

## تحلیل بقای بیماران مبتلا به سرطان معده بعد از عمل جراحی، بر اساس یک مدل انعطاف پذیر در رقابت جویی خطرات

مریم ناظمی پور: دانشجوی دوره کارشناسی ارشد، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی تهران، تهران، ایران  
محمود محمودی: استاد، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی تهران، تهران، ایران- نویسنده  
رابط: mahmoodim@tums.ac.ir

حجت زراعتی: دانشیار، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی تهران، تهران، ایران  
عباس رحیمی فروشانی: دانشیار، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی تهران، تهران، ایران  
تاریخ دریافت: ۱۳۹۱/۵/۱۴ تاریخ پذیرش: ۱۳۹۱/۱۰/۲۵

### چکیده

زمینه و هدف: در بسیاری از مطالعات تشخیصی از جمله بررسی بقای بیماران مبتلا به سرطان معده که هر فرد بعد از عمل جراحی با بیشتر از یک رخداد مواجه است و رخداد یکی سبب پوشانده شدن رخداد دیگری می شود، بحث رقابت جویی خطرات مطرح خواهد بود که برای بررسی اثر هر متغیر کمکی بر رخداد هر حادثه و برآورد تابع مخاطره، از مدل‌هایی چون کاکس و Fine و Gray استفاده می شود. حال اگر پیش فرض‌های دو مدل برقرار نباشد، بکارگیری آنها نادرست خواهد بود. یکی از راه‌های برطرف کردن چنین مشکلی استفاده از مدل‌هایی با انعطاف بیشتر است.

روش کار: در این مطالعه خصوصیات جمعیت شناختی، بالینی و درمانی ۳۳۰ بیمار مبتلا به سرطان معده که طی سال‌های ۱۳۷۴ تا ۱۳۷۸ به انستیتو کانسر مجتمع بیمارستانی امام خمینی مراجعه کرده و تحت عمل جراحی قرار گرفتند، به همراه متغیرهای، نوع و زمان تا اولین رخداد حادثه (عود بیماری / مرگ) برای هر بیمار جمع‌آوری و بررسی شد. با استفاده از این اطلاعات نمودار تابع تجمعی بروز عود بیماری (CIF)، توسط سه مدل کاکس، Fine و Gray و مدل انعطاف پذیر و با استفاده از نرم افزار R(2.14.0) رسم و با نمودار تابع تجمعی بروز عود مشاهده شده برای بیماران، مقایسه و کارایی آنها بررسی گردید.

نتایج: فرض متناسب بودن خطرات، تقریباً برای همه ی متغیرها به ازای هر حادثه برقرار می‌باشد. نمودار تابع تجمعی بروز برای حادثه مورد نظر (عود)، نشان داد که مدل کاکس در برآورد آن دچار بیش برآوردی شده و نیز نمودارهای دو مدل دیگر بسیار شبیه به هم و علاوه بر آن شبیه به نمودار مشاهده شده می باشند. با این تفاوت که نمودار تابع تجمعی بروز مربوط به مدل انعطاف پذیر، هموارتر می- باشد طوریکه در پیش بینی خطر می‌تواند دقت بیشتری داشته باشد.

نتیجه‌گیری: با وجود رخدادهای رقیب، مدل کاکس چندان مناسب چنین وضعیتی نبوده و به نظرمی‌رسد مدل انعطاف پذیر نه تنها جایگزین مناسبی برای مدل Fine و Gray می‌باشد بلکه در صورت عدم برقراری فرض متناسب بودن، نسبت به آن برتر نیز خواهد بود.  
واژگان کلیدی: تحلیل بقا، رقابت جویی خطرات، تابع بروز، اثرات وابسته به زمان، سرطان معده

### مقدمه

است. اما به هر حال در دنیای واقعی ممکن است برخی از بیماران قبل از مواجهه با سکنه قلبی به دلایلی جز آن بمیرند. در چنین حالتی مرگ سبب پوشانده شدن سکنه شده و بحث رخدادهای رقیب به میان خواهد آمد (Prentice et al. 1978; Pintilie 2006; Klein 2010).

در بسیاری از مطالعات بقا، بیماران ممکن است با بیشتر از یک نوع حادثه مواجه باشند. به عنوان مثال فرض کنید گروهی از بیماران قلبی تا زمان مشاهده‌ی سکنه قلبی دنبال شوند. اگر تا پایان مطالعه برای هر فرد سکنه رخ دهد یا زنده بماند روش‌های معمولی بقا قابل کاربرد

اما آنچه که در اینجا مورد توجه می باشد پیش فرض دو مدل (کاکس و مدل Fine و Gray) است یعنی نسبت تابع مخاطره‌ی تجمعی بروز در مدل Fine و Gray و نسبت تابع مخاطره به علت مشخص در مدل کاکس برای هر دو رخداد رقیب در طول زمان ثابت بوده و هر متغیر کمکی اثر یکسانی بر روی مخاطره‌ی هر رخداد داشته باشد. حال اگر پیش فرض‌ها برقرار نباشد، بکارگیری این مدل‌ها نادرست خواهد بود. یکی از راه‌های برطرف کردن چنین مشکلاتی استفاده از مدل‌هایی با انعطاف بیشتر است. پس به منظور وجود محدودیت‌های ذکر شده، مدلی انعطاف پذیر با استفاده از برازش رگرسیون (spline) (Regression) پیشنهاد شد که احتیاج به فرض متناسب بودن ذکر شده ندارد. در این مطالعه هدف مقایسه‌ی تابع تجمعی بروز عود مشاهده شده برای بیماران با تابع تجمعی بدست آمده از سه مدل کاکس، Fine و Gray و مدل انعطاف پذیر و بررسی کارایی آنها می باشد.

