	۰/۲۹۴	۰/۹۵ (HR=۰/۶۹)	میزان تحصیلات
۷۴۶	۰/۵۹۴	۱/۸۸ (HR=۰/۹۵)	۷۴۴	۰/۵۹۶	۱/۸۹ (HR=۰/۹۴)	شغل
۷۴۶	۰/۷۳۸	۳/۰۳ (HR=۱/۰۷)	۷۴۴	۰/۷۴۰	۳/۰۳ (HR=۱/۰۷)	استعمال دخانیات
۷۴۶	۰/۷۷۲	۳/۵۶ (HR=۰/۹۸)	۷۴۴	۰/۷۷۷	۳/۵۷ (HR=۰/۹۸)	نژاد
۷۳۷	۰/۰۸۰	۰/۳۷ (HR=۱/۹۷)	۷۳۸	۰/۰۸۰	۰/۳۷ (HR=۱/۹۸)	وضعیت اقامت

جدول شماره ۲ (ب) - نتایج مدل پارامتریک لگ لجستیک در آنالیز تک متغیری بقای بیماران مبتلا به سرطان معده

AIC	مقدار احتمال	Standardized variability (Gamma frailty)	AIC	مقدار احتمال	Standardized variability (with out frailty)	
۴۰۶	۰/۰۴۱	۰/۵۱ (RR=۱/۰۳) *	۴۲۶	۰/۱۷۶	۰/۷۴ (RR=۱/۰۲)	سن در زمان تشخیص (سال)
۴۱۰	۰/۷۰۶	۲/۶۳ (RR=۱/۱۳)	۴۲۸	۰/۵۹۶	۱/۲۶ (RR=۱/۸۹)	جنسیت
۴۰۹	۰/۳۸۹	۱/۱۶ (RR=۱/۲۹)	۴۲۹	۰/۸۸۲	۶/۶۷ (RR=۱/۰۶)	محل سکونت
۴۰۸	۰/۱۹۷	۰/۷۸ (RR=۰/۰۶)	۴۲۶	۰/۱۲۸	۰/۶۶ (RR=۰/۵)	استان
۴۰۸	۰/۰۲۶	۰/۷۹ (RR=۱/۵۳) *	۴۱۷	۰/۰۰۱	۰/۳ (RR=۳/۵۶) *	وجود سابقه فامیلی سرطان
۴۰۸	۰/۲۴۶	۰/۸۶ (RR=۱/۸۴)	۴۲۸	۰/۴۷۲	۱/۳۹ (RR=۰/۶۳)	میزان تحصیلات
۴۱۰	۰/۵۴۵	۱/۶۷ (RR=۰/۹۱)	۴۲۸	۰/۵۶۵	۱/۷۲ (RR=۰/۹)	شغل
۴۱۰	۰/۵۹۶	۱/۸۹ (RR=۰/۸۵)	۴۲۸	۰/۹۲۴	۱۰ (RR=۱/۰۴)	استعمال دخانیات
۴۰۹	۰/۴۹۶	۱/۴۷ (RR=۱/۰۵)	۴۲۸	۰/۹۵۷	۲۰ (RR=۱/۰۱)	نژاد
۴۰۶	۰/۰۷۵	۰/۵۶ (RR=۱/۹۷)	۴۲۴	۰/۲۰۵	۰/۴۵ (RR=۲/۷۵)	وضعیت اقامت

* معنی‌دار در سطح ۰/۰۵، HR: نسبت خطر، RR: خطر نسبی

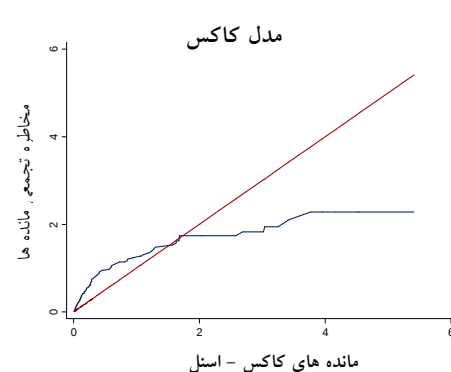
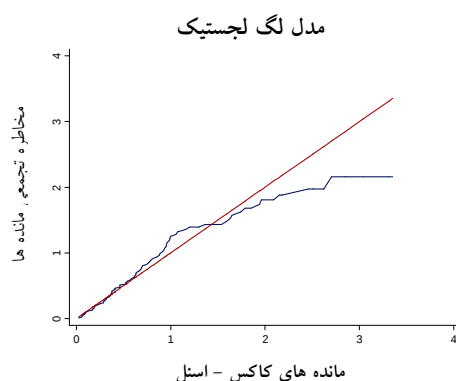
جدول شماره ۳ (الف) - نتایج مدل کاکس در آنالیز چند متغیری بقای بیماران مبتلا به سرطان معده

