مقاله پژوهشی

مقایسه عملکرد مدل کاکس و روش K ــ نزدیکترین همسایگی در تخمین بقای بیماران پیوند کلیه

دکتر جواد فردمال* ، طاهره امیدی ، دکتر جلال پورالعجل* ، دکتر قدرت اله روشنایی*****

دریافت: ۹۴/۵/۷ پذیرش: ۹۴/۹/۱۴

چکیده:

مقدمه و هدف: روش رایج در برآورد بقا، مدل کاکس است که اعتبار نتایج آن، به پذیره مخاطرات متناسب وابسته است. روش k- نزدیکترین همسایگی یک روش ناپارامتری برای احتمالات بقا در جوامع ناهمگن می‌باشد. هدف این مطالعه مقایسه کارایی مدل کاکس و روش k- نزدیکترین همسایگی (KNN) است.

روش کار: این مطالعه کوهورت گذشته‌نگر بر روی ٤٧٥ بیمار دریافت کننده پیوند کلیه طی سال‌های ۱۳۹۰-۱۳۷۳ شهر همدان می‌باشد. اطلاعات از پرونده‌های پزشکی بیماران استخراج شد. مدت زمان بین پیوند کلیه و رد برگشت ناپذیر پیوند به عنوان پاسخ در نظر گرفته شد. برای مدل‌سازی داده‌ها، از مدل کاکس و روش KNN و برای مقایسه کارایی مدل‌ها از خطای پیش‌بینی نمره بری‌یر استفاده شد.

نتایج: از ٤٧٥ گیرنده پیوند، ۵۵ نفر (۱۱/۵۰٪) رد پیوند داشتند. میزان بقای ۵، ۱۰ و ۱۵ سال به ترتیب ۹۱/۷۰٪، ۸٤/۹۰٪ و ۷٤/۵۰٪ بدست آمد. تعداد همسایگی بهینه با روش اعتبار سنجی متقاطع برابر ٤۵ بدست آمد. نمره بری‌یر برای الگوریتم KNN در زمان‌های ۵، ۱۰ و ۱۵ سال ۰/۰۰۳، ۰/۰۰۶و ۰/۰۰۷و برای مدل کاکس به ترتیب ۰/۰۳۶، ۰/۰۵۸ و ۰/۰۶۹بدست آمد. روش KNN با تعداد همسایگی ٤۵ خطای پیش‌بینی کمتری در زمان‌های ۵، ۱۰ و ۱۵ سال نسبت به مدل کاکس دارد که نشان می‌دهد این روش عملکرد بهتری دارد.

نتیجه نهایی: نتیجه این مطالعه نشان می‌دهد که پیش‌بینی روش KNN نسبت به مدل کاکس زمانی که حجم نمونه بالا و تعداد متغیرهای پیشگو زیاد است، دقت بالاتری دارد.

کلید واژه ها: پیوند کلیه / رد برگشت ناپذیر پیوند / روش کاکس / روش K- نزدیکترین همسایگی

مقدمه:

پیوند کلیه یک روش درمانی مناسب و موثرترین شیوه برای درمان بیماران مبتلا به بیماری پیشرفته کلیه است و باعث افزایش کیفیت زندگی و کاهش خطر مرگ و میر بیماران در مرحله نهایی نارسایی کلیه می‌شود (۱،۲). پیوند کلیه از نظر اقتصادی و هم از نظر زمان نسبت به دیالیز مقرون به صرفه است (۳). در بیماران مبتلا به نارسایی مزمن کلیوی، تصمیم‌گیری‌های پزشکی و برآورد طول عمر بیماران اغلب تحت تاثیر عوامل تشخیصی بیماران است. روش رایج برای بدست آوردن ارتباط عوامل خطر با مخاطره و برآورد احتمالات بقا، مدل کاکس است. اگر چه مدل کاکس یک روش ارزشمند در تحقیقات اپیدمیولوژی و مراقبت‌های پزشکی است، اما برقراری پذیره مخاطره متناسب ضروری است (۴). مدل کاکس به عنوان معیار استاندارد برای مدل‌سازی اثر متغیرهای مستقل روی مخاطره به‌کار می‌رود. مدل کاکس زمانی که حجم نمونه بالا و تعداد متغیرهای پیشگو زیاد است، امکان بررسی فرض متناسب بودن مخاطرات برای تک تک متغیرها دشوار می‌شود (۵). پس نادیده گرفتن فرض مخاطره متناسب و بکار گرفتن روش‌های مبتنی بر مدل

* دانشیار آمار زیستی، مرکز تحقیقات مدل سازی بیماریهای غیر واگیر دانشگاه علوم پزشکی همدان

** کارشناسی ارشد آمار زیستی دانشگاه علوم پزشکی همدان

*** استادیار آمار زیستی، مرکز تحقیقات مدل سازی بیماریهای غیر واگیر دانشگاه علوم پزشکی همدان (gh.roshanaei@umsha.ac.ir)

کاکس، باعث می‌شود که برخی کوریت‌ها به اشتباه برای پیش‌گویی کردن بقا مهم شناخته شوند و برخی دیگر از کوریت‌های مهم، به عنوان عوامل تاثیرگذار شناسایی نشوند. بنابراین برای برخورد با نقض پذیره مخاطره متناسب راه‌کارهایی وجود دارد که می‌توان به استفاده از متغیرهای وابسته به زمان، استفاده از مدل زمان شکست شتاب‌دار یا مدل مخاطره جمعی، وارد کردن کوریت‌هایی که پذیره مخاطره متناسب ندارند به عنوان عوامل طبقه-بندی و تقسیم‌بندی آن متغیر درون فواصلی که پذیره مخاطره متناسب درون هر فاصله برقرار باشد، نام برد (۶). همه این استراتژی‌ها زمان‌بر و نیاز به مدل‌سازی به منظور تجزیه و تحلیل مجموعه داده‌ها، کنترل فرضیات ساختاری مرتبط و انتخاب مدل مناسب دارند. روش k- نزدیکترین همسایگی (KNN) به منظور پیش‌بینی احتمالات بقا برای داده‌های سانسور از راست بکار رفته است که نیاز به بررسی فرض‌ها ندارد. با تشکیل K- نزدیکترین همسایگی احتمالات بقا تعدیل خواهد شد. برای این روش می‌توان مزایایی از جمله عدم نیاز به به شکل تابعی، عدم نیاز به فرض مخاطرات متناسب، سادگی محاسبات و ارایه مستقیم احتمال بقا به جای برآورد مخاطره برشمرد (۷).