## روش کار

زمانی بحث رقابت جویی خطرات (Competing risks) مطرح می شود که هر فرد بیشتر از یک نوع حادثه را بتواند تجربه کند و رخداد یک حادثه سبب پوشانده شدن رخداد دیگر حوادث بشود. در واقع افراد تا وقوع اولین حادثه از میان  $j \geq 2$  حادثه‌ی متفاوت، یا تا زمانی که از دسترس خارج شده و دیگر نتوان آنها را دنبال کرد پیگیری خواهند شد. (Prentice et al. 1978; Pintilie 2006).

تابع تجمعی بروز Cumulative Incidence Function (CIF): اگر  $t_1 < t_2 < \dots < t_r$  زمان‌های شکست مشاهده شده باشند،  $d_{ij}$  تعداد رخداد حادثه‌ی  $i$ ام در زمان  $t_j$ ،  $n_j$  تعداد افراد در معرض خطر در زمان  $t_j$  است که برایشان حادثه‌ی مورد نظر رخ نداده یا سانسور شده باشند. تابع تجمعی بروز، احتمال رخداد حادثه‌ی مورد نظر، قبل از زمان  $t_j$  برای همه افرادی است که تا قبل از زمان  $t_j$  هیچ یک از حوادث برایشان رخ نداده و هنوز در معرض خطر

همانطور که می دانیم اغلب هدف، برآورد احتمال رخداد حادثه مورد نظر بعد از نقطه‌ای از زمان می باشد. (Pintilie 2006)  $(1-KM_i)$  که در آن  $KM_i$  برآورد کاپلان مایر برای زمان  $i$ ام است، احتمال رخ داد حادثه مورد نظر را در حالتی ایده آل و غیر واقعی، یعنی عدم حضور دیگر حوادث رقیب بیان خواهد داشت. بنابراین برای بررسی اثر هر متغیر کمکی (Covariate) بر تابع مخاطره حادثه مورد نظر، استفاده از مدل مخاطرات متناسب کاکس (Cox proportional hazard model) و برآورد کاپلان مایر (KM) چندان عملی نخواهد بود (Kaplan and Meier 1958; Pepe and Mori 1993; Lunn and McNeil 1995; Tai et al. 2001). پس استفاده از برآورد کاپلان مایر، بیش برآوردی از احتمال رخداد حادثه‌ی مورد نظر بدست خواهد داد. چرا که در چنین حالتی آنالیز بقا بر روی یک رخداد تکیه دارد و سایر رخدادها به عنوان سانسور و مستقل از هم در نظر گرفته می شوند. گویی بخشی از اطلاعات دور ریخته می شود.

این گونه بود که در سال ۱۹۸۰، روشی به نام Kalbfleisch (CIF) توسط Prentice Kalbfleisch and Prentice (1980) پیشنهاد داده شد که وجود رخدادها را رقیب را نیز در نظر می گرفت (Kalbfleisch and Prentice 1980).

با استفاده از این روش احتمال رخداد همه ی حوادث به صورت کلی (هم حادثه مورد نظر و هم حوادث رقیب) به احتمال رخداد هر حادثه به صورت مجزا تفکیک می شد و همانند مدل کاکس فرض را بر استقلال حوادث رقیب نمی گذاشت.

در سال ۱۹۹۹ با مدل کردن تابع مخاطره ی تجمعی بروز  $\gamma(t)$  یا hazard of CIF یا hazard of subdistribution (Fine و Gray پیشرفت قابل توجهی در مطالعات بقا در حضور رخدادهای رقیب بوقوع پیوست (Fine 1999; Fine and Gray 1999).

در این مدل، توابع

$$\beta_{jp}(t), \alpha_p(t), b_j(t), \log(\lambda_1(t))$$

توسط برازش رگرسیونی درجه ۳ بکار برده می‌شود که این برازش، مجموعه‌ای از چند جمله‌ای‌هایی از درجه  $m$  به صورت زیر می‌باشد.

$$S(t) = \sum_{j=0}^m \beta_j t^j + \sum_{i=1}^k \theta_i (t-t_i)_+^m$$

+

$m$  درجه‌ی چندجمله‌ای،  $k$  تعداد نقاط اتصال این

چندجمله‌ای‌ها،  $t_i$  ها مکان اتصال  $\theta_i$  و  $\beta_j$  ضرایب این چند جمله‌ای می‌باشد. در این مطالعه  $m=3$  و  $k=1$  در نظر گرفته شده است (Durrleman and Simon 1989; De Boor 1978; Sleeper and Harrington 1990; Hess 1994; Abrahamowicz and MacKenzie 2007; Abrahamowicz et al. 1996)

- چگونگی اجرای مدل به تفصیل در Appendix (Belot et al. 2010) و همچنین در بخش ضمایم آورده شده است. با استفاده از این مدل می‌توان پیش‌فرض‌های ذکر شده برای دو مدل قبلی (۱ و ۲) یعنی: یکسان بودن شکل

$$H_0: b_j(t) = b_j$$

تابع مخاطره اولیه، به ازای هر حادثه مرجع:  $H_0$  - ثابت بودن اثر هر متغیر کمی بر روی حادثه مرجع:  $H_0$

$$\alpha_p(t) = \alpha_p$$

و ثابت بودن نسبت اثر هر متغیر کمی بر روی تابع مخاطره حادثه‌ی  $j$  ام، به اثر آن بر روی حادثه‌ی مرجع  $\beta_{jp}(t) = \beta_{jp}$

$H_0$  را توسط آزمون نسبت درست‌نمایی آزمون کرد (Abrahamowicz et al. 1996; Giorgi et al. 2003; Remontet et al. 2007)

