

# مقایسه‌ی مدل‌های رگرسیون اکولوژیکی فضایی بیزی در بررسی بروز سرطان سینه در ایران

احمد حداد خوشکار<sup>۱</sup>، توحید جعفری کشکی<sup>۱،۲</sup>، بهزاد مهکی<sup>۳</sup>

## مقاله پژوهشی

### چکیده

**مقدمه:** در ایران سرطان سینه با بروز استاندارد شده‌ی سنی ۲۸/۲۵ در رتبه‌ی اول سرطان زنان قرار دارد. با توجه به روند پیر شدن جمعیت و سیر صعودی سرطان در ایران، اهمیت پیشگیری از سرطان و وجود پراکندگی سرطان‌ها در مناطق مختلف کشور، شناسایی مناطق دارای بروز بالای سرطان سینه و آگاهی از وضعیت پراکندگی و نقش مهم‌ترین عوامل خطر بالقوه‌ی این سرطان با استفاده از مدل‌های پیشرفته‌ی آماری ضروری می‌باشد.

**روش‌ها:** در این مطالعه‌ی کاربردی اکولوژیکی، جامعه مورد مطالعه، تمام بیماران مبتلا به سرطان سینه بین سال‌های ۱۳۸۴ تا ۱۳۸۷ در ایران می‌باشند. عوامل خطر: مصرف سیگار، مصرف میوه و سبزیجات، فعالیت فیزیکی، اضافه وزن و وضعیت اقتصادی-اجتماعی می‌باشد. جهت برآورد خطر نسبی سرطان سینه، از مدل‌های گاما-پوآسن، لگ-نرمال و (BYM یا Besag, York, Mollie) استفاده شد.

**یافته‌ها:** تعداد کل سرطان سینه ثبت شده در سال‌های ۱۳۸۴ تا ۱۳۸۷ در ایران ۲۵۱۵۲ مورد بوده است. از میان مدل‌های استفاده شده، مدل BYM تعدیل نشده بهترین برازش را داشت. بدون تعدیل اثر عوامل خطر، استان‌های اصفهان و یزد و تهران بیشترین خطر بروز سرطان سینه و سیستان و بلوچستان و چهارمحال و بختیاری کمترین میزان خطر ابتلا را دارند. با تعدیل عوامل خطر، استان‌های خراسان رضوی، خرم‌آباد و همدان بیشترین خطر بروز سرطان سینه و استان‌های اردبیل و کهگیلویه و بویراحمد، کمترین خطر بروز را دارند. ارتباط مستقیم معنی‌داری بین سرطان سینه و شاخص توسعه انسانی مشاهده شد.

**نتیجه‌گیری:** با توجه به برازش بهتر مدل BYM، توصیه می‌شود در مطالعات آینده از مدل BYM استفاده شود و با توجه به معنی‌داری متغیر شاخص توسعه انسانی، در مطالعات مشابه این متغیر در نظر گرفته شود. این مطالعه نشان داد که استان‌های اصفهان، یزد و تهران دارای بیشترین خطر بروز سرطان سینه می‌باشند.

**واژه‌های کلیدی:** نقشه‌بندی بیماری، سرطان سینه، شاخص توسعه انسانی، مدل BYM

**ارجاع:** حداد خوشکار احمد، جعفری کشکی توحید، مهکی بهزاد. مقایسه‌ی مدل‌های رگرسیون اکولوژیکی فضایی بیزی در بررسی

بروز سرطان سینه در ایران. مجله تحقیقات نظام سلامت ۱۳۹۴؛ ۱۱(۳): ۶۲۶-۶۱۳

تاریخ پذیرش: ۱۳۹۴/۰۴/۱۹

تاریخ دریافت: ۱۳۹۳/۱۰/۰۶

Email: Behzad.mahaki@gmail.com

اندام ترشچی است که از بافت غددی و مجاری آن تشکیل شده است. دلیل این سرطان نامشخص است ولی به دلیل نادر بودن این سرطان در مردان از نظر اتیولوژی مشخص می‌شود

### مقدمه

سرطان سینه شایع‌ترین سرطان و شایع‌ترین علت مرگ ناشی از سرطان در زنان ساکن کشورهای غربی است. سینه یک

که دلیل آن می‌تواند هورمون‌های زنانگی باشد. این تومورها رشد آرامی دارند ولی در زمان کوتاهی پس از پیدایششان به مرحله‌ی پیشرفته می‌رسند. تغییرپذیری توزیع این سرطان در مناطق مختلف تأثیر نقش عوامل محیطی را در این سرطان نشان می‌دهد (۱).

بار جهانی سرطان در کشورهای توسعه یافته به طور فزاینده‌ای در حال افزایش است. دلیل این امر از یک سو پیر شدن و رشد جمعیت و از سوی دیگر افزایش رفتارهای مؤثر بر افزایش سرطان مانند مصرف سیگار می‌باشد. بر اساس گزارش GLOBOCAN برآورد شده است که در سال ۲۰۰۸، ۱۲/۷ میلیون مورد سرطان و ۷/۶ میلیون مورد مرگ (حدود ۱۳٪ از کل مرگ‌ها) ناشی از سرطان روی داده است (۲).

در جهان سرطان از مهم‌ترین عوامل مرگ می‌باشد. در کشورهای توسعه یافته سرطان بعد از بیماری‌های قلبی عروقی دومین عامل مرگ و در کشورهای در حال توسعه، سرطان سومین عامل مرگ می‌باشد. حدود ۷۰٪ از کل مرگ‌های ناشی از سرطان در کشورهای با درآمد پایین و متوسط اتفاق می‌افتد. برآورد شده است که در سال ۲۰۳۰ مرگ‌های ناشی از سرطان به ۱۳/۱ میلیون برسد با تشخیص و کنترل عوامل خطر می‌توان جلوی ۴۰٪ از مرگ‌های ناشی از سرطان را گرفت. بنابراین برای مقابله و پیشگیری از سرطان، شناسایی عوامل خطر در هر منطقه امری ضروری می‌باشد (۳).

امروزه سرطان عامل ۴۵ تا ۵۰٪ از مرگ‌ها در زنان ۴۵ تا ۶۴ سال می‌باشد. این نرخ بالا نشان دهنده‌ی افزایش معنی‌دار مرگ ناشی از سرطان نسبت به قرن گذشته می‌باشد. این موضوع می‌تواند دلایل مختلفی داشته باشد. آلودگی هوا، سیگار کشیدن، تغییر رژیم غذایی، فعالیت فیزیکی ناکافی، چاقی، استرس و ... می‌تواند از عوامل خطرزا برای سرطان باشد که در قرن گذشته این عوامل رشد چشمگیری داشته است (۴).