در این مقاله، روش k- نزدیکترین همسایگی که مبتنی بر منحنی کاپلان- مایر بر اساس بقای مشاهده شده از K مشاهده برای پیش‌بینی احتمال بقا در جوامع ناهمگن معرفی می‌شود. با تشکیل این K همسایگی احتمالات بقا تعدیل خواهد شد. این روش برای اولین بار بر روی این داده‌ها انجام شده است. در مطالعه ساعتچی و همکاران (۸) از مدل کاکس برای تعیین عوامل تاثیر گذار بر روی بقای پیوند کلیه استفاده شده است.

روش KNN و مدل کاکس برای پیش‌بینی احتمال بقای بیمارانی که پیوند کلیه دریافت کرده‌اند به کار می‌رود و هدف اصلی از این مطالعه، مقایسه عملکرد و انعطاف پذیری دو روش KNN و مدل کاکس زمانی که مدل به فرض مخاطرات متناسب نیاز ندارد، می‌باشد.

روش کار:

در این مطالعه کوهورت گذشته‌نگر، از داده‌های پیوند کلیه ۴۷۵ نفر از افرادی که طی سال‌های ۱۳۹۰-۱۳۷۳ تحت عمل پیوند قرار گرفته‌اند، استفاده شده است. اطلاعات استخراج شده از پرونده‌های پزشکی بیماران با استفاده از چک لیست شامل سن، جنس اهداکنندگان و دریافت کنندگان، نوع اهدا کننده (زنده یا مرده)، رابطه خانوادگی، گروه خونی جنس اهداکنندگان و دریافت کنندگان کلیه، طول مدت دیالیز قبل از پیوند، سطح کراتنین در ترشحات، وضعیت ابتلا به عفونت هپاتیت B، کلیه راست‌یا چپ، اسکیمی سرد کلیه پیوندی، مدت زمان بستری در بیمارستان، داروی‌سرکوب کننده سیستم ایمنی بدن، وقوع رد حاد یا خیلی حاد پیوند، میزان هموگلوبین، علت ابتلا به مرحله نهایی بیماری کلیه جمع‌آوری شد. این چک لیست با استناد به فرم استاندارد معرفی گیرنده و دهنده پیوند در پرونده بیماران تهیه شده است. در این مطالعه روش انتخاب نمونه به صورت سر شماری بود زیرا تنها مرکز پیوند کلیه در استان همدان ابتدا بیمارستان اکباتان و سپس بیمارستان شهید بهشتی بوده و جامعه آماری تمامی بیمارانی که از سال ۱۳۷۳ تا زمان انجام این مطالعه که تا پایان سال ۱۳۹۰ در همدان عمل پیوند کلیه را انجام داده‌اند می‌باشند. بیمارانی که جهت پیگیری به این مراکز مراجعه نکردند از طریق تماس تلفنی از وضعیت بیمار کسب اطلاع شد. پیوند کلیه درمان انتخابی اکثریت بیماران مبتلا به نارسایی پیشرفته مزمن کلیه است. مرحله‌ایی است که کاهش غیر قابل برگشت عملکرد کلیه بیشتر از ۳ ماه طول‌کشیده باشد را نارسایی پیشرفته‌مزمن کلیه گویند و به مرحلهٔ پیشرفته CKD که جهت ادامه حیات به دیالیز یا پیوند نیاز باشد را مرحله انتهایی بیماری کلیه یا End Stage Renal Disease (ESRD) گویند.

معیارهای ورود به مطالعه شامل تمام بیمارانی که به مرحله نهایی بیماری کلیه رسیده‌اند و عمل پیوند را برای اولین بار انجام داده‌اند و معیارهای خروج شامل کلیه بیمارانی که بعد از عمل پیوند هیچ گونه اطلاعاتی درباره آنها موجود نباشد و نیز بیمارانی پیوند مجدد داشته‌اند، از مطالعه خارج شده‌اند. بیمارانی که از مطالعه خارج یا قبل از عمل پیوند به هر دلیل فوت شده‌اند به عنوان سانسور در نظر گرفته شده‌اند. رد برگشت ناپذیر پیوند به عنوان رخداد مورد نظر و طول زمان بین پیوند کلیه تا رد برگشت ناپذیر پیوند به عنوان پاسخ در نظر گرفته شده است.

مدل کاکس: مدل کاکس شامل دو جزء است: یک جزء ناپارامتری که بر اساس منحنی بقا کاپلان مایر مبتنی بر جمعیت است (S_0) و جزء پارامتری مدل کاکس، یک مدل خطی که تابعی نمایی از مجموع پیشگوها، x_i و پارامترها،

β_i است (۹). تابع بقا برآورد شده مدل کاکس برای فرد i ام در زمان t بصورت زیر تعریف می‌شود:

$$\hat{S}_t^{cox}(t,X) = \left[\hat{S}_0(t)\right]^{e^{\sum_{i=1}^{p}\beta_i X_i}}$$

در این مطالعه برای بررسی فرض مخاطره متناسب کاکس از آزمون باقیمانده‌های شونفلد استفاده شد.

روش K- نزدیکترین همسایگی: فرض کنید که j یک مشاهده جدید باشد. در این روش یک پیش‌بینی برای منحنی بقای j- امین فرد با استفاده از منحنی کاپلان مایر از زمان‌های بقا مشاهدات مشابه موجود در مجموعه داده‌ها ایجاد می‌شود. همسایگی‌ها بر اساس متغیرهای مستقل که با مشاهده مورد نظر بر اساس معیار فاصله ماهالانوبیس بیشترین شباهت را دارند تعیین می‌شوند (۱۰). تعداد همسایگی‌های بهینه با روش اعتبار سنجی متقاطع برابر ۴۵ بدست آمد (۱۱). هر مشاهده در ارتباط با مجموعه‌ای از پیشگوها است. $t_1,\ldots, t_n$ زمان‌های بقا مشاهده شده برای n فرد می‌باشد و $\delta_1, \ldots, \delta_n$ وضعیت نهایی فرد بر حسب وقوع حادثه مورد نظر می‌باشد.δ_i=۱ به معنای این است که فرد واقعه مورد نظر را تجربه کرده است و δ_i=۰ به معنای این است که برای فرد واقعه مورد نظر رخ نداده است. در روش نزدیکترین همسایگی، مجموعه‌ای از پیشگوها، x_i ، در ارتباط با مشاهده جدید، x_j ، هستند. فرض کنید $d(x_i.x_j)$ یک فاصله متریک روی پیشگوها است که اندازه شباهت بین مشاهدات x_i و x_j را اندازه‌گیری می‌کنند. بدین صورت که تعداد همسایه‌ها حول مشاهده جدید با حداقل فاصله انتخاب می‌شوند. روش انتخاب نزدیکترین همسایه از مشاهده جدید در مجموعه داده‌ها، بر طبق فاصله متریک $d(x_i.x_j)$ می‌باشد. مجموعه مشاهدات k- نزدیکترین همسایگی حول مشاهده جدید، تشکیل یک مجموعه کوچکتری را می‌دهد که این مجموعه از مشاهدات مشابه زیر مجموعه‌ای از مجموعه داده اصلی می‌باشد. منحنی بقا کاپلان مایر از مشاهدات در مجموعه کوچکتر از داده‌ها بصورت زیر است:

$$\hat{S}_j^{knn}(t) = \prod_{j:t_j \le t}\left(1 - \frac{d_j}{n_j}\right)$$

که d_j تعداد افرادی که در زمان $t_{(j)}$ فوت شده‌اند و n_j تعداد افراد در معرض خطر قبل از زمان t_j می‌باشد. در این مطالعه تعداد همسایگی برابر ۴۵ بر اساس معیارهای مناسبت مدل انتخاب شد.

روش مقایسه مدل‌ها: برای ارزیابی کارایی مدل‌های مختلف بقا چندین روش وجود دارد. یکی از این روش‌ها، نمره بری‌یر(Brier Score) است. اختلاف میانگین مربع از خطای پیش‌بینی اندازه‌گیری شده، به عنوان نمره بری‌یر شناخته می‌شود. مقدار نمره بری‌یر بین صفر و یک است. حداقل نمره بری‌یر، نشان دهنده پیش‌بینی دقیق‌تر است یعنی اینکه هر چه مقدار نمره بری‌یر به صفر نزدیکتر باشد حاکی از بهتر بودن مدل است و بدست آوردن شاخص نمره بری‌یر و رسم نمودار آن در طول زمان معیار مفیدی برای ارزیابی قدرت پیش‌بینی مدل‌های بقا بکار رفته است (۱۲-۱۴).

فرمول کلی برای نمره بری‌یر بصورت زیر است:

$$BS(t,\hat{S}) = E(Y_i(t) - \hat{S}(t|X_i))^2$$

که $Y_i(t)$ وضعیت بقای فرد مشاهده شده در زمان t است و $\hat{S}(t|X_i)$ احتمالات بقا پیش‌بینی شده در زمان t برای فرد i ام با متغیرهای پیش‌بین x_i است. تحلیل داده‌ها با استفاده از نرم‌افزار R با نسخه ۳,۱,۲ آنالیز شد.

نتایج:

در این مطالعه تعداد ۴۷۵ بیمار که طی سال‌های ۱۳۹۰-۱۳۷۳ پیوند کلیه دریافت کرده بودند، وارد مطالعه شدند. از ۴۷۵ گیرنده پیوند، ۲۴۴ نفر (۵۱/۴۰٪) مرد و ۲۳۱ نفر (۴۸/۶۰٪) زن بودند. از ۴۷۵ دهنده پیوند، ۳۸۳ نفر (۸۰/۶۰٪) مرد و ۹۲ نفر (۱۹/۴۰٪) زن بودند. میانگین (انحراف معیار) سنی دهندگان و گیرندگان به‌ترتیب ۲۸/۰۸ (۵/۶۳) و ۳۷/۲۹ (۱۲/۰۹) سال بود. ۳۴ نفر (۷/۲۰٪) کلیه سمت راست و ۴۴۱ نفر (۹۲/۸۰٪) کلیه سمت چپ را پیوند داده بودند. مقادیر کراتینین و هموگلوبین زمان ترخیص به ترتیب ۱/۳۸ (۰/۹۲) و ۱۰/۶۸ (۱/۵۹) و میانگین و انحراف معیار زمان پیگیری بیماران ۶/۲۶ (۴/۷۰) سال بود. از ۴۷۵ بیمار، ۵۵ نفر (۱۱/۵۰٪) رد پیوند رخ داده است. سایر اطلاعات بیماران در جدول ۱ آمده است.

جدول ۲ و شکل ۱ برآورد بقای کاپلان مایر بیماران پیوند کلیه را نشان می‌دهد.

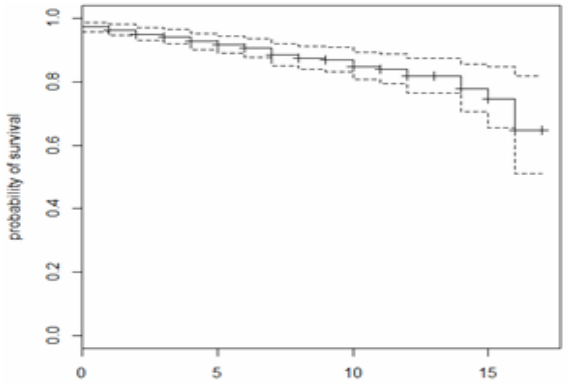
جدول ۱: توزیع فراوانی متغیرهای مورد بررسی در بیماران پیوند شده