کاربرد در سرطان معده: اطلاعات ۳۳۰ بیمار مبتلا به سرطان معده که از ابتدای سال ۱۳۷۴ تا فروردین ۱۳۸۷ به انستیتو کانسر مجتمع بیمارستانی امام خمینی مراجعه کرده و تحت عمل جراحی قرار گرفته، بررسی شد. در همه‌ی آنها محل اولیه سرطان معده بوده که تشخیص آن بر اساس بیوپسی صورت گرفته است. یکسری خصوصیات جمعیت شناختی، چون، سن (هنگام عمل جراحی) که به دو گروه

باشند. که احتمال آن با  $\hat{S}(t_{j-1})$  نشان داده می‌شود (Kalbfleisch and Prentice 1980; Klein 2010).

$$CF_i(t_j) = F_i(t) = \sum_{all\ j, t_j \leq t} \frac{\hat{h}_{ij} S(t_{j-1})}{\sum_{n_j} S(t_{j-1})}$$

تابع مخاطره به علت مشخص

(Cause-specific hazard function)  $h_{ij} \rightarrow i$

مروری بر مدل‌های کاکس، Gray و Fine، انعطاف پذیر (Flexible model):

در مدل کاکس فرض بر ثابت بودن، نسبت تابع مخاطره به علت مشخص به ازای هر حادثه و اثر هر متغیر کمی بر تابع مخاطره، در طول زمان می‌باشد (Cox 1972). که به صورت زیر نشان داده می‌شود.

رابطه ۱:

$$\lambda_j(t, X) = \lambda_1(t) \exp(b_j + \sum_{p=1}^P \alpha_p x_p + \sum_{p=1}^P \beta_{jp} x_p)$$

$\beta_{1p} = 0$  و  $b_1 = 0$  مرجع و  $j=1$  حادثه

در این مدل  $(\alpha_p)$  اثر مشترک و  $(\beta_{jp})$  اثر متفاوت بر روی حوادث می‌باشد.

اگر فرض ثابت بودن نسبت تابع مخاطره، به ازای حوادث مختلف برقرار نباشد می‌توان نسبت به حوادث طبقه بندی کرده از مدل کاکس طبقه‌ای (Stratify proportional hazard model) استفاده کرد.

مدل Gray و Fine با قبول وجود رخدادهای رقیب، تابع مخاطره‌ی تجمعی بروز به صورت زیر بکار برده میشود (Lunn and McNeil 1995; Fine and Gray 1999; Belot et al. 2010) رابطه ۲

$$\gamma(t, X) = \gamma_0(t) \exp(\beta X)$$

مدل انعطاف پذیر (Flexible model): با توسعه‌ی مدل (۱) بوسیله‌ی برازش رگرسیونی درجه ۳ (Cubic) Regression splines می‌توان به مدل انعطاف پذیرتری به شکل زیر دست یافت.

رابطه ۳:

$$p=1,2,\dots,P \quad \beta_{1p}(t) = 0 \quad \text{و} \quad b_1(t) = 0$$

به نتایج آن، متغیری که نشان‌دهنده‌ی نوع درمان تکمیلی دریافت شده است، بر عود بیماری، موثر و نسبت مخاطره برای آن ۱/۳۹ می‌باشد. بزرگتر بودن نسبت مخاطره از ۱ بدلیل این می‌باشد که افرادی که تقاضای دریافت درمان تکمیلی کرده‌اند مخاطره‌ی بیشتری نسبت به کسانی که تقاضا نکرده‌اند دارند.

در شکل ۱ نمودار تابع تجمعی بروز مشاهده شده و تابع تجمعی بروز بدست آمده از هر سه مدل برای حادثه مورد نظر (عود) نشان داده شده است. با مقایسه‌ی آنها دیده می‌شود مدل کاکس (b) بدلیل سانسور قرار دادن دیگر رخدادهای رقیب و گذاشتن فرض بر استقلال آنها، در برآورد تابع تجمعی بروز دچار بیش برآوردی (Overestimate) شده و نیز نمودار دو مدل دیگر (c,d) بسیار شبیه به هم و علاوه بر آن شبیه به نمودار مشاهده شده (a) می‌باشند. با این تفاوت که نمودار تابع تجمعی بروز مدل انعطاف پذیر (d)، هموارتر است.

البته بدلیل انعطاف پذیر بودن مدل ۳ نسبت به پیش فرض متناسب بودن، در صورت عدم برقراری پیش فرض‌ها، تفاوت قابل ملاحظه‌ای بین نمودار تابع تجمعی بروز مدل Fine و Gray و مدل انعطاف پذیر دیده خواهد شد که این از عمده علل برتری مدل انعطاف پذیر نسبت به مدل Fine و Gray می‌باشد.

## بحث

مطالعه‌ی حاضر تکیه بر روش رقابت جویی خطرات به منظور دستیابی به برآورد دقیق‌تری از بقای بیماران پس از عمل جراحی داشته و از میان روش‌های گوناگون برآورد توابع تجمعی بروز رخدادهای رقیب (عود بیماری و مرگ)، از مدل انعطاف پذیر نیز بهره جسته که با وجود تمرکز این مدل بر تابع مخاطره به علت مشخص (ESHF) و فرض بر استقلال رخدادهای رقیب، از نتایج خوبی نسبت به دیگر روش‌ها برخوردار می‌باشد. و همچنین نسبت تابع مخاطره‌ی اولیه (BHF) برای رخدادهای رقیب و اثر هر متغیر کمکی بر تابع مخاطره‌ی هر