سرطان سینه در زنان با میزان بروز استاندارد شده‌ی ۲۸/۲۵ در رتبه‌ی اول سرطان زنان قرار دارد. این سرطان در استان‌های مختلف دارای پراکندگی یکسانی می‌باشد. این سرطان در همه‌ی استان‌ها به غیر از استان‌های اردبیل و کهگیلویه و بویر احمد در رتبه‌ی اول قرار دارد. در استان اردبیل سرطان سینه بعد از سرطان مری و معده رتبه‌ی سوم و در استان کهگیلویه و بویر احمد بعد از سرطان پوست رتبه‌ی دوم را دارد. میزان بروز استاندارد شده‌ی این سرطان در کشورهای پیشرفته ۸۶/۴ و در کشورهای کمتر توسعه یافته ۲۷/۳، در جهان ۳۸/۹، در آسیا ۲۶ و در آسیای شرقی ۳۲/۶ می‌باشد (۳).

در سال ۲۰۰۴ بروز استاندارد سنی برای مردان ۱۱۰ و برای زنان ۹۸ در ۱۰۰۰۰۰ نفر بوده است و نرخ مرگ‌ومیر ناشی از سرطان برای مردان ۶۵ و برای زنان ۴۱/۱ در ۱۰۰۰۰۰ نفر بوده است (۱).

پیش‌بینی می‌شود که به دلیل افزایش امید زندگی در آینده و گرایش به سبک زندگی غربی نرخ سرطان در آینده افزایش زیادی داشته باشد (۵).

با توجه به بسیاری از مطالعات انجام شده مصرف سیگار، مصرف میوه و سبزیجات، فعالیت فیزیکی، اضافه وزن و وضعیت اقتصادی، اجتماعی از مهم‌ترین عوامل مؤثر و قابل اندازه‌گیری بروز سرطان می‌باشد (۱۵-۵).

مطالعات زیادی بر سرطان سینه و عوامل خطر مؤثر بر آن صورت گرفته است (۱۰، ۱۴، ۱۶-۸، ۷، ۵). این مطالعات در سطح اکولوژیک نمی‌باشند و نقش همسایگی مناطق را در نظر نمی‌گیرند. مطالعات بسیار کمی در سطح اکولوژیک با لحاظ کردن پراکندگی جغرافیایی و نقش همسایگی مناطق انجام شده است (۱۷). هم‌چنین این مطالعات نقش عوامل خطر در بروز سرطان سینه را بررسی نکرده است.

با توجه به روند پیر شدن جمعیت و سیر صعودی سرطان در ایران، اهمیت پیشگیری از سرطان و وجود پراکندگی سرطان‌ها در مناطق مختلف کشور، شناسایی مناطق دارای بروز بالای سرطان سینه و آگاهی از وضعیت پراکندگی و

نقش مهم‌ترین عوامل خطر بالقوه‌ی این سرطان با استفاده از مدل‌های پیشرفته‌ی آماری ضروری می‌باشد. بنابراین در مطالعه‌ی حاضر به بررسی بروز سرطان سینه در سطح استان‌های ایران و نقش عوامل خطر مصرف سیگار، اضافه وزن یا چاقی، فعالیت فیزیکی، مصرف میوه و سبزیجات و شاخص توسعه انسانی (HDI یا Human Development Index) یا با استفاده‌ی مدل‌های بیزی که نقش همبستگی‌های مکانی بین بروز سرطان در نواحی مورد مطالعه را لحاظ می‌کند، پرداخته شده است.

### روش‌ها

این پژوهش از نوع کاربردی - بوم‌شناختی است. دلیل انتخاب این نوع مطالعه، نبود اطلاعات در سطح افراد مورد مطالعه بود. جامعه‌ی مورد مطالعه‌ی این پژوهش تمام بیماران ثبت شده مبتلا به سرطان سینه در استان‌های کشور به صورت سرشماری در فاصله‌ی زمانی بین سال‌های ۱۳۸۴ تا ۱۳۸۷ است که بدون هیچ محدودیتی از نظر سن، جنسیت و سایر متغیرها مورد مطالعه قرار گرفت. این از گزارش‌های سالانه‌ی کشوری ثبت موارد سرطانی که توسط اداره سرطان مرکز مدیریت بیماری‌های غیر واگیر معاونت بهداشتی وزارت بهداشت، درمان و آموزش پزشکی منتشر شده استخراج گردید. ثبت سرطان شیوه‌ای است که در آن به صورت مداوم اطلاعات مربوط به بروز و مشخصات سرطان‌ها جمع‌آوری، ذخیره، تجزیه و تحلیل می‌گردد.

در مطالعه حاضر، اطلاعات مربوط به چهار عامل خطر: مصرف سیگار، مصرف میوه و سبزیجات، فعالیت فیزیکی، اضافه وزن از گزارش‌های سالانه‌ی نظام مراقبت عوامل خطر بیماری‌های غیر واگیر وزارت بهداشت استخراج گردید. متغیر مصرف سیگار حاصل ضرب درصد افراد سیگاری در میانگین تعداد نخ سیگار مصرفی روزانه هر استان، متغیر اضافه وزن یا چاقی به صورت درصد افراد دارای BMI > ۲۵ در هر استان، مصرف میوه و سبزیجات به صورت جمع متوسط تعداد واحد میوه و تعداد واحد سبزیجات مصرفی روزانه در هر استان در نظر گرفته شد. متغیر فعالیت فیزیکی براساس شاخص ترکیبی مت (MET) یا

(Metaolic Equivalent) یا به دست آمده است. شاخص HDI، میانگین هندسی موفقیت‌های به دست آمده در هر استان در سه بعد اصلی توسعه انسانی یعنی زندگی طولانی و سالم، دستیابی به دانش و استانداردهای زندگی آبرومندانه می‌باشد. شاخص توسعه انسانی میانگین هندسی شاخص توسعه‌های نرمالی است که موفقیت‌های هر بعد را اندازه می‌کند. این شاخص از امید به زندگی، تحصیلات، تولید سرانه ناخالص ملی به دست می‌آید و نسبتی بین ۰ و ۱ است و از گزارشات سالانه بانک مرکزی استخراج گردید (۳).