AIC	مقدار احتمال	Standardized variability (Gamma frailty)	AIC	مقدار احتمال	Standardized variability (with out frailty)	
	۰/۲۰۵	۰/۸۲ (HR=۱/۰۱)		۰/۲۰۴	۰/۸۲ (HR=۱/۰۱)	سن در زمان تشخیص (سال)
	۰/۳۰۰	۰/۸۷ (HR=۱/۴۵)		۰/۲۹۹	۰/۸۸ (HR=۱/۴۵)	جنسیت
	۰/۴۸۰	۱/۲۵ (HR=۰/۸۲)		۰/۴۷۸	۱/۲۵ (HR=۰/۸۳)	محل سکونت
	۰/۱۱	۰/۴۸ (HR=۰/۵۲)		۰/۱۲	۰/۴۹ (HR=۰/۵۳)	استان
۷۳۸	<۰/۰۰۱	۰/۲۵ (HR=۲/۸۴) *	۷۳۶	<۰/۰۰۱	۰/۲۵ (HR=۲/۸۴) *	وجود سابقه فامیلی سرطان
	۰/۵۳۳	۱/۵۴ (HR=۰/۷۷)		۰/۵۳۱	۱/۵۴ (HR=۰/۷۸)	میزان تحصیلات
	۰/۴۲۰	۱/۰۳ (HR=۱/۱۳)		۰/۴۱۳	۱/۰۴ (HR=۱/۱۴)	شغل
	۰/۳۰۰	۰/۸۲ (HR=۰/۷۱)		۰/۲۹۳	۰/۸۲ (HR=۰/۷۲)	استعمال دخانیات
	۰/۶۱۰	۱/۷۴ (HR=۰/۹۶)		۰/۶۰۹	۱/۷۵ (HR=۰/۹۶)	نژاد
	۰/۰۸۰	۰/۵۳ (HR=۱/۷۰)		۰/۰۷۹	۰/۵۴ (HR=۱/۷۱)	وضعیت اقامت

جدول شماره ۳ (ب) - نتایج مدل پارامتریک لگ لجستیک در آنالیز چند متغیری بقای بیماران مبتلا به سرطان معده