کمتر و مساوی ۶۰ سال و بالای ۶۰ سال طبقه بندی شده، جنس (مرد/زن) و خصوصیات بالینی آنها همچون وجود متاستاز (بلی/خیر)، مرحله بیماری (I - II- III- IV) و خصوصیات درمانی همانند دریافت درمان تکمیلی پس از عمل (شیمی درمانی، رادیوتراپی، جراحی و یا ترکیبی از سه روش) به همراه متغیرهای، زمان تا اولین رخداد حادثه (عود بیماری/ مرگ) که به عنوان حوادث رقیب در نظر گرفته شده‌اند برای هر بیمار جمع‌آوری شد که اطلاعات توصیفی راجع به این بیماران به تفصیل در جدول ۱ آورده شده‌اند. (نرم افزار مورد استفاده R(2.14.0 می‌باشد)

با استفاده از این اطلاعات تابع تجمعی بروز عود بیماری پس از عمل جراحی، توسط سه مدل کاکس، Fine و Gray و مدل انعطاف پذیر رسم و با نمودار تابع تجمعی بروز عود مشاهده شده برای بیماران مقایسه و کارایی آنها بررسی گردید.

## نتایج

۳۳۰ بیمار مورد مطالعه در بازه‌ی سنی ۳۲-۹۶ سال (  $65.61 \pm 11.00$ : انحراف معیار  $\pm$  میانگین) قرار داشتند. اطلاعات توصیفی راجع به بیماران و متغیرهای مورد بررسی در جدول ۱ مختصراً آمده که به تفصیل در (Zeraati et al. 2005) آورده شده است.

پس از رسم نمودارهای  $\text{Log-Log } \hat{S}$  و  $\text{Log-Log } (1-\text{CIF})$  در برابر  $\text{Log}(\text{time})$  برای هر سطح از هر متغیر و مقایسه نمودار برای سطوح مختلف و یا با استفاده از مدل کاکس توسعه یافته (Extended cox model) برای همه‌ی متغیرها به ازای هر حادثه فرض متناسب بودن خطرات (Proprtionality) برقرار می‌باشد.

بنابراین پیش فرض برای هر دو مدل کاکس و Fine و Gray رعایت شده است. همچنین با در نظر گرفتن وجود رخدادهای رقیب و با استفاده از مدل Fine و Gray اثر هر متغیر کمکی نیز مورد آزمون قرار گرفت که نتایج آن در جدول ۲ قابل ملاحظه می‌باشد. که باتوجه

در همین سال‌ها کاکس، مدل رگرسیونی که بعدها بنام خود او معروف شد را ارائه نمود، که تمامی ایده‌های مطرح شده تا آن زمان را تحت الشعاع خویش قرار داد. با کاربرد این مدل رگرسیونی امکان بررسی اثر متغیرهای مختلف بر طول عمر بطور همزمان فراهم گردید. این مدل چندی بعد توسط خود او توسعه یافت و چگونگی برآورد ضرایب مدل رگرسیونی با روش‌های دقیقتر از جمله روش درست‌نمایی جزئی (Partial likelihood) تشریح گردید. در مدل توسعه یافته کاکس بررسی اثر متغیرهای کمی و کیفی و نیز متغیرهای وابسته به زمان بر احتمالات بقاء امکان پذیر است (Cox 1972; Cox 1975).

با ارائه مدل کاکس، مطالعات بسیاری از آمارشناسان بر این روش و چگونگی برطرف کردن ضعف‌های آن متمرکز شد.

در دههٔ اخیر مطالعات تنوریک در زمینهٔ تحلیل بقاء در دو شاخهٔ اصلی انجام می‌شود. گروه اول، مطالعاتی است که مدل رگرسیونی کاکس را برای نقطهٔ شروع پذیرفته و تلاش می‌کنند راه‌هایی را برای افزایش کارآمدی نتایج حاصل از این مدل بیابند. در چند سال گذشته استفاده از برآوردگرهای بییزی، و استفاده از الگوریتم‌های مختلف برای کاهش واریانس در این رابطه بسیار مرسوم شده است. دلیل تلاش پژوهشگران برای بهبود مدل کاکس، سادگی کاربرد آن و استنتاج ساده از نتایج حاصل از آن است. اما گروه دوم، که مطالعات خویش را بر مبنای استفاده از فرآیندهای تصادفی گذارده‌اند، تلاش می‌کنند با توجه به رفتار پدیده‌های مورد بررسی با استفاده از مدل‌های مختلف فرآیندی مدل‌های جدید تری را ارائه نموده و کاربرد مدل‌های موجود را ساده‌تر نمایند. و اما آنالیز داده‌های رقابت جویی خطرات، در ابتدا به Daniel Bernoulli در سال ۱۷۶۰، و برای برآورد میزان مرگ در جامعه برمی‌گردد (Bernoulli 1760; Klein 2010).

رایج‌ترین روش برای این داده‌ها، بررسی داده‌های بدست آمده از زمان‌های شکست است. یکی از روش‌هایی که برای تحلیل داده‌های بقا وجود دارد، روش توصیف شده

رخداد را در طول زمان، با برازش رگرسیونی درجه ۳ مدل کرده و می‌توان فرض‌های متناسب بودن را نیز به-راحتی و با استفاده از آزمون نسبت درست‌نمایی، آزمون کرد. به صورت کلی، اولین دیدگاه‌های نظری برای تحلیل بقا، مبتنی است بر روش‌های کلاسیک موسوم به "جدول طول عمر" که توسط Kaplan و Meier و در سال ۱۹۵۸ ارائه شد. آنها در بررسی خویش با توجه به داده‌های ناتمام موجود در مطالعات بقاء، روشی ناپارامتری را برای برآورد احتمالات بقا ارائه نمودند (Kaplan and Meier 1958). Sverdrup در سال ۱۹۶۵، ضعف‌های این روش را مطرح و استفاده از مدل‌های مبتنی بر فرآیندهای تصادفی را پیشنهاد نمود (Sverdrup 1965).