تحلیل بوم‌شناختی یک روش اکتشافی اپیدمیولوژی است که رابطه‌ی شیوع بیماری‌ها و عوامل خطر را که در گروه‌ها به جای افراد اندازه‌گیری می‌شود بررسی می‌کند. ساده‌ترین روش، استفاده از نسبت مرگ و میر استاندارد شده (SMR یا Sample Mortality Rate) می‌باشد که از تقسیم موارد مشاهده شده ( $Y_i$ ) به مقادیر مورد انتظار ( $E_i$ )، به دست می‌آید. این روش از برآورد ماکسیم درست‌نمایی استفاده می‌کند. این میزان دو عیب عمده دارد. اول این که نسبت مرگ و میر استاندارد شده همبستگی فضایی پاسخ‌ها را در نظر نمی‌گیرد، دوم این که پیش‌پراکنندگی موجود در داده‌ها را به حساب نمی‌آورد.

تحلیل داده‌ها و رسم نقشه‌ها با استفاده از نرم‌افزار OpenBUGS 3.2.1 انجام شده است و مقایسه‌ی مدل‌ها با استفاده از شاخص DIC انجام شده است.

### روش‌های آماری:

مدل‌های گاما- پواسن، لگ‌نرمال و BYM زیر مجموعه‌ی مدل‌های بیزی می‌باشد. در داده‌های مربوط به نقشه‌بندی بیماری‌ها به علت کمبود داده‌ها استفاده از روش‌های بیزی که داده‌ها را با اطلاعات پیشین ترکیب می‌کند روش مناسب‌تری نسبت به روش مرسوم SMR می‌باشد.

### مدل‌های بوم‌شناختی:

مدل‌های زیر برای بررسی میزان بروز و همچنین ارزیابی میزان ارتباط بروز با عوامل خطر مورد استفاده قرار گرفت. در این مدل‌ها  $\{Y_i, i = 1, \dots, n\}$  و  $\{E_i, i = 1, \dots, n\}$  به

$$v_i \sim N(0, \tau_v^2)$$

جز تصادفی این مدل  $v_i$  جز تصادفی مدل می‌باشد.

مدل بیز کامل (BYM):

این مدل، علاوه بر متغیرهای مستقل، دو منبع تغییرات برای ناهمگنی میزان بروز در هر ناحیه لحاظ می‌کند و میزان بروز را از طریق رابطه خطی زیر مدل‌بندی می‌کند.

$$\log \theta_i = \alpha + u_i + v_i + \sum_{h=1}^H \beta_h x_{ih}$$

$u_i$  و  $v_i$  عضو تصادفی مدل BYM می‌باشند و به ترتیب نشان‌دهنده ناهمگنی غیرساختاری (پیش پراکندگی) و ناهمگنی ساختاری می‌باشد. برای متغیرهای  $u_i$  و  $v_i$  به ترتیب توزیع پیشین نرمال و توزیع نرمال خودهمبسته‌ی شرطی<sup>۱</sup> و برای ضرایب  $\beta$  ها توزیع پیشین نرمال در نظر گرفته می‌شود. با در نظر گرفتن توزیع خودهمبسته شرطی، میزان بروز در هر ناحیه به نواحی مجاور آن مشروط شده و از میزان بروز آن نواحی تأثیر می‌پذیرد (۱۸-۱۹).

برای اطلاع بیشتر از چگونگی محاسبه‌ی توزیع‌های پسین به مقاله‌ی Besag و همکاران (۲۰) مراجعه نمایید.

شاخص (DIC یا Deviance information criterion) یا یک شاخص برای مقایسه‌ی مناسب بودن مدل‌ها می‌باشد. این شاخص از ترکیب دو شاخص دیگر به دست آمده است. شاخص اول مربوط به انحراف از برازش و شاخص دوم مربوط به تعداد پارامترهای مدل می‌باشد. شاخص DIC هرچه کوچکتر باشد نشان می‌دهد که برازش مدل بهتر است (۲۱).

برای برازش مدل‌ها، توزیع‌های پیشین مناسب به کلیه‌ی پارامترهای نامعلوم مدل اختصاص داده شد و برآوردهای توزیع پسین پارامترها با استفاده از روش‌های (MCMC یا Markov chain Monte Carlo) یا به دست آمد. به منظور آزمون همگرایی اثرات از ابزار تشخیصی بروکس-گلمن-روبین استفاده شد. با توجه به معیارهای بروکس-گلمن-روبین زمان سوختن مدل ۳۰۰۰۰۰ در نظر گرفته شد. آزمون معنی‌داری در مورد پارامترها با استفاده از فواصل باورپذیر بیزی

ترتیب نشانگر تعداد رویدادهای مشاهده شده و تعداد رویدادهای مورد انتظار برای سرطان سینه در  $\Omega$  منطقه  $i$  ام می‌باشد. فرض می‌شود که  $Y_i$  دارای توزیع پواسن با نرخ  $\mu_i = \theta_i E_i$  باشد که در آن  $\theta_i$  خطر نسبی برای منطقه‌ی  $i$  ام است. در مدل‌های زیر  $\sum_{h=1}^H \beta_h x_{ih}$  نشان‌دهنده‌ی ترکیب خطی عوامل خطر  $x_1, x_2, \dots, x_H$ ، با ضرایب  $\beta_h$  متناظر می‌باشد. در همه معادلات زیر  $\alpha$  سطح کلی خطر نسبی است.

مدل رگرسیون پواسن-گاما:

این مدل ارتباط بین میزان بروز و عوامل خطر در هر استان را به صورت لگ خطی زیر در نظر می‌گیرد.

$$\log \theta_i = \alpha + \sum_{h=1}^H \beta_h x_{ih}$$

مهم‌ترین مشکل این مدل در نظر نگرفتن همبستگی فضایی می‌باشد. در بسیاری از موارد تغییرپذیری را نمی‌توان با متغیرهای بوم‌شناختی توضیح داد و تغییرپذیری بیشتر از تغییرپذیری مورد انتظار از مدل پواسن می‌شود. در این شرایط پیش پراکندگی رخ می‌دهد. بنابراین پارامتر پیش‌پراکندگی به مدل اضافه می‌شود. میزان بیش‌پراکندگی به درجه ناهمگنی بین خطرهای نسبی بستگی دارد در ساختار بیزی، برای این مدل فرض می‌شود که  $\theta_i$  دارای توزیع گاما است. لذا توزیع پسین حاصل نیز گاما خواهد بود. این مدل فاقد جز تصادفی می‌باشد.