	زیر گروه‌ها	تعداد (درصد)
جنس دهنده	مرد	۳۸۳ (۸۰/۶۰)
	زن	۹۲ (۱۹/۴۰)
جنس گیرنده	مرد	۲۴۴ (۵۱/۴۰)
	زن	۲۳۱ (۴۸/۶)
گروه خونی دهنده	O	۱۶۷ (۳۵/۲۰)
	A	۲۱۶ (۴۵/۵۰)
	B	۸۷ (۱۸/۳۰)
	AB	۵ (۱/۱۰)
گروه خونی گیرنده	O	۱۲۶ (۲۶/۵۰)
	A	۲۳۶ (۴۹/۷۰)
	B	۸۹ (۱۸/۷۰)
	AB	۲۴ (۵/۱۰)
استفاده از داروهای سرکوبگر	ایموران،پریدنیزولونسیکلوسپورین	۱۵۴ (۳۲/۴)
	سلسپت،پریدنیزولون، سیکلوسپورین	۳۲۱ (۶۷/۶)
رابطه خویشاوندی دهنده با گیرنده	ندارد	۴۵۷ (۹۶/۲۰)
	دارد	۱۸ (۳/۸۰)
سمت کلیه اهدا شده	راست	۳۴ (۷/۲۰)
	چپ	۴۴۱ (۹۲/۸۰)
نوع دهنده	زنده	۴۶۱ (۹۷/۱۰)
	جسد	۱۴ (۲/۹۰)
شرایط بیمار قبل از پیوند	بدون عارضه	۴۴۹ (۹۴/۵۰)
	رد حاد پیوند	۲۳ (۴/۸۰)
	رد فوق حاد پیوند	۳ (۰/۶۰)
	دیابت	۱۱ (۲/۳۰)
	پر فشار خون	۸۷ (۱۸/۳۰)
	گلومرونفریت	۵۰ (۱۰/۵۰)
علت ابتلا به مرحله نهایی بیماری کلیه	کلیه پلی کیستیک	۸ (۱/۷۰)
	اختلالات ادراری	۱۰۵ (۲۲/۱۰)
	مادرزادی	۶ (۱/۳)
	نامعلوم	۱۷۲ (۳۶/۲۰)
	عوامل دیگر	۳۶ (۷/۶۰)
وضعیت ابتلا به عفونت هپاتیت (B) HBSAg	مثبت	۳ (۰/۶۰)
	منفی	۴۷۲ (۹۹/۴۰)
سن گیرنده (سال)*	۳۷/۳ (۰/۵)	
سن اهداکننده (سال)*	۲۸/۰۸ (۵/۶)	
وضعیت آخرین کراتنین(mg/dl)*	۱/۳۸ (۰/۹۲)	
مدت زمان بستری در بیمارستان(روز)*	۲۷/۲۳ (۸/۷)	
اسکیمی سرد کلیه پیوندی(دقیقه)*	۳۶/۸۶ (۴/۹)	
میزان هموگلوبین(mg/dl) *	۱۰/۶۸ (۱/۵۹)	

* میانگین (انحراف معیار)

جدول ۲: میزان بقای بیماران پیوند کلیه از زمان پیوند تا رد برگشت ناپذیر پیوند با استفاده از کاپلان مایر

	احتمال	انحراف	فاصله اطمینان ۹۵٪	
زمان(سال)	بقا	استاندارد	حد پایین	حد بالا
۱-۰	۰/۹۷۳	۰/۰۰۷	۰/۹۵۸	۰/۹۸۷
۲-۱	۰/۹۶۳	۰/۰۰۹	۰/۸۴۶	۰/۹۸۱
۳-۲	۰/۹۵۰	۰/۰۱۰	۰/۹۳۰	۰/۹۷۱
۴-۳	۰/۹۴۲	۰/۰۱۱	۰/۹۲۰	۰/۹۶۵
۵-۴	۰/۹۲۷	۰/۰۱۳	۰/۹۰۲	۰/۹۵۳
۶-۵	۰/۹۱۷	۰/۰۱۴	۰/۸۹۰	۰/۹۴۵
۷-۶	۰/۹۰۶	۰/۰۱۵	۰/۸۷۷	۰/۹۳۷
۸-۷	۰/۸۸۴	۰/۰۱۸	۰/۸۵۰	۰/۹۲۰
۹-۸	۰/۸۷۵	۰/۰۱۹	۰/۸۳۹	۰/۹۱۲
۱۰-۹	۰/۸۷۰	۰/۰۱۹	۰/۸۳۲	۰/۹۰۸
۱۱-۱۰	۰/۸۴۹	۰/۰۲۲	۰/۸۰۷	۰/۸۹۴
۱۲-۱۱	۰/۸۴۰	۰/۰۲۳	۰/۷۹۵	۰/۸۸۸
۱۴-۱۲	۰/۸۱۸	۰/۰۲۸	۰/۷۶۵	۰/۸۷۴
۱۵-۱۴	۰/۷۷۸	.۰/۰۳۸	۰/۷۰۶	۰/۸۵۶
۱۶-۱۵	۰/۷۴۵	۰/۰۴۸	۰/۶۵۶	۰/۸۴۸
۱۸-۱۶	۰/۶۴۶	۰/۰۷۸	۰/۵۱۰	۰/۸۱۷



شکل۱: منحنی بقای پیوند کلیه در بین بیماران پیوند شده مورد مطالعه

همانطور که شکل فوق نشان می‌دهد میانه بقا قابل برآورد نیست ولی برآورد میانگین (انحراف معیار) بقا در این بیماران برابر ۱۴/۷ (۲/۱) سال بوده است.

در ادامه به‌منظور ارزیابی عملکرد روش KNN و مقایسه آن با مدل کاکس و کاپلان مایر، از مقدار خطای پیش‌بینی بر اساس معیار نمره بری‌یر استفاده شد. مقدار نمره بریر در زمان‌های ۵، ۱۰ و ۱۵ سال در روش KNN به ترتیب برابر ۰/۰۰۳، ۰/۰۰۶ و ۰/۰۰۷، در مدل کاکس به ترتیب برابر ۰/۰۳۶، ۰/۰۵۸ و ۰/۰۶۹و در روش کاپلان مایر به ترتیب برابر ۰/۰۵، ۰/۰۷ و ۰/۱ بوده است. با توجه به این‌که کم بودن نمره بری‌یر نشان دهنده کم بودن

میزان خطای پیش‌بینی و بالا بودن دقت آن است، لذا روش KNN نسبت به دو روش دیگر عملکرد بهتری در پیش‌بینی دارد.

همان‌طور که گفته شد روش کاپلان مایر بدون توجه به کووریت، بقا را پیش‌بینی می‌کند و پیش‌بینی بقای بیماران با متغیرهای زمینه‌ای متفاوت در این روش یکسان است ولیکن در روش KNN پیش‌بینی بقا بر اساس کووریت‌های افراد انجام می‌گیرد. برای نشان دادن مزیت این روش نسبت به کاپلان مایر، برای سه بیمار فرضی با متغیرهای پیشگو مختلف که دارای شرایط رد پیوند در سه حالت بدون عارضه، رد حاد پیوند و رد فوق حاد پیوند، بقا به روش KNN برآورد شده است.