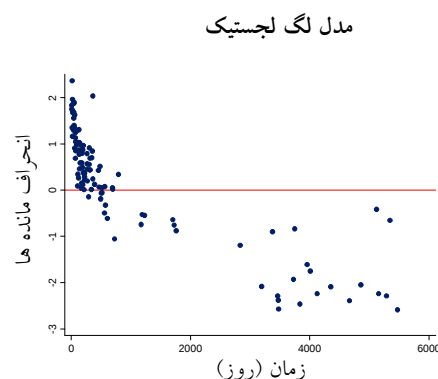
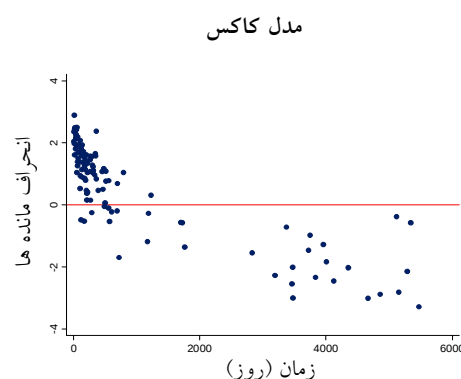
AIC	مقدار احتمال	Standardized variability (Gamma frailty)	AIC	مقدار احتمال	Standardized variability (with out frailty)	
۴۱۴	۰/۰۱۷	۰/۴۱ (RR=۱/۰۴) *	۴۲۴	۰/۰۳۹	۰/۴۹ (RR=۱/۰۳) *	سن در زمان تشخیص (سال)
	۰/۸۱۵	۴/۵۵ (RR=۰/۹۱)		۰/۴۵۶	۱/۲۸ (RR=۱/۴۷)	جنسیت
	۰/۳۴۶	۱/۰۳ (RR=۱/۳۹)		۰/۹۷۷	۲۰ (RR=۰/۹۸)	محل سکونت
	۰/۱۱۸	۰/۷۲ (RR=۰/۵۵)		۰/۱۳۷	۰/۵۳ (RR=۰/۴۰)	استان
	۰/۰۳۴	۰/۴۶ (RR=۲/۲۲) *		<۰/۰۰۱	۰/۲۴ (RR=۴/۷۱) *	وجود سابقه فامیلی سرطان
	۰/۴۹۶	۱/۳۹ (RR=۰/۶۵)		۰/۲۲۲	۰/۸۱ (RR=۰/۵)	میزان تحصیلات
	۰/۳۰۸	۱/۱ (RR=۰/۸۵)		۰/۹۲۶	۶/۲۵ (RR=۱/۰۳)	شغل
	۰/۲۱۵	۰/۷۶ (RR=۰/۶۵)		۰/۳۴۱	۱/۰۱ (RR=۰/۶۹)	استعمال دخانیات
	۰/۷۸۱	۲/۷۸ (RR=۱/۰۳)		۰/۸۲۵	۶/۲۵ (RR=۰/۹۹)	نژاد
	۰/۲۵۰	۰/۸۳ (RR=۱/۵۸)		۰/۰۸۴	۰/۵۷ (RR=۲/۱۳)	وضعیت اقامت

* معنی‌دار در سطح ۰/۰۵، HR: نسبت خطر، RR: خطر نسبی



شکل شماره ۱- مانده‌های کاکس - اسنل در مدل پارامتریک لگ لجستیک و مدل کاکس

خطوط متناظر با مانده‌های کاکس - اسنل در مدل لگ لجستیک نزدیکتر به خط هستند که نمایانگر برازش بهتر این مدل به داده‌هاست. و این نتایج با معیار آکاییک نیز همخوانی دارد.



شکل شماره ۲- انحراف مانده‌ها در مدل پارامتریک لگ لجستیک و مدل کاکس

مقایسه انحراف مانده‌ها نمایانگر انحراف کمتر مدل لگ لجستیک و مناسب‌تر بودن این مدل نسبت به مدل کاکس است.

بحث و نتیجه گیری

سرطان معده یک از بیماری‌های بدخیم شایع در ایران است که بروز آن به ویژه در نواحی شمالی کشور بالاتر است. بر اساس گزارش سازمان جهانی سرطان (IRAC)، بدخیمی معده دومین سرطان مردان و ششمین سرطان زنان است [۲۳، ۳۳]. در مقاله‌ای که در سال ۲۰۰۲ توسط سجادی و همکاران منتشر شده است، میزان وقوع سرطان معده در مردان، رتبه اول و در زنان، رتبه دوم گزارش شده است [۳۲]. نتایج بدست آمده نشان می‌دهد که میزانهای بقا در این بیماری نسبتاً پایین است و به طور کلی افراد پس از تشخیص قطعی بیماری خود مدت زمان زیادی را زندگی نخواهند کرد. در مطالعات مختلف عوامل متعددی به عنوان فاکتورهای مؤثر بر پیش‌آگهی بیماری شناخته و معرفی گردیده است [۳۵، ۳۴، ۲۰، ۱۱]. مطالعات مختلفی در مورد سرطان‌های معده انجام شده است، اما پژوهشگران فیلد پزشکی اغلب تمایل دارند که در تحلیل اثر متغیرهای توضیحی بر بقای بیماران، از مدل‌های خطر متناسب کاکس به جای مدل‌های پارامتریک مشابه استفاده کنند. یک مطالعه سیستماتیک روی ژورنال‌های سرطان نشان می‌دهد که فقط ۵ درصد از مطالعاتی که در آنها مدل کاکس استفاده شده، پیش‌فرض‌های مورد نیاز این مدل را بررسی کرده‌اند [۳۶]. اگر پیش‌فرض‌های خطر متناسب برقرار نباشد، نتایج مدل کاکس قابل اعتماد نخواهد بود. در نتیجه مدل‌های پارامتریکی چون لگ نرمال، لگ لجستیک، وایبل و نمایی در چنین وضعیتی می‌توانند انتخاب‌های مناسب‌تری باشند. این مدل‌ها با فرض وجود یک توزیع خاص برای متغیر زمان و بدون نیاز به فرض‌های خطر متناسب، مدل را برازش می‌دهند [۳۷، ۲۹، ۲۵، ۲۴].