مدل لگ نرمال:

توزیع پیشین گاما برای نرخ خطر از نظر ریاضی مناسب به نظر می‌رسد ولی این مدل دارای محدودیت‌هایی می‌باشد زیرا سازگاری متغیر مستقل مشکل است و امکان ورود همبستگی فضایی بین نرخ‌های مناطق وجود ندارد. مدل لگ نرمال برای خطر نسبی انعطاف‌پذیرتر می‌باشد که به صورت زیر تعریف می‌شود.

$$\log \theta_i = \alpha + v_i + \sum_{h=1}^H \beta_h x_{ih}$$

در مدل لگ-نرمال با در نظر گرفتن ناهمگنی غیر ساختاری اثرات متغیرها تعدیل شده و عوامل خطر فعالیت فیزیکی و شاخص توسعه‌ی انسانی معنی‌دار شده‌اند. مثبت بودن این دو ضریب نشان می‌دهد که افزایش اضافه وزن و چاقی و شاخص توسعه انسانی باعث افزایش میزان بروز سرطان می‌شود. با در نظر گرفتن ناهمگنی غیر ساختاری اثرات متغیرها تعدیل شده و عوامل خطر فعالیت فیزیکی و شاخص توسعه‌ی انسانی معنی‌دار شده‌اند. مثبت بودن این دو ضریب نشان می‌دهد که افزایش اضافه وزن و چاقی و شاخص توسعه انسانی باعث افزایش میزان بروز سرطان می‌شود. در جدول زیر برآورد واریانس  $\alpha$ ،  $u$  و  $v$  برای مدل‌ها آمده است.

عوامل خطر مصرف سیگار، اضافه‌وزن و چاقی، فعالیت فیزیکی، مصرف میوه و سبزیجات و شاخص توسعه انسانی به ترتیب دارای میانگین‌های  $۱۲۹/۶۳۶۵$ ،  $۴۳/۴۱۹۳$ ،  $۳۲/۸۹۷۱$ ،  $۲/۵۳۴۲۵$ ،  $۰/۷۳۵۷$  و واریانس‌های  $۹۵۰/۵۶۸۳$ ،  $۳۲/۹۵۶۸$ ،  $۴۷/۱۰۴۸$ ،  $۰/۲۷۸۵$ ،  $۰/۰۰۲۴$  می‌باشند.

مدل لگ-نرمال همبستگی ساختاری بین استان‌ها را در نظر نمی‌گیرد. مدل کامل‌تر مدل BYM می‌باشد که هر دو همبستگی ساختاری و غیر ساختاری را در نظر می‌گیرد.

(CrI یا Credible interval) یا استفاده شد که معادل  $p$ -value در ساختار بی‌زی است.

### یافته‌ها

تعداد کل سرطان سینه ثبت شده در سال‌های ۱۳۸۴ تا ۱۳۸۷ در ایران ۲۵۱۵۲ مورد بوده است. در بین استان‌ها بیشترین مورد مشاهده شده مربوط به استان‌های اصفهان با ۲۲۱۹ مورد و استان خراسان رضوی با ۲۲۱۸ مورد می‌باشد. کمترین میزان بروز مربوط به استان کهگیلویه و بویراحمد با ۶۷ مورد می‌باشد. در جدول زیر تعداد موارد سرطان ثبت شده در ۴ سال، جمعیت در معرض خطر و خطر نسبی به‌دست آمده با استفاده از سه مدل گاما-پوآسن، لگ-نرمال و BYM با تعدیل و بدون تعدیل اثر به تفکیک استان‌ها آمده است. نتایج حاصل از مدل‌های مختلف رگرسیون اکولوژی فضایی بی‌زی به شرح زیر است. این مدل‌ها، میزان‌های بروز و میزان‌های بروز تعدیل شده بر حسب عوامل را به‌دست می‌دهند. جدول ۱ نتایج مربوط به ضرایب متغیرهای مستقل را برای سه مدل مورد استفاده در این مطالعه را نشان می‌دهد. با توجه به جدول ۲ قسمت مربوط به مدل گاما-پوآسن، نقش تمام عوامل خطر معنی‌دار شده‌اند. این مدل همبستگی بین استان‌ها را در نظر نمی‌گیرد. به همین دلیل استفاده از جدول مربوط به مدل گاما-پوآسن می‌تواند گمراه‌کننده باشد.

جدول ۱. جدول اطلاعات توصیفی و برآورد خطر نسبی به تفکیک استان‌ها

با در نظر گرفتن عوامل خطر		بدون در نظر گرفتن عوامل خطر		جمعیت در معرض خطر	تعداد کل موارد سرطان سینه در ۵ سال	نام استان‌ها
BYM	لگ-نرمال	BYM	لگ-نرمال			
۰/۶۱۴۹	۰/۷۳۷۴	۰/۷۱۸۲	۰/۵۹۰۹	۶۱۰۸۱۳	۱۹۷	اردبیل
۱/۰۱۱	۰/۹۵۸۴	۰/۶۰۳۰	۰/۷۴۵۰	۱۴۱۴۶۷۸	۶۵۲	آذربایجان غربی
۰/۸۷۰۱	۰/۹۳۱۸	۰/۸۰۴۴	۰/۸۵۴۵	۱۷۶۴۱۱۵	۹۶۴	آذربایجان شرقی
۱/۴۸۸۱	۰/۹۴۸۱	۰/۷۵۷۱	۱/۱۹۱۴	۴۱۸۲۰۵	۲۹۹	بوشهر
۱/۰۹۴۵	۱/۱۴۸۵	۰/۶۷۶۷	۰/۶۳۹۷	۴۲۶۲۶۶	۱۷۱	چهارمحال و