سپس (Hoem) نیز در سال ۱۹۶۹ برای اولین بار امکان استفاده از فرآیندهای "زنجیر مارکف (Markov chain process)" را اعلام نمود (Hoem 1969).

اما Nelson در سال ۱۹۶۹، با ارائه مباحث نظری خود در مورد "تابع مخاطره (Hazard function)" هم به توسعهٔ مدل کاپلا نمایر پرداخت و هم باعث شد که برای مدتی بحث استفاده از فرآیندهای تصادفی در مطالعات بقاء به حاشیه رانده شود. Altshuler در مقاله‌ای برای اولین بار در مورد رقابت جویی مخاطرات سخن گفت، و در این مسیر از روش Nelson استفاده نمود (Altshuler 1970).

در سال ۱۹۷۲ Aalen در پایان نامه فوق لیسانس خود برآوردگری را برای تابع مخاطره ارائه نمود که بعدها به برآوردگر نلسون - آلن "معروف شد (Borgan 2005).

در این سال‌ها اقدامات گسترده‌ای نیز در زمینه استفاده از روش‌های ناپارامتری صورت گرفت. از جمله آنها افرادی چون Gehan، Mantel و Breslow بودند (Gehan 1965; Mantel 1966; Breslow 1970; Efron 1988).

همچنین برتری تحلیل در سه مدل ذکر شده در تحلیل بقای بیماران، با مقایسه‌ی نمودار تابع تجمعی بروز صورت گرفت. با توجه به نمودار ۱، مقایسه‌ی بین آنها نشان داد با وجود رخدادهای رقیب، مدل کاکس مقدار تابع تجمعی بروز را برای حادثه‌ی مورد نظر (عود) بیش برآورد کرده و چندان مناسب چنین وضعیتی نمی‌باشد. همچنین نمودار تابع تجمعی بروز دو مدل دیگر بسیار شبیه به هم و شبیه به نمودار تابع تجمعی بروز مشاهده شده بوده و حاکی از آن است که نه تنها مدل انعطاف پذیر جایگزین مناسبی برای مدل Fine و Gray می‌باشد بلکه در صورت عدم برقراری فرض متناسب بودن، نسبت به آن برتر نیز خواهد بود. زیرا در چنین حالتی پیش فرض برای مدل Fine و Gray رعایت نشده و با رسم نمودار تابع تجمعی بروز تفاوت قابل ملاحظه‌ای بین تابع تجمعی بروز دو مدل دیده خواهد شد. علاوه بر این موارد، مفید بودن استفاده از مدل انعطاف پذیر بدلیل هموار بودن نمودار تابع تجمعی بروز بدست آمده و آزمون پیش فرض‌های ذکر شده نیز قابل ملاحظه می‌باشد.

### نتیجه گیری

عود بعد از عمل، رخداد مورد علاقه در این مطالعه بوده که با بررسی اثر متغیرهای کمکی مورد نظر، بر تابع تجمعی بروز آن، تنها متغیر معنی‌دار، نوع درمان تکمیلی دریافت شده دیده شد. اما آنچه که در اینجا حائز اهمیت است بدست آوردن مدلی مناسب با توجه به حضور رخدادهای رقیب است. با مقایسه نمودار تابع تجمعی بروز سه مدل نامبرده شده، به نظر مدل انعطاف پذیر توصیف شده، دارای کارایی و دقت بیشتری نسبت به بقیه بوده و از انعطاف بیشتری برخوردار می‌باشد و مطمئناً با مطالعات بیشتر می‌توان به دیگر نقاط قوت این مدل پی برده و براحتی از آن جهت تحلیل اطلاعات بهره جست.

توسط Kalbfleisch و Prentice (۱۹۸۰) است که بوسیله‌ی Kay (۱۹۸۶) بکار برده شد که در آن برای هر رخداد به صورت جداگانه مدل زده و دیگر رخدادهای به عنوان سانسور در نظر گرفته می‌شوند. اشکال روش نامبرده، عدم بررسی رخدادهای به صورت همزمان است که سبب پیچیدگی مقایسه‌ها بین پارامترهای برآورد شده برای رخدادهاست.

روش دیگر که توسط بسیاری از افراد از جمله Larson و Dinse (۱۹۸۵) و Kuk (۱۹۹۲) مورد استفاده قرار گرفت، شامل مدل‌های بسیار پیچیده‌ای برای رخدادهای بوده که اشکال آن نیز نبود نرم افزارهایی استاندارد، و احتیاج به شبیه‌سازی Monte karlo اشاره شده است.

در این میان مدل پیشنهادی توسط Lunn و McNeil به دلیل در نظر گرفتن مدل Cox (PH) به صورتی گسترده و بررسی توام رخدادهای وجود نرم-افزارهای استاندارد و همچنین آزمون‌هایی همزمان برای پارامترهای رخدادهای مختلف، از کاربرد بیشتری برخوردار است (Prentice et al. 1978).

البته در بررسی داده‌های بقا، می‌توان از مدل‌های Multistate، که از ابزارهای قوی Counting process بهره می‌گیرند را نیز مورد استفاده قرار داد. در این مطالعه با توجه به نتایج بدست آمده از مدل Fine و Gray بر عود و مرگ بیماری، متغیری که نشان‌دهنده‌ی نوع درمان تکمیلی دریافت شده می‌باشد، بر بقای بیماران از هر دو حادثه (عود و مرگ) و متغیر مرحله‌ی بیماری بر مرگ بیماران، معنی‌دار می‌باشد. که نتایج نسبتاً مشابهی با استفاده از آزمون Gray، بدست آمده که متغیر نوع درمان تکمیلی بر هر دو حادثه و متغیرهای سن و وجود متاستاز بر مرگ بیماران پس از عمل جراحی موثر می‌باشند که نتایج مطالعه‌ی حاضر با یافته‌های مطالعات خدمت و همکاران و زراعتی و همکاران مطابقت دارد.