هدف از این مطالعه بررسی و مقایسه مدل پارامتریک لگ لجستیک و مدل کاکس و ارزیابی ورود اثر شکنندگی در تحلیل بقای بیماران مبتلا به سرطان معده بود که برای مقایسه مدل‌ها نیز از ملاک آکاییک و بررسی مانده‌ها استفاده شد. یکی از مزایای مطالعه حاضر کم بودن درصد سانسور از راست است (۱۷/۳ درصد)، چون برای برازش بهتر مدل‌های پارامتریک بهتر است درصد سانسور از راست از ۴۰ الی ۵۰ درصد فزونی نکند [۳۸].

در این بررسی، سن تشخیص بیماری در مدل لگ لجستیک به همراه شکنندگی گاما هم در آنالیز چند متغیری و هم در آنالیز تک متغیری به عنوان یک عامل مهم پیشگویی کننده بقا در بیماران مبتلا به سرطان معده شناخته شد، بدین مفهوم که سنین بالاتر با

یک افزایش محسوس در میزان خطر مرگ ناشی از بیماری مواجه بودند. این یافته با نتایج بسیاری از مطالعات در این زمینه مورد تأیید قرار گرفت [۴۰، ۳۹].

همچنین بسیاری از مطالعات نشان داده‌اند که بیماران دارای سابقه فامیلی سرطان با بقا کمتری همراه هستند [۴۲، ۴۱، ۱۲]. در مطالعه حاضر نیز بیمارانی که سابقه مثبت خانوادگی داشتند با خطر مرگ بالاتری همراه بودند که با نتایج سایر مطالعات مطابقت دارد.

نتایج نشان دادند که بر اساس ملاک آکاییک و با توجه به رد فرضیه‌های خطر متناسب، مدل لگ لجستیک به همراه شکنندگی گاما از سایر مدل‌ها کاراتر است، که در این مدل به جز سن و سابقه مثبت خانوادگی، بقیه متغیرها تأثیر معنی‌داری روی میزان بقای بیماران نداشتند.

ناردی (Nardi) و همکاران مدل کاکس و مدل‌های پارامتریک جایگزین را در سه مطالعه بالینی مقایسه کردند [۳۸]. آنها از مانده‌های میزان انحراف از نرمال برای بررسی فرضیات مدل‌های پارامتریک استفاده کردند [۴۳]. در مطالعه ناردی و همکاران، مدل وایبل بر اساس معیار میزان تغییرات پارامتر برآورد شده، از سایر مدل‌ها بهتر بود. در مطالعه ما نیز این وضعیت برای مدل لگ لجستیک (با شکنندگی و بدون آن) برقرار بود. اربه (HRbe) و همکاران در یک شبیه‌سازی، رگرسیون کاکس و مدل زمان شکست شتابیده را مقایسه کردند [۲۸]. آنها از روش پیشنهاد شده استوت (Stute) که برای برازش مدل‌های رگرسیون خطی با سانسور راست به کار می‌رود استفاده کردند [۴۴]. نتایج نشان داد چه در وضعیتی که فرض‌های خطر متناسب برقرار نیست و چه در حالتی که این فرض برقرار است مدل‌های لگ لجستیک، لگ نرمال و مدل استوت از کاکس کاراترند. آنها همچنین این مقایسه را در مطالعه‌ای پیرامون بقای بیماران مبتلا به سرطان معده انجام دادند که در این مطالعه فرض‌های خطر متناسب برقرار نبود و نتایج حاصل نشان دهنده برازش مناسب برای مدل‌های استوت و لگ نرمال با برآوردهای مشابه برای پارامترها بود. بردبرن (Bradburn) و همکاران کفایت انواع مدل‌های پارامتریک و مدل‌های خطر متناسب کاکس را از طریق بررسی مانده‌ها و ملاک آکاییک بررسی کردند [۴۵] [۴۵]. در مطالعه بردبرن و همکاران که بر روی سرطان ریه و تخمدان انجام شد، مدل تعمیم یافته گاما در مقایسه با دیگر مدل‌های پارامتریک و مدل کاکس کاراتر و دارای کمترین مقدار