							بختیاری
۰/۹۹۸۱	۱/۲۰۸۳	۱/۲۲۱۱	۱/۶۳۲۴	۱/۹۵۷۲	۲۲۲۳۸۵۷	۲۲۱۹	اصفهان
۱/۱۵۰۳	۱/۳۱۳۲	۱/۰۶۸۳	۱/۴۰۱۸	۱/۶۶۱۱	۲۱۳۲۰۲۶	۱۷۹۵	فارس
۰/۷۴۰۹	۰/۸۸۵۸	۱/۰۵۶۵	۰/۹۳۱۰	۱/۲۱۲۵	۵۵۹۳۳۱	۳۴۴	قزوین
۱/۰۱۲۲	۱/۲۴۳۷	۱/۱۳۷۱	۱/۴۰۳۶	۱/۶۶۷۳	۱۲۰۶۳۴۲	۹۸۹	گیلان
۱/۰۱۲۲	۱/۰۷۲۶	۱/۰۸۷۱	۱/۲۱۸۶	۱/۵۱۰۵	۸۱۲۹۲۵	۵۹۵	گلستان
۱/۸۹۹۸	۱/۶۱۰۳	۱/۵۱۷۲	۱/۰۴۲۳	۱/۰۹۸۷	۸۴۶۴۳۰	۴۹۸	همدان
۱/۰۵۵۰	۰/۸۸۰۶	۰/۷۲۰۲	۰/۵۸۷۲	۰/۸۱۰۳	۶۷۸۴۷۸	۲۹۹	هرمزگان
۱/۲۳۹۲	۱/۰۱۴۵	۰/۶۹۰۱	۰/۷۶۵۰	۰/۸۴۰۹	۲۶۷۲۲۱	۱۰۵	ایلام
۰/۸۷۸۶	۰/۸۰۲۸	۰/۷۲۲۰	۰/۶۶۱۷	۰/۸۳۶۷	۱۲۹۸۵۴۶	۵۱۹	کرمان
۰/۹۶۳۴	۱/۲۵۳۲	۰/۹۹۹۶	۱/۱۳۹۷	۱/۳۴۴۳	۹۲۲۰۶۰	۶۲۴	کرمانشاه
۰/۰۱۲۰	۰/۹۱۹۴	۰/۸۸۹۷	۰/۶۴۴۸	۰/۹۴۹۱	۳۱۴۲۷۸	۱۲۳	خراسان جنوبی
۱/۷۳۲۹	۱/۷۱۶۵	۰/۸۸۶۴	۱/۳۰۰۴	۱/۵۲۷۲	۲۷۸۳۶۷۶	۲۱۱۸	خراسان رضوی
۰/۷۲۴۹	۰/۶۳۵۳	۰/۸۷۹۴	۰/۶۲۷۶	۰/۷۲۵۳	۴۰۹۷۰۸	۱۲۶	خراسان شمالی
۰/۹۹۹۹	۱/۱۵۳۵	۱/۰۳۱۸	۱/۳۶۳۸	۱/۷۰۹۰	۲۰۹۰۰۴۸	۱۷۵۶	خوزستان
۰/۷۷۴۲	۰/۶۰۴۴	۰/۶۴۴۸	۰/۳۵۶۷	۰/۴۸۰۰	۲۱۳۳۱۶	۶۷	کهگیلویه و بویراحمد
۱/۲۹۹۴	۱/۲۲۷۰	۰/۵۷۶۷	۰/۶۹۹۲	۰/۸۲۴۳	۷۱۰۴۳۸	۳۲۵	کردستان
۱/۷۰۷۳	۱/۴۴۳۱	۰/۶۰۷۰	۰/۷۶۰۹	۰/۹۹۸۷	۸۴۲۸۳۵	۴۲۲	لرستان
۱/۰۱۶۱	۱/۱۴۸۷	۰/۸۴۳۱	۰/۹۰۴۷	۰/۹۹۶۷	۶۶۸۸۹۰	۳۵۵	مرکزی
۱/۰۱۲۲۲	۱/۰۷۲۶	۱/۰۹۷۹	۱/۴۲۶۳	۱/۶۴۳۹	۱۴۵۵۵۶۲	۱۲۱۲	مازندران
۰/۷۸۴۹	۰/۸۰۸۳	۱/۱۵۳۵	۱/۰۹۸۴	۱/۲۶۶۲	۵۰۹۸۳۳	۳۳۳	قم
۰/۹۱۹۴	۰/۹۵۶۰	۱/۰۸۹۱	۱/۱۹۶۴	۱/۵۶۶۵	۲۸۷۳۰۹	۲۰۹	سمنان
۱/۰۲۸۸	۰/۷۴۳۱	۰/۳۳۰۰	۰/۲۷۲۴	۰/۳۵۰۰	۱۱۸۴۵۰۲	۲۱۲	سیستان و بلوچستان
۰/۸۸۱۷	۰/۹۷۷۱	۱/۴۶۲۲	۱/۷۲۴۷	۲/۰۸۸۴	۶۵۲۱۱۴۹	۷۰۳۰	تهران
۰/۶۷۹۷	۰/۹۰۷۲	۱/۳۶۲۷	۱/۳۹۲۵	۱/۷۶۷۷	۴۷۳۶۸۱	۴۲۸	یزد
۰/۸۵۹۳	۰/۸۴۰۶	۰/۶۶۳۸	۰/۵۶۴۴	۰/۶۳۳۴	۴۸۲۹۰۳	۱۶۶	زنجان

جدول ۲. برآورد ضرایب اثرات عوامل خطر برای مدل‌های رگرسیون اکولوژی

BYM		لگ نرمال		پواسن-گاما		عوامل خطر
95% CrI	برآورد ضریب	95% CrI	برآورد ضریب	95% CrI	برآورد ضریب	
(-۷/۷۱۸۰, -۳/۸۵۳۰)	-۵/۵۵۵۰	-۷/۲۰۲۰, -۴/۰۳۲۰)	-۵/۵۶۲۰	-۴/۵۶۲۰, -۳/۹۱۵۰)	-۴/۲۳۵۰	عرض از مبدا
(-۰/۰۰۵۷, ۰/۰۰۲۰)	-۰/۰۰۱۸۶	(-۰/۰۰۵۴, ۰/۰۰۱۸۸)	-۰/۰۰۱۸۴۱	(-۰/۰۰۲۵, -۰/۰۰۰۷)	-۰/۰۰۱۶۵	مصرف سیگار
(-۰/۰۰۱۷۹, ۰/۰۴۲۶)	۰/۰۲۰۷	(۰/۰۰۱۰, ۰/۰۴۲۰)	۰/۰۲۱۵	(-۰/۰۲۲۳, -۰/۰۱۳۷)	-۰/۰۱۸۱	اضافه وزن و چاقی
(-۰/۰۱۱۰, ۰/۰۱۹۶)	۰/۰۰۴۱۵	(-۰/۰۱۱۰, ۰/۰۱۹۰)	۰/۰۰۳۸۸	(-۰/۰۰۸۰۲, -۰/۰۰۱۱)	-۰/۰۰۰۴۵	فعالیت فیزیکی
(-۰/۲۲۵۸, ۰/۱۸۰۵)	-۰/۰۲۴۵	(-۰, ۲۳۰۷, ۰, ۱۸۳۹)	-۰/۰۲۲۰	(-۰/۱۴۷۸, -۰/۰۴۸۶)	-۰/۰۹۷۸	مصرف میوه و سبزیجات
(۳/۷۹۴۰, ۸/۵۲۷۰)	۶/۱۷۲۰	(۳/۹۰۸۰, ۸/۴۱۸۰)	۶/۱۲۲۰	(۶/۹۷۰۰, ۷/۹۰۰۰)	۷/۴۲۶۰	شاخص توسعه انسانی
	۲۶۹/۱۰۰۰		۲۹۸/۲۰۰۰		۸۸۸/۶۰۰۰	DIC برای مدل با عوامل خطر
	۱۷۱/۴۰۰۰		۲۹۸/۱۰۰۰		-	DIC برای مدل بدون عوامل خطر