## تشکر و قدردانی

حس تعهد و مسؤولیت در زندگیمان تالوویی خدایست. سپاسگزار تمام کسانی هستیم که مسؤولانه، صبورانه و صادقانه ما را همراهی نمودند.

ضمایم: لازمه‌ی اجرای مدل انعطاف پذیر تکرار اطلاعات هر بیمار به تعداد رخدادهای رقیب است که سبب بوجود آمدن متغیرهایی مجازی (  $\Delta.1$ ,  $\Delta.2$ ) به تعداد رخدادهای رقیب ( $k=2$ ) نیز خواهد شد. پس با این وجود تعداد کل داده‌ها از  $N$  به  $N * K$  افزایش خواهد یافت. علاوه بر آن احتیاج به تقسیم هر جزئیات بیشتر با یک مثال در زیر آورده شده.

یک از فواصل زمانی به فواصل زمانی کوچکتر، و ایجاد دو متغیر که یکی نشان دهنده‌ی طول (Tik) و دیگری نشان-دهنده‌ی میانه‌ی هر فاصله (Intnum) است نیز می‌باشد. که برای رسیدن به این مقاصد به ترتیب از توابعی به نام Duplicate و W.Lexi در نرم افزار R می‌توان بهره جست. با استفاده از لگاریتم تابع درست‌نمایی پواسن، نهایتاً با بکارگیری Glm برای خانواده‌ی پواسن پارامترهای مدل انعطاف پذیر برآورد خواهد شد.

```
> colrec <- data.frame(id=c(1, 2, 3),sex=c(2, 1, 1),time=c(5, 2/82, 0/38),status=c(0, 1, 1),type.evt=c(0, 2, 1))
> colrec
  id    sex  time  status  type.evt
1  1     2  5/00     0         0
2  2     1  2/82     1         2
3  3     1  0/38     1         1
> colrecx2 <- duplicate(status=status, event=type.evt, data=colrec)
> colrecx2
  id    sex  time  status  type.evt  delta.1  delta.2
1  1     2  5/00     0         1         1         0
2  1     2  5/00     0         2         0         1
3  2     1  2/82     0         1         1         0
4  2     1  2/82     1         2         0         1
5  3     1  0/38     1         1         1         0
6  3     1  0/38     0         2         0         1
> library(Epi)
> split.colrecx2 <- W.Lexis(exit=time, fail=status,breaks=seq(0,5,by=0/05),
include=list(id,sex,delta.1,delta.2),data= colrecx2)
> split.colrecx2$tik <- split.colrecx2$Exit - split.colrecx2$Entry
> split.colrecx2$intnum <- (split.colrecx2$Entry + split.colrecx2$Exit)/2
> mod <- glm(Fail ~ bs(intnum, knots=c(1), Boundary.knots=c(0,5)) + delta.2:(bs(intnum,
knots=c(1), Boundary.knots=c(0,5))) + sex + delta.2 + delta.2:sex + offset(log(tik)),
family=poisson,data=split.colrecx2)
(Belot, Abrahamowicz et al. 2010)
```

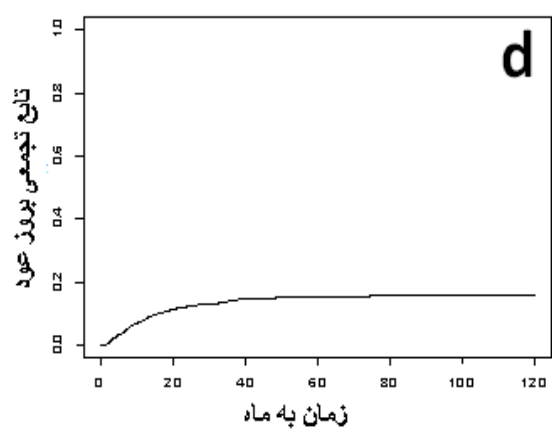
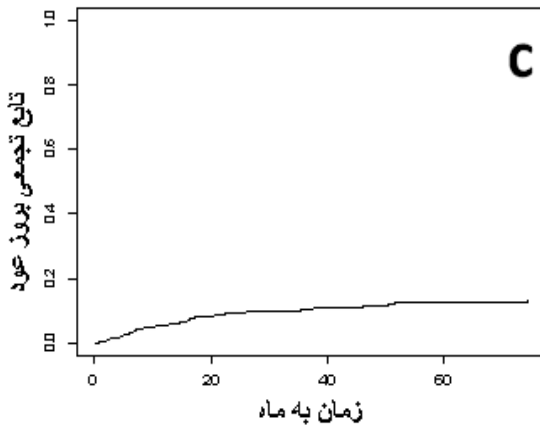
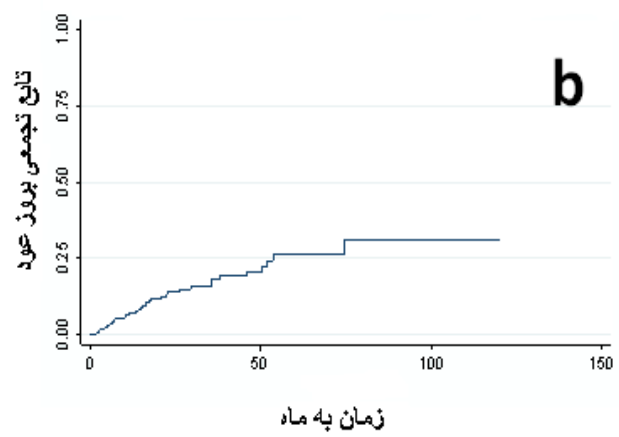
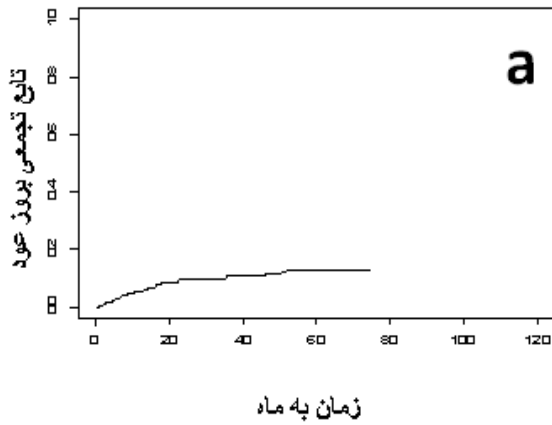
جدول ۱- اطلاعات توصیفی بیماران مبتلا به سرطان معده پس از عمل جراحی