نتایج حاصل از دو مدل و ضرایب آنها متفاوت بود، ولی با توجه به معیار آکاییک و نیز رد فرضیه خطرات متناسب، مدل لگ لجستیک به همراه شکنندگی گاما برازش مناسب‌تری برای داده‌ها بود که در این مدل متغیرهای سن و سابقه خانوادگی سرطان به عنوان مهم‌ترین عوامل پیشگویی کننده بقا بیماران مطرح هستند. پیشنهاد می‌گردد اقدامات لازم در جهت تشخیص زودرس بیماری از طریق غربالگری‌های مفید، مؤثر و کم هزینه در مناطق با بروز بالای بیماری (از جمله مناطق شمال و شمال غربی کشور) صورت پذیرد.

سهم نویسندگان

محمودرضا قدیمی: طراحی طرحنامه، جمع‌آوری و تحلیل داده‌ها، تهیه و تدوین مقاله

محمود محمودی: نظارت علمی بر کلیه مراحل اجرای طرح

کاظم محمد: نظارت علمی بر کلیه مراحل اجرای طرح

مصطفی حسینی: مشاور آماری

محبوبه رسولی: جمع‌آوری و تحلیل داده‌ها، تهیه و تدوین مقاله

تشکر و قدردانی

نگارندگان این مقاله، مراتب سپاس و قدردانی خود را از حمایت‌های مالی و اجرایی مؤسسه تحقیقات ملی سلامت دانشگاه علوم پزشکی تهران ابراز می‌دارند. همچنین از تمامی کارکنان ایستگاه تحقیقاتی سلامت بابل که همکاری صمیمانه‌ای در انجام این پژوهش داشته‌اند، کمال تشکر و سپاس را داریم.

آکاییک بود. عبارت شکنندگی ابتدا توسط والپول (Vaupel) و همکاران به منظور تشریح پیامدهای ناشی از وجود چندین منبع تغییرات برای داده‌های یک متغیره طول عمر پیشنهاد شده است [۴۶]. اما بر اساس اطلاعات موجود و نتایج مطالعه حاضر مناسب‌تر است که مدل شکنندگی برای تبیین ناهمگنی موجود در داده‌های بقا به کار رود. مدل‌های اثر تصادفی در تحلیل بقا مدل‌های شکنندگی نامیده می‌شوند. انگیزه به کارگیری این مدل تبیین عوامل خطر ناشناخته و در نظر گرفتن وابستگی بین مشاهدات برای برازش بهتر روی داده‌های بقا است. این مدل‌ها، مدل‌های نسبتاً جدیدی در بقا بوده و به صورت گسترده در دهه ۱۹۹۰ مورد مطالعه قرار گرفته و هم اکنون نیز موضوع بسیاری از تحقیقات را به خود اختصاص داده است. مشکلات تکنیکی در برآورد پارامترها باعث شده است که در مقایسه با مدل کاکس و مدل‌های پارامتریک کمتر استفاده شوند. هندرسون (Henderson) و همکاران یک روش نظری عدم استفاده از مدل شکنندگی را هنگامی که اثر شکنندگی وجود دارد بررسی نموده و نتیجه گرفته‌اند که در این حالات اریبی در برآوردهای رگرسیون ایجاد می‌گردد [۴۷]. شوماخر (Schumacher) و همکاران نشان دادند که چگونه حذف یک عامل مهم، باعث کاهش برآورد خطر نسبی می‌گردد [۴۸]. کیدینگ (Keiding) و همکاران به صورت نظری نشان دادند که چگونه حذف یکی از دو متغیر کمکی باعث افزایش واریانس تابع خطر و اریبی برآورد دیگر متغیر موجود در مدل می‌گردد [۴۹]. همچنین در این مطالعه پیشنهاد شد که برای مقابله با اثر متغیرهای ناشناخته در داده‌های بقا یک متغیره بهتر است از مدل زمان شکست شتابیده استفاده شود.