جدول ۳. برآورد واریانس آلفا، U و V

با در نظر گرفتن عوامل خطر			بدون در نظر گرفتن عوامل خطر			برآورد واریانس
BYM	لگ نرمال	گاما-پواسن	BYM	لگ نرمال		
۰/۶۸۴۴	۰/۶۸۱۴۵	۰/۰۱۱۳	۰/۰۰۵۶۰	۰/۰۰۶۵۲		آلفا
۰/۰۰۰۳۱	-	-	۰/۰۰۴۳۷	-		U
۰/۰۴۲۶۵	۰/۰۵۴۰۰	-	۰/۱۳۳۸۰	۰/۱۸۴۲۴		V

غیرساختاری نشان می‌دهد. طبق این نقشه، استان‌های یزد، قزوین و اردبیل و خراسان شمالی دارای کمترین خطر بروز سرطان سینه و استان‌های خراسان رضوی و خرم آباد و همدان دارای بیشترین خطر ابتلا بروز می‌باشند.

ساختاری نشان می‌دهد. طبق این نقشه، استان‌های اصفهان و تهران دارای بیشترین خطر بروز و استان سیستان و بلوچستان و کهگیلویه و بویر احمد دارای کمترین خطر بروز می‌باشند. شکل (۱)، نقشه‌ی (ه)، خطر نسبی استان‌ها را با تعدیل اثر عوامل خطر و با در نظر گرفتن ناهمگنی‌های ساختاری و

شکل (۱)، نقشه‌ی (د)، خطر نسبی استان‌ها را بدون تعدیل اثر عوامل خطر و با در نظر گرفتن ناهمگنی‌های ساختاری و غیر عوامل خطر مصرف سیگار، اضافه‌وزن و چاقی، فعالیت فیزیکی، مصرف میوه و سبزیجات و شاخص توسعه انسانی به‌ترتیب دارای میانگین‌های ۱۲۹/۶۳۶۵، ۴۳/۴۱۹۳، ۳۲/۸۹۷۱، ۲/۵۳۴۲۵، ۰/۷۳۵۷ و واریانس‌های ۹۵۰/۵۶۸۳، ۳۲/۹۵۶۸، ۴۷/۱۰۴۸، ۰/۲۷۸۵، ۰/۰۰۲۴ می‌باشند.

مدل لگ نرمال همبستگی ساختاری بین استان‌ها را در نظر نمی‌گیرد. مدل کامل‌تر مدل BYM می‌باشد که هر دو همبستگی ساختاری و غیر ساختاری را در نظر می‌گیرد.

با در نظر گرفتن ناهمگنی ساختاری و غیر ساختاری تنها متغیر توضیحی شاخص توسعه انسانی معنی‌دار می‌باشد و نشان می‌دهد که خطر ابتلا به سرطان سینه در نواحی دارای شاخص توسعه انسانی بالاتر، بیشتر است.

شکل ۱، نقشه‌های مربوط به خطر بروز سرطان سینه را برای مدل‌های BYM، لگ‌نرمال، گاما-پوآسن با تعدیل و بدون تعدیل عوامل خطر نشان می‌دهد.

### بحث

در مدل BYM بدون تعدیل اثر عوامل خطر، استان‌های (مرکزی) اصفهان و یزد و تهران بیشترین خطر بروز سرطان سینه و بعد از آن‌ها استان‌های شمالی و فارس و خوزستان خراسان شمالی بالاترین خطر بروز را دارند و استان‌های شمال شرقی و استان‌های جنوب غربی کمترین میزان خطر ابتلا را دارند. درمیان استان‌ها، استان‌های سیستان و بلوچستان و چهارمحال و بختیاری کمترین میزان خطر ابتلا را دارند. این نتایج با نتایج مطالعه جعفری کشکی و همکاران (۱۷) مطابقت دارد. با این تفاوت که در آن مطالعه اثر عوامل خطر تعدیل نشده است ولی در مطالعه حاضر اثر عوامل خطر نیز تعدیل شده است.

با تعدیل عوامل خطر، استان‌های خراسان رضوی، خرم‌آباد و همدان بیشترین خطر بروز سرطان سینه و استان‌های اردبیل و کهگیلویه و بویراحمد، کمترین خطر بروز را دارند و در کل استان‌های جنوب غربی دارای خطر بروز کمتری می‌باشند.

جدول ۲، مقایسه‌ی نیکویی برازش مدل‌های گاما پوآسن، لگ‌نرمال و BYM را با حضور و بدون عوامل خطر با استفاده از شاخص DIC نشان می‌دهد. مدل BYM بدون حضور عوامل خطر، بهترین برازش را دارد، زیرا ناهمگنی‌های ساختاری و غیر ساختاری را در بر دارد. در این مدل اثر عوامل خطر با در نظر گرفتن ناهمگنی‌های ساختاری و غیر ساختاری، تعدیل می‌شود و نیازی به در نظر گرفتن عوامل خطر نمی‌باشد. مدل گاما-پوآسن بدون حضور عوامل خطر، بدترین برازش را دارد. با در نظر گرفتن عوامل خطر، دو مدل BYM و لگ‌نرمال برازش تقریباً یکسانی دارند. در مدل لگ‌نرمال، در نظر گرفتن عوامل خطر تغییری در نیکویی برازش مدل ایجاد نکرد. مدل گاما-پوآسن برازش ضعیفی دارد زیرا این مدل همبستگی‌های فضایی بین استان‌ها را در نظر نمی‌گیرد.

شکل (۱)، نقشه‌ی (الف)، خطر نسبی استان‌ها را بدون تعدیل عوامل خطر و بدون در نظر گرفتن ناهمگنی‌های ساختاری و غیر ساختاری نشان می‌دهد. با توجه به این نقشه، استان‌های، یزد و تهران دارای بیشترین خطر و استان سیستان و بلوچستان دارای کمترین خطر بروز سرطان سینه می‌باشند و در کل استان‌های مرکزی دارای خطر بیشتری می‌باشند.

شکل (۱)، نقشه‌ی (ب)، خطر نسبی استان‌ها را با در نظر گرفتن ناهمگنی‌های غیر ساختاری و بدون تعدیل اثر عوامل خطر نشان می‌دهد. با توجه به این نقشه، استان‌های شمال غرب و جنوب شرق خطر بروز سرطان سینه‌ی کمتری دارند و استان‌های سیستان و بلوچستان و کهگیلویه و بویر احمد کمترین بروز سرطان سینه را دارند.