حالت اول: حالتی که فرد رد پیوند کلیه را بدون عارضه داشته است : فرض کنید فرد گیرنده پیوند مردی با سن ۴۷ سال بوده است. فرد دهنده پیوند مرد با سن ۲۸ سال است. فرد دهنده و گیرنده پیوند دارای گروه خونی O هستند. فرد گیرنده پیوند از داروی سرکوب کننده ایمنی بدن سلسپت، پریدنیزولون، سیکلوسپورین استفاده کرده است. عفونت هپاتیت (B) نداشته است. علت ابتلا به مرحله نهایی بیماری کلیه، اختلات ادراری بوده است. این فرد به مدت ۶ سال دیالیز شده است. در سال ۲۰۱۰ پیوند کلیه را از دهنده زنده دریافت کرده است و رابطه خویشاوندی بین دهنده‌با گیرنده پیوند وجود نداشته است. کلیه سمت چپ، پیوند شده است. بر اساس روش KNN با در نظر گرفتن همسایگی ۴۵، میانه زمان بقای این فرد برابر با ۱۶ سال و میانگین زمان بقا برابر با ۱۳/۱۳ سال برآورد شد. جدول ۳ پیش بینی احتمالات بقا برای این فرد را نشان می‌دهد.

جدول ۳: پیش بینی میزان بقای فرد رد پیوند کلیه بدون عارضه با استفاده از روش KNN

زمان (سال)	احتمال بقا	انحراف استاندارد	فاصله اطمینان ۹۵٪	
			حد پایین	حد بالا
۱-۰	۰/۹۵۶	۰/۰۳۰۷	۰/۸۹۷	۱/۰۰۰
۴-۱	۰/۹۲۸	۰/۰۴۱	۰/۸۵۳	۱/۰۰۰
۵-۴	۰/۸۸۴	۰/۰۵۸	۰/۷۷۸	۱/۰۰۰
۷-۵	۰/۸۳۵	۰/۰۷۲	۰/۷۰۴	۰/۹۹۰
۱۶-۷	۰/۷۵۹	۰/۰۹۷۸	۰/۵۹۰	۰/۹۷۷
۱۸-۱۶	۰/۴۷	۰/۱۲	۰/۵۵	۰/۹۲

حالت دوم: حالتی که فرد رد حاد پیوند کلیه داشته است: فرض کنید فرد گیرنده پیوند زنی با سن ۲۲ سال بوده است. فرد دهنده پیوند مرد با سن ۴۷ سال است. فرد دهنده و گیرنده پیوند دارای گروه خونی A هستند. فرد گیرنده پیوند از داروی سرکوب کننده ایمنی بدن ایموران، پریدنیزولون، سیکلوسپورین استفاده کرده است. عفونت هپاتیت (B) نداشته است. علت ابتلا به مرحله نهایی بیماری کلیه، پرفشاری خون بوده است. این فرد به مدت۲ سال دیالیز شده است. در سال ۲۰۰۰ پیوند کلیه را از دهنده زنده دریافت کرده است و رابطه خویشاوندی بین دهنده با گیرنده پیوند وجود ندارد. کلیه سمت راست، پیوند شده است. میانه و میانگین زمان بقای این فرد بر اساس روش KNN با در نظر گرفتن همسایگی ۴۵ به ترتیب برابر ۱۲ و ۱۰/۱ سال برآورد شد. جدول ۴ پیش‌بینی احتمالات بقا برای این فرد را نشان می‌دهد.

جدول ۴. پیش‌بینی میزان بقای برای فرد رد حاد پیوند کلیه با استفاده از روش KNN

زمان (سال)	احتمال بقا	انحراف استاندارد	فاصله اطمینان ۹۵٪	
			حد پایین	حد بالا
۱-۰	۰/۹۵۶	۰/۰۳۰۷	۰/۸۹۷	۱/۰۰۰
۲-۱	۰/۹۲۶	۰/۰۴۲	۰/۸۴۷	۱/۰۰۰
۳-۲	۰/۸۹۴	۰/۰۵۱	۰/۷۹۹	۱/۰۰۰
۴-۳	۰/۸۶۲	۰/۰۵۸	۰/۷۵۵	۰/۹۸۴
۶-۴	۰/۸۲۷	۰/۰۶۶	۰/۷۰۹	۰/۹۶۶
۱۱-۶	۰/۷۴۰	۰/۰۸۳	۰/۵۹۵	۰/۹۲۱
۱۲-۱۱	۰/۶۱۷	۰/۱۳۲	۰/۴۰۶	۰/۹۳۸
۱۸-۱۲	۰/۴۶۳	۰/۱۶۶	۰/۲۲۹	۰/۹۳۶

حالت سوم: حالتی که فرد رد فوق حاد پیوند کلیه داشته است: فرض کنید فرد گیرنده پیوند زنی با سن ۶۴ سال بوده است. فرد دهنده پیوند مرد با سن ۳۴ سال است. فرد دهنده و گیرنده پیوند دارای گروه خونی O هستند. فرد گیرنده پیوند از داروی سرکوب کننده ایمنی بدن سلسپت، پریدنیزولون، سیکلوسپورین استفاده کرده است. عفونت هپاتیت (B) نداشته است. علت ابتلا به مرحله نهایی بیماری کلیه، دیابت بوده است. این فرد به مدت۱ سال دیالیز شده است. در سال ۲۰۰۸ پیوند کلیه را دریافت کرده است. از دهنده زنده کلیه پیوندی را دریافت کرده است و رابطه خویشاوندی بین دهنده با گیرنده پیوند وجود ندارد. کلیه سمت چپ، پیوند شده است. میانه و میانگین زمان بقا این فرد بر اساس روش KNN با در نظر گرفتن همسایگی ۴۵ به ترتیب برابر۱۰ و ۷/۸ سال برآورد

شد. جدول ۵ پیش بینی احتمالات بقا برای این فرد را نشان می‌دهد.