منابع

1. Stewart BW, Kleihues P. World cancer report. 1 st Edition, IARC Press-WHO: Lyon, 2003
2. Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of eighteen major cancers in 1985. International Journal of Cancer 1993; 54: 594-606
3. Jemal A, Thomas A, Murray T, Thun M. Cancer statistics, 2002. CA Cancer Journal for Clinicians 2002; 52: 23-47
4. Ries LA, Kosary CL, Hankey BF. SEER Cancer Statistics Review 1973-1995. Bethesda: U.S. Dept. of Health and Human Services, Public Health Service.
5. National Institutes of Health: National Cancer Institute, 1998
5. Parkin DM, Bray FI, Devesa SS. Cancer burden in the year 2000: the global picture. European Journal of Cancer 2001; 37: S4-S66
6. Parkin DM. International variation. Oncogene 2004; 23: 6329-40
7. Siewert JR, Maruyama K. What's new in gastric cancer? World Journal of Surgery 1995; 19:483
8. Whelan SL, Parkin DM, Masuyer E. Trends in cancer incidence and mortality. 1 st Edition, IARC Scientific Publications: Lyon, 1993

9. Ghadimi R, Taheri H, Suzuki S, Kashifard M, Hosono A, et al. Host and environmental factors for gastric cancer in Babol, the Caspian Sea Coast, Iran. *European Journal of Cancer Prevention* 2007; 16: 192-95
10. Glade MJ. Food, nutrition, and the prevention of cancer: a global perspective. American Institute for Cancer Research / World Cancer Research Fund. American Institute for Cancer Research, 1997. Nutrition 1999; 15: 523-6
11. Ikeda M, Furukawa H, Imamura H, Shimizu J, Ishida H, Masutani S, et al. Poor prognosis associated with thrombocytosis in patients with gastric cancer. *Annals of Surgical Oncology* 2002; 9: 287-91
12. Munoz S, Ferraroni M, Vecchia C. Gastric Cancer Risk Factors in Subjects with Family History. *Cancer Epidemiology, Biomarkers & Prevention* 1997; 6: 137-40
13. Nagayato T. Background data to study of advance gastric cancer. 1 st Edition, Springer-Verlag Inc: New York, 1986
14. Zali M. Indices related to gastric cancer in Tehran and seven city provinces in the years 1999 to 2002. *Journal of Medical Sciences, Islamic Azad University* 2005; 15: 8-15 [Persian]
15. Akbari ME, Abachizadeh, K, Tabatabaei SM, Ghanbari Motlagh A, Majd Jabari Z, Khaiamzadeh M, et al. Cancer in Iran. 1st Edition, Darolfekr Publications: Qom. Available at: http://wwwwcrcir/cancer_bookhtm#id [Persian]
16. Hajiani E, Sarmast Shoostari MH, Masjedizadeh R, Hashemi J, Azmi M, Rajabi T. Clinical profile of gastric cancer in Khuzestan, southwest of Iran. *World Journal of Gastroenterology* 2006; 12: 4832-5
17. Iranian Annual of Cancer Registration. 2001-2002. Tehran, Iran: Cancer Office, Center for Disease Control, Deputy for Health, Ministry of Health and Medical Education, 2002 [Persian]
18. Mosavi M, Ramezani H. National report on registered cancer cases in 2005. Tehran, Iran: Cancer Office, Centre for Disease Control, Deputy for Health, Ministry of Health and Medical Education, 2005 [Persian]
19. Naghavi M. Iranian annual of national death registration report. Tehran; Ministry of Health and Medical Education, 2005 [Persian]
20. Sadighi S, Raafat J, Mohagheghi M, Meemary F. Gastric cancer: 5 years experience of a single institute. *Asian Pacific Journal of Cancer Prevention* 2005; 6: 195-6
21. Samadi F, Babaei M, Yazdanbod A, Fallah M, Nouraei M, Nasrollahzadeh D, et al. Survival rate of gastric and esophageal cancers in Ardabil province, North-West of Iran. *Archives of Iranian Medicine* 2007; 10: 32-7
22. Alidina A, Gaffar A, Hussain F, Islam M, Vaziri I, Burney I, et al. Survival data and prognostic factors seen in Pakistani patients with esophageal cancer. *Annals of Oncology* 2004; 15: 118-22
23. Hougaard P. Modeling heterogeneity in survival data. *Applied Probability* 1991; 28: 695-701
24. Hougaard P. Analysis of Multivariate Survival Data. 1 st Edition, Springer-Verlag Inc: New York, 2000
25. Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data. 1 st Edition, Springer-Verlag Inc: New York, 2003
26. Duchateau L, Janssen P. The Frailty Model. 1 st Edition, Springer-Verlag Inc: New York, 2008
27. Hail RN. Frailty models in survival analysis. 1 st Edition Halle-Wittenberg: Wittenberg, 2007
28. HRbe J, Ferreira E, Nunez-Anton V. Comparing proportional hazards and accelerated failure time models for survival analysis. *Statistics in Medicine* 2002; 21: 3493-510
29. Kleinbaum DG, Klein M. Survival analysis: a self-learning text. 1 st Edition, Springer-Verlag Inc: New York, 2005
30. Lawless JF. Statistical models and methods for lifetime data. 2nd Edition, Wiley: New York, 2002
31. Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 1974; 19: 716-23
32. Sajadi A, Nouraei M, Mohagheghi MA. Cancer occurrence in Iran in 2002, an international perspective. *Asian Pacific of Cancer Prevention* 2005; 6: 359-63
33. Sajadi A, Raafat J, Mohagheghi MA. Gastric carcinoma, 5 years experience of a single institute. *Asian Pacific of Cancer Prevention* 2005; 6: 195-6
34. Correa P. Clinical implications of recent developments in gastric cancer pathology and epidemiology. *Seminars in Oncology* 1985; 12: 2-10
35. Yazdanbod A, Arshi S, Derakhshan MH, Sadjadi AR, Malekzadeh R. Gastric cardia cancer; the most common type of upper gastrointestinal cancer in Ardebil, Iran: an endoscopy clinic experience. *Archives in Iran Medicine* 2001; 4: 76-9