شکل (۱)، نقشه‌ی (ج)، خطر نسبی استان‌ها را با در نظر گرفتن ناهمگنی غیر ساختاری و با تعدیل عوامل خطر نشان می‌دهد. با توجه به این نقشه، استان خراسان رضوی و همدان دارای بیشترین خطر بروز سرطان سینه و استان کهگیلویه و بویر احمد، سیستان و بلوچستان و اردبیل دارای کمترین خطر بروز سرطان سینه می‌باشند.



## نتیجه‌گیری

با توجه به نتایج به دست آمده بروز سرطان سینه در استان‌های با شاخص توسعه انسانی بالاتر، بیشتر است. این موضوع ممکن است به دلیل آلودگی بیشتر هوا و تماس بیشتر با مواد سرطان‌زا در این شهرها باشد، یا ممکن است به دلیل تشخیص بیشتر سرطان در این استان‌ها باشد.

با مقایسه‌ی دو نقشه‌ی (ب) و (ج)، می‌توان مشاهده کرد که با وجود تفاوت در خطر نسبی‌های به دست آمده، رتبه‌ی خطر نسبی استان‌ها در اکثر موارد برابر می‌باشد. این موضوع با مطالعه‌ی Clayton D, Kaldor J (۲۲) مطابق است.

با توجه به جدول (۲) اثر مستقیم شاخص توسعه انسانی بر بروز سرطان سینه تأیید می‌شود. این نتیجه با نتیجه‌ی مطالعه‌ی Yost و همکاران (۱۴) مطابقت دارد. این موضوع می‌تواند به دلیل افزایش سبک زندگی غربی و افزایش آلودگی با افزایش شاخص توسعه انسانی باشد. همچنین افزایش شاخص توسعه انسانی می‌تواند باعث افزایش شناسایی موارد سرطانی شود.

ارتباط معنی‌داری بین مصرف میوه و سبزیجات و بروز سرطان سینه یافت نشد در حالی که در مطالعه‌ی عباس‌تبار و همکاران (۱۶) ارتباط معنی‌داری بین سرطان سینه و مصرف میوه و سبزیجات مشاهده شد. این تفاوت می‌تواند به این دلیل باشد که در مطالعه‌ی حاضر ارتباط در سطح بوم‌شناختی (استان‌ها) بررسی شد ولی در مطالعه‌ی عباس‌تبار و همکاران، این ارتباط در سطح اشخاص بررسی شد.

در سال ۸۲ بختیاری و همکاران مطالعه‌ای بر ۷۲ بیمار مبتلا به سرطان سینه و ۸۵ زن سالم در شهر بابل‌سر انجام دادند (۹) و ارتباط معنی‌داری بین سطح اقتصادی-اجتماعی و سرطان سینه یافتند. در مطالعه حاضر نیز وجود این ارتباط تأیید شد. در این مطالعه رابطه بین فعالیت فیزیکی کم، مصرف سیگار و خطر بروز سرطان سینه رابطه معنی‌داری مشاهده شد ولی به دلیل متفاوت بودن جامعه هدف و نوع مطالعه، این ارتباط در مطالعه‌ی حاضر یافت نشد.

شریف زاده و همکاران (۸) بر روی ۸۵ بیمار مبتلا به سرطان سینه و ۸۵ شاهد در شهر بیرجند در سال‌های ۸۷-۸۵ مطالعه کردند و ارتباط مستقیم و معنی‌داری بین بروز سرطان سینه و مصرف سیگار مشاهده کردند. این امر می‌تواند به دلیل متفاوت بودن جامعه هدف و متفاوت بودن نوع مطالعه باشد.

در مطالعه‌ی هلاکویی (۱۳) و همکاران که بر روی عوامل خطر موثر بر بروز سرطان سینه در سال ۱۳۸۳ در استان مازندران انجام گرفت اثر BMI بالا بر بروز سرطان سینه تأیید. در صورتی که در مطالعه‌ی حاضر این عامل معنی‌دار نشد. زیرا مطالعه‌ی هلاکویی و همکاران یک مطالعه مورد-شاهدی در استان مازندران در سال ۱۳۸۳ بوده است ولی مطالعه‌ی حاضر یک مطالعه اکولوژی بر روی کل بیماران ایران در سال‌های ۸۳ تا ۸۷ می‌باشد.

از مشکلات این پژوهش می‌توان به در دسترس نبودن عوامل خطر در سطح اشخاص اشاره کرد. برای رفع این مشکل از رگرسیون اکولوژی استفاده شد. و برای جلوگیری از اریبی اکولوژی نتایج این تحقیق نباید در سطح اشخاص تفسیر شود. همچنین مستقل نبودن استان‌ها از دیگر مشکلات این پژوهش بود که برای رفع این مشکل، از روش‌های بیزی استفاده شد.

از اولین ثبت آمار سرطان (به‌عنوان علت مرگ) در دنیا نزدیک به ۴۰۰ سال می‌گذرد و این در حالی است که با توجه به پیشرفت‌های بسیار زیاد در زمینه‌ی مکانیزه شدن، این برنامه همچنان درگیر مشکلات ناشی از عدم ثبت صحیح و پوشش ناکامل ثبت مبتنی بر جمعیت است. سابقه‌ی ثبت منسجم و هماهنگ سرطان در ایران به حدود سال ۱۳۶۸ برمی‌گردد. در سال ۱۳۸۷ ادغام آمار پاتولوژی با اولین سری نتایج آمار مبتنی بر جمعیت از مراکز غیر پاتولوژی نظیر مراکز تشخیص بالینی و درمانی سرطان، رادیولوژی و مراکز ثبت مرگ، سطح پوشش ثبت سرطان را از ۸۶/۷٪ در سال ۱۳۸۶ به ۱۰۵٪ در سال ۱۳۸۷ ارتقا داد (۳).

با توجه روند رو به رشد زندگی شهری و رابطه‌ی مستقیم شاخص توسعه انسانی و نرخ بروز سرطان سینه، برای جلوگیری از افزایش بروز سرطان سینه لازم است تدابیری برای اطلاع رسانی و اصلاح سبک زندگی، اتخاذ گردد.

در این مطالعه به دلیل در دسترس نبودن اطلاعات بسیاری از عوامل خطر در سطح اکولوژی، عوامل موثر زیادی مانند آلودگی هوا، سابقه‌ی خانوادگی، وضعیت شیردهی، ... مورد مطالعه قرار نگرفت. بهتر است در مطالعات آینده این عوامل در نظر گرفته شوند.