جدول ۵. پیش بینی میزان بقای برای فرد رد حاد پیوند کلیه با استفاده از روش KNN

زمان (سال)	احتمال بقا	انحراف استاندارد	فاصله اطمینان ۹۵٪	
			حد پایین	حد بالا
۱-۰	۰/۹۳۳	۰/۰۳۷	۰/۸۶۳	۱/۰۰۰
۲-۱	۰/۸۸۹	۰/۰۴۷	۰/۸۰۲	۰/۹۸۶
۵-۲	۰/۸۶۵	۰/۰۵۱	۰/۷۷۱	۰/۹۷۲
۶-۵	۰/۸۲۸	۰/۰۶۱	۰/۷۱۶	۰/۹۵۷
۷-۶	۰/۷۷۶	۰/۰۷۶	۰/۶۴۰	۰/۹۴۱
۸-۷	۰/۷۱۶	۰/۰۹۱	۰/۵۵۹	۰/۹۱۸
۱۰-۸	۰/۶۵۱	۰/۱۰۳	۰/۴۷۷	۰/۸۸۹
۱۸-۱۰	۰/۴۳۴	۰/۱۹۰	۰/۱۸۴	۱/۰۰۰

نتایج جداول ۵-۳ نشان می‌دهد که با در نظر گرفتن اثر متغیرهای‌پیشگو افراد میتوان پیش بینی بهتری برای افراد با ویژگی مشابه ارایه داد درحالی که روش کاپلان مایر برای هر سه حالت فوق بقای مشابهی را پیش‌بینی می‌کند.

بحث:

با پیشرفت روش‌های جراحی، امروزه میزان بقای پیوند در مقایسه با دهه‌های گذشته افزایش داشته است. بر اساس نتایج حاصل از این مطالعه، میزان بقای ۱، ۵، ۱۰، ۱۵ و ۱۸ ساله به ترتیب ٪۹۷/۳۰، ٪۹۱/۷۰، ٪۸۴/۹۰، ٪۷۴/۵۰، ٪۶۴/۶۰ بدست آمد. بر اساس گزارش اعضای پیوندی ایران، میزان بقای یک ساله‌ی پیوند کلیه در ایران٪ ۹۴/۷۰ گزارش شده است (۱۵) که نتیجه حاصل از این مطالعه، نشان دهنده‌ی این مطلب می‌باشد که میزان بقای یک ساله پیوند کلیه از میانگین کشوری بالاتر است. در مطالعه‌ای که توسط اشرفی و همکاران در اصفهان انجام شد، میزان بقا یک ساله ۹۶٪ گزارش شده است (۱۶) که از بقای یک ساله مطالعه حاضر بیشتر است. ساعتچی و همکاران (۸) مطالعه ای را با عنوان میزان بقا طولانی مدت پیوند کلیه و عوامل مرتبط انجام دادند نتایج این مطالعه نشان داد که مهمترین عوامل برای رد پیوند شامل سن گیرنده پیوند، نوع دهنده (زنده، جسد) و وقوع رد حاد یا خیلی حاد پیوند است و میزان بقای یک، پنج، ده و ۱۸ساله بیماران پیوند شده به ترتیب برابر ۹۷/۱٪، ۹۲/۳٪، ۸۶/۲٪ و ۷۷/۶٪ گزارش شده است. سلطانیان و همکاران در ارومیه مطالعه ای را به جهت تعیین عوامل اثر گذار مرحله اول پیوند کلیه انجام دادند. میزان بقای یک، پنج و ده ساله بیماران به ترتیب برابر ۹۶٪، ۸۹٪ و ۴۶٪ گزارش شده است (۱۷). در مطالعه الحسینی و همکاران در کشور مصر، میزان بقای یک و پنج ساله به ترتیب برابر ۹۳٪ و ۷۵٪ گزارش شده است (۱۸). در این مطالعه، میزان بقای یک و پنج ساله بیشتر از کشور مصر می‌باشد. کی ون و همکاران (۱۹) یک مطالعه مشابه در کره جنوبی بر روی ۶۱۴ بیمار که تحت پیوند کلیه بودند، انجام دادند. مطالعه آنان مشخص کرد که میزان بقای ۱، ۵ و ۱۰ ساله‌ی پیوند به ترتیب، ۹۲/۶۲٪، ۸۲/۳۷٪ و ۷۶/۰۷٪ بود. عوامل تشخیصی و فردی می‌تواند بر بقای پیوند تاثیر گذار باشد. با توجه به یافته‌های مطالعه حاضر، چند عوامل فردی و دموگرافی اثرات قابل توجهی در رد پیوند دارند. مهم‌ترین آنها سن اهد کننده، نوع اهدا کننده (زنده یا فوت) و رد حاد یا بیش از حد حاد پیوند می‌باشد. نتایج مشابهی از جمله تاثیر سن، جنس و اهدا کنندگان و دریافت کنندگان در بقای پیوند کلیه (۲۰،۱۹)، شاخص توده بدن (۲۱،۲۲)، نوع اهدا کننده (زنده یا فوت دهنده دهنده) (۲۳) وجود دارد. در مجموع بر اساس نتایج به دست آمده در این مطالعه و مقایسه با مطالعات مشابه، می‌توان نتیجه‌گیری کرد که بقای پیوند بیماران تحت پیوند کلیه در این مرکز تقریبا مشابه با گزارش‌های سایر مراکز بزرگ پیوند در دنیا می‌باشد. واضح است که تناقض بین یافته‌های مطالعات مختلف وجود دارد. این تغییرات می‌تواند به عوامل مختلف، از جمله همزمان نبودن انجام مطالعه، متفاوت بودن ویژگی‌های فردی و وضعیت بالینی بیماران نسبت داد. همچنین یکسان نبودن تجربه‌های مراکز پیوندی و پزشکان نیز می‌تواند به عنوان علت دیگری برای وجود اختلاف در بقای پیوند مطرح باشد. با این حال، از یافته‌های مهم در مطالعه حاضر این بود که میزان بقای بیماران پیوند کلیه در طول زمان از شروع مطالعه در سال ۱۳۷۳ تا پایان مطالعه در سال ۱۳۹۰ کاهش یافته است.