متغیرها	وضعیت	تعداد(درصد)	
جنس	مرد	۲۲۸ (۶۹/۱)	
	زن	۱۰۲ (۳۰/۹)	
سن	کوچکتر و مساوی ۶۰ سال	۸۷ (۲۶/۴)	
	بزرگتر از ۶۰ سال	۲۴۳ (۷۳/۶)	
وجود متاستاز و محل آن	بلی	۱۲۸ (۳۸/۸)	
	غدد لنفاوی	۱۸ (۵/۵)	
	کبد	۲۸ (۸/۵)	
	دوردست	۳ (۰/۹)	
	غدد لنفاوی و کبد	۱۲ (۳/۶)	
	غدد لنفاوی و دوردست	۳ (۰/۹)	
	کبد و دوردست	۱۳۸ (۴۱/۸)	
	خیر	۱۴۵ (۴۳/۹)	
	محل درگیری معده با تومور	آنتروم	۶۳ (۱۹/۱)
	کاردیا	۱۲۲ (۳۷/۳)	
نوع بیماری براساس گزارش پاتولوژی	آدنوکارسینوما	۲۸۱ (۸۵/۲)	
	سایر	۴۹ (۱۴/۸)	
مرحله بیماری	I	۲۲ (۶/۷)	
	II	۶۰ (۱۸/۲)	
	III	۵۴ (۱۶/۴)	
	IV	۱۹۴ (۵۸/۸)	
دریافت درمان تکمیلی	بلی	۲۶۳ (۷۹/۷)	
	خیر	۶۷ (۲۰/۳)	



جدول ۲- برآورد ضرایب رگرسیون مدل Fine و Gray به همراه p-value برای عودبیماری پس از عمل جراحی

متغیرها	$\beta$	$s.e(\beta)$	p-values	HR
جنس	۰/۱۳	۰/۳۳	۰/۷۰	۱/۱۴(۰/۶، ۲/۱۷)
سن	۰/۶۲	۰/۳۸	۰/۱۱	۱/۸۶(۰/۸۸، ۳/۹۱)
وجود متاستاز	۰/۰۲	۰/۶۴	۰/۹۸	۱/۰۲(۰/۲۹، ۳/۵۸)
محل متاستاز	۰/۰۶	۰/۰۴	۰/۱۷	۱/۰۶(۰/۹۸، ۱/۱۵)
محل درگیری معده با تومور	۰/۰۵	۰/۱۹	۰/۷۹	۱/۰۵(۰/۷۲، ۱/۵۳)
نوع بیماری	-۰/۴۸	۰/۵۱	۰/۳۵	۰/۶۲(۰/۲۳، ۱/۶۸)
مرحله بیماری	-۰/۱۶	۰/۲۹	۰/۵۹	۰/۸۵(۰/۴۸، ۱/۵)
دریافت درمان تکمیلی	۰/۳۳	۰/۱۳	۰/۰۱	۱/۳۹(۱/۰۸، ۱/۷۹)



شکل ۱- تابع تجمعی بروز مشاهده شده و بدست آمده از هر سه مدل برای حادثه مورد نظر (عود)