36. Altman DG, Destavola BL, Love SB, Stepniowska KA. Review of survival analyses published in cancer journals. *British Journal of Cancer* 1985; 72: 511-8
37. Hoesmer D, Lemeshow S. *Applied survival analysis*. 1 st Edition, Wiley: New York, 1999
38. Nardi A. Comparing cox and parametric models in clinical studies. *Statistics in Medicine* 2003; 22: 3597-610
39. Alidina A, Gaffar A, Hussain F, Islam M, Vaziri I, Burney I, et al. Survival data and prognostic factors seen in Pakistani patients with esophageal cancer. *Annals of Oncology* 2004; 15: 118-22
40. Petrequin P, Huguier M, Lacaine F, Houry S. Surgically treated esophageal cancers: predictive model of survival. *Gastroenterology and Clinical Biology* 1997; 21: 12-6
41. Larson P. Patients with a family history of cancer a guide to primary care. *Sussex Cancer Network* 2007; 2: 1-15
42. Boccia B. *Genetic Determinants of Gastric Cancer*. Rome: Erasmus University Rotterdam, 2009
43. Nardi A, Schemper M. New residuals for cox regression and their application to outlier screening. *Biometrics* 1999; 55: 523-9
44. Stute W. Consistent estimation under random censorship when co variables are present. *Journal of Multivariate Analysis* 1993; 45: 89-103
45. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part III: multivariate data analysis-choosing a model and assessing its adequacy and fit. *British Journal of Cancer* 2003; 89: 605-11
46. Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamic of mortality. *Demography* 1997; 16: 439-54
47. Henderson R, Oman P. Effect of frailty on marginal regression. *JRStatistSocB* 1999; 61: 367-79
48. Schumacher M, Olschewski M, Schmoor C. The impact of heterogeneity on the comparison of survival times. *Statistics in Medicine* 1987; 6: 773-84
49. Keiding N, Anderson PK, Klein PJ. The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates. *Statistics in Medicine* 1997; 16: 215-24