در این مطالعه برای مدل سازی داده‌ها از روش KNN برای پیش‌بینی احتمال بقا بر اساس منحنی بقای کاپلان-مایر استفاده شد. در مطالعه حاضر احتمال بقا برای مجموعه‌های کوچک از بیماران مشابه، پیش‌بینی شد. روش KNN برای انواع پیش‌بینی به‌کار می‌رود و متکی بر برآوردهای آماری از نمونه‌های کوچک از همسایگی‌ها

است. همیلتون و همکاران (۲۴) از الگوریتم KNN با تعداد همسایگی‌ها ۴، برای پیش‌بینی زمان بقای بر روی ۲۱۶ بیمار مبتلا به سرطان روده بزرگ استفاده کردند. داده‌ها به دو مجموعه تست و آموزش تقسیم شدند. زنده ماندن بیمار تا پایان ۵ سال پیگیری به عنوان زمان سانسور در نظر گرفته شد. برای سه متغیر سن بیمار، مرحله Dukes و وضعیت لوله کوچک، فاصله اقلیدسی اندازه‌گیری شد و تعداد افراد با کوچکترین فاصله دارای بیشترین شباهت در مجموعه آموزش در نظر گرفته شدند. نتایج این مطالعه نشان داد که اختلاف بین زمان بقا مشاهده شده و زمان بقا پیش‌بینی شده و همچنین پراکندگی مقادیر زیاد است. این مطالعه معایب استفاده از این روش برای داده‌های بقا با حجم نمونه و تعداد همسایگی کم را نشان داد. در این حالت ناتوانی در دقت پیش‌بینی پاسخ بیمار بر مجموعه‌های کوچک از ویژگی‌های کلینیکی آشکار شد، بنابراین برای بهبود عملکرد این روش در پیش‌بینی زمان بقا باید از داده‌هایی با حجم نمونه بالا و تعداد متغیرهای پیشگوهای زیاد استفاده شود. پرایس و همکاران (۲۵) از KNN برای پیش‌بینی زمان بقا بیماران مزمن تنفسی استفاده کردند. مطالعه آنان نشان داد که روش نزدیکترین همسایگی در مقایسه با مدل کاکس، عملکرد بهتری با تعداد همسایگی ۲۰ دارد. لوسکی و همکاران (۷) با استفاده از روش KNN بر روی بیماران دریافت کننده پیوند کلیه احتمال بقا را پیش‌بینی کردند. مطالعه آنان مشخص کرد که در جامعه مورد بررسی، روش KNN دارای مزایایی از جمله عدم نیاز به شکل تابعی، عدم نیاز به پذیره مخاطره متناسب و سادگی محاسبات و نیاز به زمان محاسبات کمتر نسبت به مدل کاکس دارد.

از دیگر استراتژی‌های این روش این است که می‌تواند با هر فاصله متریک استفاده شود و با توجه به این‌که مطالعات دیگر نیز عملکرد این معیار فاصله را تایید کرده‌اند در این مطالعه نیز از فاصله ماهالانوبیس استفاده شده است (۲۶). این اندازه‌گیری بدون واحد است و به‌طور گسترده‌ای در تجزیه و تحلیل خوشه‌ای و روش‌های طبقه‌بندی استفاده می‌شود. از جمله محدودیت‌های روش KNN تعیین مقدار پارامتر کنترل است زیرا به شدت می‌تواند کیفیت پیش‌بینی را تحت تاثیر قرار دهد. یک مقدار کوچک از K در پیش‌بینی باعث می‌شود واریانس بزرگ و اریبی کاهش یابد. باید k طوری انتخاب شود که بین اریبی و واریانس تعادل برقرار شود (۲۷). روش KNN بدلیل نداشتن پذیره خاصی، انعطاف بیشتری نسبت به مدل کاکس دارد (۷). در این مطالعه عملکرد پیش‌بینی روش KNN با مدل کاکس مقایسه شد. مدل کاکس پرکاربردترین مدل برای پیش‌بینی زمان بقا است ولی پیش فرض متناسب بودن مخاطرات برای استفاده از این مدل لازم است. در مواجهه با داده‌هایی که حجم بالایی دارند تعداد متغیرهای زیادی دارند امکان بررسی فرض متناسب بودن مخاطرات برای تک تک متغیرها دشوار می‌شود، بنابراین به جای رگرسیون کاکس استفاده از مدل‌هایی که این فرض را نداشته باشند مطلوب است. به هر حال، به خاطر محدودیت‌های مدل کاکس، روش نزدیکترین همسایگی یک جایگزین مناسب برای مدل کاکس است. در مجموعه داده‌های حاضر، نمره بری‌یر برای الگوریتم KNN در زمانهای ۵، ۱۰ و ۱۵ سال ۰/۰۰۳، ۰/۰۰۶و ۰/۰۰۷ و برای مدل کاکس به ترتیب ۰/۰۳۶، ۰/۰۵۸ و ۰/۰۶۹ بدست آمد. روش KNN با تعداد همسایگی ۴۵ خطای پیش‌بینی کمتری در زمان‌های ۵،۱۰و ۱۵ سال نسبت به مدل کاکس دارد که نشان می‌دهد این روش عملکرد بهتری دارد. در این مقاله با وجود این‌که انتخاب مدل تابع متریک ساده باعث عملکرد بهتر آن نسبت به مدل کاکس شد لذا پیشنهاد می‌شود با انتخاب جایگزین برای تابع متریک و انتخاب تابع وزن مناسب انتظار می‌رود عملکرد این روش بهبود بیشتری یابد که برای تحقیقات آینده پیشنهاد می‌شود.

نتیجه نهایی:

نتیجه این مطالعه نشان داد که روشKNN با توجه به این‌که تاثیر متغیرهای پیشگو افراد را در برآورد بقا در نظر می گیرد نسبت به مدل کاکس و کاپلان مای‌ر دارای دقت پیش‌بینی بالاتری است و حتی در صورتی که مفروضات مدل کاکس برقرار نباشد این مدل کارایی لازم را دارد. لذا این روش جایگزین مناسبی برای تحلیل داده‌های بقا است که در اکثر موارد شرایط استفاده از مدل کاکس را ندارند.

سپاسگزاری:

این مقاله برگرفته از پایان‌نامه مصوب شورای پژوهشی دانشگاه علوم پزشکی همدان می‌باشد که بدین‌وسیله از زحمات معاونت محترم تحقیقات و فناوری قدردانی و تشکر می‌گردد. ضمناً منافع شخصی نویسندگان با نتایج این مطالعه ارتباطی نداشته است.