## Reference

- Abrahamowicz, M. and MacKenzie, T.A., 2007. Joint estimation of time-dependent and non-linear effects of continuous covariates on survival. *Statistics in Medicine*. 26, pp. 392-408.
- Abrahamowicz, M., MacKenzie, T. and Esdaile, JM., 1996. Time-dependent hazard ratio: modeling and hypothesis testing with application in lupus nephritis. *Journal of the American Statistical Association*. 91, pp. 1432-1439.
- Altshuler, B., 1970. Theory for the measurement of competing risks in animal experiments. *Mathematical Biosciences*. 6, pp. 1-11.
- Belot, A., Abrahamowicz, M., Remontet, L. and Giorgi, R., 2010. Flexible modeling of competing risks in survival analysis. *Statistics in Medicine*. 29, pp. 2453-2468.
- Bernoulli, D., 1760. Essai d'une nouvelle analyse de la mortalité causée par la petite Vérole, et des avantages de l'Inoculation pour la prévenir. *Historic avec les mémoires Academie Royal des Sciences Paris*. pp. 1-45.
- Borgan, Q., 2005. *Nelson-Aalen Estimator*. *Encyclopedia of Biostatistics*. John Wiley and Sons, Inc.
- Breslow, N.E., 1970. A generalized Kruskal-Wallis test for comparing K samples subject to unequal pattern of censorship. *Biometrika*. 57, pp. 579-594.
- Cox, D.R., 1972. Regression models and life tables. *J. Royal Statistical Society*. 34(B), pp. 187-202.
- Cox, D.R., 1975. Partial likelihood. *Biometrika*. 62(2), pp. 269-276.
- De Boor, C., 1978. *A Practical Guide to Splines*. New York.
- Durrleman, S. and Simon, R., 1989. Flexible regression models with cubic splines. *Statistics in Medicine*. 8, pp. 551-561.
- Efron, B., 1988. Logistic regression, survival analysis, and the Kaplan Meier curve. *Journal of the American Statistical Association*. 83, pp. 414-425.
- Fine, J.P., 1999. Analyzing competing risks data with transformation models. *Journal of the Royal Statistical Society*. 61(B), pp. 817-830.
- Fine, J.P. and Gray, R.J., 1999. A proportional hazards model for the sub distribution of a competing risk. *Journal of the American Statistical Association*. 94, pp. 496-509.
- Gehan, E.A., 1965. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika*. 52, pp. 203-223.
- Giorgi, R., Abrahamowicz, M., Quantin, C., Bolard, P., Esteve, J., Gouvernet, J. and Faivre, J., 2003. A relative survival regression model using B-spline functions to model non-proportional hazards. *Statistics in Medicine*. 22, pp. 2767-2784.
- Hess, KR., 1994. Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. *Statistics in Medicine*. 13, pp. 1045-1062.
- Hoem, J., 1969. Markov Chain Models in Life Insurance. *Blätter der DGVM*. 9(2), pp. 91-107.
- Kalbfleisch, J.D. and Prentice, R.L., 1980. *The Statistical Analysis of Failure Time Data*. New York, John Wiley & Sons, Inc.
- Kaplan, E.L. and Meier, P., 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*. 53, pp. 457-481.
- Klein, J.P., 2010. Competing risks. *John Wiley & Sons, Inc* 2(1), 333-339.
- Lunn, M. and McNeil, D., 1995. Applying Cox regression to competing risks. *Biometrics* 51, pp. 524-532.
- Mantel, N., 1966. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Report*. 50, pp. 163-170.
- Pepe, MS. and Mori, M., 1993. Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? *Statistics in Medicine*. 12, pp. 737-751.
- Pintilie, M., 2006. *Competing Risks: A Practical Perspective*. England, John Wiley & Sons.

- Prentice, RL., Kalbfleisch, JD., Peterson, Jr AV., Flournoy, N., Farewell, VT. and Breslow, NE., 1978. The analysis of failure times in the presence of competing risks. *Biometrics*, 34, pp. 541-554.
- Remontet, L., Bossard, N., Belot, A. and Esteve, J., 2007. An overall strategy based on regression models to estimate relative survival and model the effects of prognostic factors in cancer survival studies. *Statistics in Medicine*. 26, pp. 2214-2228.
- Sleeper, LA. and Harrington, D., 1990. Regression splines in the cox model with application to covariate effects in liver disease. *Journal of the American Statistical Association*. 85, pp. 941-949.
- Sverdrup, E., 1965. Estimates and test procedures in connection with stochastic models for deaths, recoveries and transfers between deferent states of health. *Skandinavisk Akuarietidskrift*. 48, pp. 184-211.
- Tai, BC., Machin, D., White, I. and GebSKI, V., 2001. Competing risks analysis of patients with osteosarcoma: a comparison of four different approaches. *Statistics in Medicine* 20, pp. 661-684.
- Zeraati, H., Mahmoudi, M., Kazemnejad, A. and Mohammad, K., 2005. Postoperative life expectancy in gastric cancer patients and its associated factors. *Saudi Medical Journal*. 26, pp. 1203-1207.

## Survival Analysis of Gastric Cancer Patients following Surgery Based on a Flexible Model in Competing Risks

Nazempour, M., M.Sc. Student, Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

Mahmoudi, M., Ph.D. Professor, Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran-Corresponder author: mahmoodim@tums.ac.ir

Zeraati, H., Ph.D. Associate Professor, Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

Rahimi Froushani, A., Ph.D. Associate Professor, Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

Received: Aug 5, 2012

Accepted: Jan 15, 2013

### ABSTRACT

**Background and Aim:** In many diagnostic studies, including investigation of survival of patients with gastric cancer, where an individual may, following surgery, experience more than one type of event and the occurrence of one type of event hinders the occurrence of other types of events, the question of competing risk arises. To check the effect of each covariate on the occurrence of any event and estimate the hazard function, models such as Cox and Fine-Gray are used. In cases where the assumptions of the two models do not hold, using them will be incorrect. One way to overcome this problem is to use models that have a higher flexibility.

**Materials and Methods:** Data on demographic, clinical and therapeutic characteristics of 330 patients with gastric cancer who had been referred to the Cancer Institute of Iran Imam Khomeini Hospital, Tehran, Iran between January 1996 and April 2000 and undergone surgery were collected. Variables such as type and time of occurrence of the first event (locoregional recurrence and death) were also recorded for each patient from their medical records. The cumulative hazard function of relapse of disease was plotted (R (2.14.0)) software using three models, namely Cox, Fine-Gray and the flexible model and checked against the observed cumulative incidence function of recurrence of the disease and, finally, their performance was evaluated.

**Results:** Almost for all events, the proportionality assumption held for all the variables. Based on the graph of cumulative incidence function for the event of interest (recurrence), it could be seen that the Cox model had overestimated the cumulative incidence function and that the curves of the two other models were very similar to each other and also similar to the observed curve. However, the cumulative incidence function of the flexible model was smoother than the others.

**Conclusion:** In the competing risk framework, the Cox model is not very useful in practice, while it seems that the flexible model is not only a good alternative to the Fine-Gray model but it is also superior to it when the assumption of proportionality does not hold.

**Key words:** Survival analysis, Competing risk, cumulative index function (CIF), time-dependent effects, Gastric cancer