References

1. Meier- Kriesche H, Port F, Ojo A, Rudich S, Hanson J, Cibrik D. Effect of waiting time on renal transplant outcome. Kidney Int 2000; 58(3):1311-7.
2. Shrestha B, Haylor J. Factors influencing long-term outcomes following renal transplantation: a review. JNMA J Nepal Med Assoc. 2007; 46 (167): 136-42.
3. Perovic S, Jankovic S. Renal transplantation vs hemodialysis: cost-effectivencess analysis. Vojnosanit Pregl 2009;66(8):939-44.
4. Cox D. Regression models and life-tables. J Royal Statistical Soc 1972; 34(2): 187–220.
5. Therneau T, Grambsch P. Modeling Survival Data: Extending the Cox Model: Springer; 2000.
6. Aalen O. A linear regression model for the analysis of life times. Stat Med 1989; 8: 907–25.
7. Lowsky D, Ding Y, Lee D, Mculloch C, Ross L. A K-nearest neighbors survival probability prediction method. Stat Med 2013; 32(12): 2062-9.
8. Saatchi M, Poorolajal J, Amirzargar M, Mahjub H, Esmailnasab N. Long-term survival rate of kidney graft and associated prognostic factors: A retrospective cohort study, 1994-2011. Ann Transplant 2013:153-60.
9. Kleinbaum D, Klein M. Survival Analysis: A Self-Learning Text, 2012.
10. De Maesschalck R, Jouan-Rimbaud D, Massart D. The Mahalanobis distance. Chemometrics and Intelligent Laboratory Systems 2000; 50: 1-18.
11. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection proceedings of the 14th international koint conference on artificial intelligence. San Mateo, CA: Morgan Kaufmann 1995;2(12): 1137–43.
12. Brier G. Verification of forecasts expressed in terms of probability. Monthly Weather Rev 1950;78:1–3.
13. Graf E, Schmoor C, Sauerbrei W, Schumacher M. Assessment and comparison of prognostic classification schemes for survival data. Stat Med 1999;18:2529-45.
14. Gerds T, Schumacher M. Consistent estimation of the expected brier score in general survival models with right-censored event times. Biometric J 2006;48(6):1029–40.
15. Iranian network for Organ Procurement. Available from: www.irantransplant. org. ipaddress.com. (Persian)
16. Ashrafi M, Hamidi Beheshti M, Shahidi S, Ashrafi F. Application of artificial neural network to predict graft survival after kidney transplantation: Reports of 22 years follow up of 316 patients in Isfahan. Tehran Univ Med J 2009;67(5):353-9. (Persian)
17. Soltanian A, Mahjub H, Taghizadeeh-Afshari A, Gholami G, Sayyadi H. Identify Survival Predictors of the First Kidney Transplantation: A Retrospective Cohort Study. Iran J Public Health 2015;44(5):683-89.
18. El-Husseini A, Foda M, Shokeir A, Shehab E-DA,Sobh M, Ghoneim M. Determinants of graft survival in pediatric and adolescent live donor kidney transplant recipients: a single center experience. Pediatr Transplan 2005; 9(6): 763-9.
19. Kwon O, Kwak J, Kanga C. The impact of gender and age matching for long-term graft survival in living donor renal transplantation. Transplant Proc 2005;37(2):726–28.
20. Głyda M, Włodarczyk Z, Czapiewski W. Results of renal transplantation from expanded criteria deceased donors-a single-center experience. Ann Transplant 2012;17(1):35-42.
21. Tetaz R, Trocmé C, Roustit M. Predictive diagnostic of chronic allograft dysfunction using urinary proteomics analysis. Ann Transplant 2912;17(3):52–60.
22. Ditonno P, Lucarelli G, Impedovo S. Obesity in kidney transplantation affects renal function but not graft and patient survival. Transplant Proc 2011;4(1):367–72.
23. Cheung C, Chan H, Chan Y. Impact of delayed graft function on renal function and graft survival in deceased kidney transplantation. Hong Kong Med J 2010;16(5):378-82.
24. Hamilton P, Bartels P, Anderson N, Thompson D, Montironi R, Sloan J. Case-based prediction of survival in colorectal cancer patients. Anal Quant Cytol Histol 1999;21(4):283-91.
25. Prijs M, Peelen L, Bresser P, Peek N, 2007;4594:. A nearest neighbor approach to predicting survival time with an application in chronic respiratory disease. Lecture Notes in Computer Science. 2007: 77-86
26. Bobrowski L, Topczewska M. Improving the K-NN classification with the Euclidean distance through linear data transformations Lecture Notes in Computer Science Advances in Data Mining: Springer; 2005: 23-32.
27. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th international joint conference on artificial intelligence. 1995; 2(12):1137–43.

Original Article

Comparison of Cox Model and K-Nearest Neighbor to Estimation of Survival in Kidney Transplant Patients

J. Faradmal, Ph.D.[*]; T. Omidi, M.Sc.[**]; J. Pourolajal, Ph.D.[*]; Gh. Roshanaei, Ph.D.[***]

Received: 29.7.2015 Accepted: 5.12.2015

Abstract

Introduction & Objective: Cox model is a common method to estimate survival and validity of the results is dependent on the proportional hazards assumption. K- Nearest neighbor is a nonparametric method for survival probability in heterogeneous communities. The purpose of this study was to compare the performance of k- nearest neighbor method (K-NN) with Cox model.

Materials & Methods: This retrospective cohort study was conducted in Hamadan Province, on 475 patients who had undergone kidney transplantation from 1994 to 2011. Data were extracted from patients' medical records using a checklist. The duration of the time between kidney transplantation and rejection was considered as the survival time. Cox model and k-nearest neighbor method were used for Data modeling. The prediction error Brier score was used to compare the performance models.

Results: Out of 475 transplantations, 55 episodes of rejection occurred**.** 5, 10 and 15 year survival rates of transplantation were 91.70% , 84.90% and 74.50%, respectively. The number of neighborhood optimized using cross validation method was 45. Cumulative Brier score of *k*-NN algorithm for t=5, 10 and 15 years were 0.003, 0.006 and 0.007, respectively. Cumulative Brier of score Cox model for t=5, 10 and 15 years were 0.036, 0.058 and 0.058, respectively. Prediction error of *k*-NN algorithm for t=5, 10 and 15 years was less than Cox model that shows that the k-NN method outperforms.

Conclusions: The results of this study show that the predictions of KNN has higher accuracy than the Cox model when sample sizes and the number of predictor variables are high.

(Sci J Hamadan Univ Med Sci 2015; 22 (4):300-308)

Keywords: Cox Model / Kidney Transplantation / K- Nearest Neighbor Transplantation Rejection

---

* Associate Professor of Biostatistics, Modeling of Non-Communicable Diseases Research Center Hamadan University of Medical Sciences & Health Services, Hamadan, Iran.

** M.Sc. in Biostatistics & Epidemiology, Hamadan University of Medical Sciences & Health Services, Hamadan, Iran.

*** Assistant Professor of Biostatistics, Modeling of Non-Communicable Diseases Research Center Hamadan University of Medical Sciences & Health Services, Hamadan, Iran. (gh.roshanaei@umsha.ac.ir)