## References

1. Fehm T, Ruckhäberle E, Lammert E, Zeeb M. Introduction to Breast Cancer. Principles of Metabolism in Health and Diseases. 2013:379.
2. Bray F, Jemal A, Grey N, Ferlay J, Forman D. Global cancer transitions according to the Human Development Index (2008–2030): a population-based study. The lancet oncology. 2012;13(8):790-801.
3. Akbari M, Abachizadeh K, Khayamzadeh M, Tabatabaee M, Esnaashari F, Motlagh A. Iran cancer report. Cancer Research Center Shahid Beheshti University of Medical Sciences Tehran, Qom: Darolfekr. 2008.
4. Kerenyi NA, Pandula E, Feuer G. Why the incidence of cancer is increasing: the role of 'light pollution'. Medical hypotheses. 1990; 33(2):75-8.
5. Rahmati Yamchi M, Zarghami N, Rahbani Noubar M, Najafipour R, Mobsser M. Correlation between telomerase gene expression and different stages of breast cancer and obesity. 2012;2:35-43.
6. Tehrani M, Mohseni S. Case finding and clinical study of obese patients. Iranian Journal of Diabetes and Metabolism. 2013;12(5):393-8.
7. Saki A, Hajianzadeh E, Tehranian N. Evaluation of breast cancer risk factors using analysis of tree models. Journal of Ofogh Danesh. 2011;17(1):60-7.
8. Sharif Zadeh G, Hosseni M, Kermani T, Atai M, Akhbari H. Factors associated with breast cancer - a case-control study. Journal of Birjand University of Medical Sciences. 2011;18(3):191-9.
9. Bakhtiari A, Hajian K, Faramarzi M. Examine the relationship between demographic, social and lifestyle in women with breast cancer. Journal of Babol University of Medical Sciences. 2002;9(2):55-62.
10. Vahedian Shahroodi M, Poorhaji F, Ismaili H, Poorhaji F. The role of physical activity in the prevention of breast cancer. Proceedings of the Eighth Student Conference of Medical Universities in East. 2013;1(1).
11. Kahn HS, Tatham LM, Patel AV, Thun MJ, Heath Jr CW. Increased cancer mortality following a history of nonmelanoma skin cancer. Jama. 1998;280(10):910-2.
12. Arkani M, Safa'i A, Karimi K, Arbabi E, Rostami F, et al. Insulin gene polymorphism and colorectal cancer. 2011;13(2):172-176.
13. Holakouee Naeeni K, Ardalan A, Mahmoudi M, Motovallian A, Yahia Poor Y. study on Risk factors for breast cancer in Mazandaran in 1383. Public Health and Institute of Health Research. 2006;4(1):27-36.
14. Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. Cancer Causes & Control. 2001;12(8):703-11.
15. Snowdon DA, PHILLIPS RL, Choi W. Diet, obesity, and risk of fatal prostate cancer. American journal of epidemiology. 1984;120(2):244-50.
16. Abbastabr H, Soleimani Dodran M, Hamidi Fard P, Jalilian F, Alvijeh M, Nasir Zadeh M. The Relationship between Breast Cancer and the Most Common Non-contagious Disease Risk Factors: an Ecologic Study. 2012;8(3):369-374.
17. Jafari-Koshki T, Schmid VJ, Mahaki B. Trends of breast cancer incidence in Iran During 2004-2008: A Bayesian space-time model. Asian Pac J Cancer Prev. 2014;15:1557-61.
18. Lawson AB, Browne WJ, Rodeiro CLV. Disease mapping with WinBUGS and MLwiN: John Wiley & Sons; 2003.

19. Lawson A, Biggeri A, Böhning D, Lesaffre E, Viel J-F, Bertollini R. Disease mapping and risk assessment for public health: John Wiley & Sons; 1999.
20. Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics*. 1991;43(1):1-20.
21. Berg A, Meyer R, Yu J. Deviance information criterion for comparing stochastic volatility models. *Journal of Business & Economic Statistics*. 2004;22(1):107-20.
22. Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*. 1987;43(3):671-81.

Archive of SID

## A Comparative Study on Ecological Regression Models in Breast Cancer Risk in Iran, 2005- 2008

Ahmad Haddad Khoshkar<sup>1</sup>, Tohid Jafari-Koshki<sup>1,2</sup>, Behzad Mahaki<sup>3</sup>

### Original Article

#### Abstract

**Background:** Breast cancer is the leading cancer among women in Iran with standardized incidence rate of 28.25. With regard to the aging and increasing cancer trends in Iranian population, the importance and geographical variations of cancer, it is necessary to identify those regions with high breast cancer incidence and get information on geographical its patterns and prominent risk factors using advanced statistical methods.

**Methods:** The used population in the present ecological study is all those breast cancer patients in Iran 2005-2008. We assessed the impact of cigarette, fruit and vegetable consumption, physical activity, overweight, and socioeconomic status on breast cancer using gamma-poisson, log-normal and BYM models.

**Findings:** Total number of registered breast cancer patients was 25152. From models used to analyze the data, unadjusted BYM had the best performance. In unadjusted models, Isfahan, Yazd, and Tehran had the highest and Sistan-and Baluchestan and Chaharmahal-and-Bakhtiari incidence had the lowest incidence rate. After adjustment, the incidence was estimated to be the highest in Khorasan-Razavi, Khoramabad, and Hamadan and the lowest in Ardebil and Kohgiluyeh-and -Boyerahmad. Direct association between human development index (HDI) and breast cancer incidence was found to be statistically significant.

**Conclusion:** BYM model had the best fit and may cast so in future applications. Due to significant relationship between HDI breast cancer, it is recommended to take it into account as a risk factor in future researches in this field. The incidence is generally high in central provinces.

**Key Words:** Disease Mapping, Breast Cancer, Human Development Index, BYM Model

**Citation:** Haddad Khoshkar A, Jafari-Koshki T, Mahaki B. A Comparative Study on Ecological Regression Models in Breast Cancer Risk in Iran, 2005- 2008. J Health Syst Res 2015; 11(3):613-626

Received date: 27.12.2014

Accept date: 09.06.2015

1. Department of Biostatistics, School of Public Health, Isfahan University of Medical Sciences, Isfahan, Iran
2. Department of Biostatistics, School of Public Health, Kermanshah University of Medical Sciences Kermanshah., Iran
3. Assistant Professor, School of Health, Isfahan University of Medical Sciences, Iran (Corresponding Author) Email: Behzad.mahaki@gmail.com

**OpenBUGS code for BYM**

```

model
{
for (i in 1:m)
{
# Poisson likelihood for observed counts
y[i]~dpois(mu[i])
log(mu[i])<-log(e[i])+alpha+v[i]+u[i]+b1*AFF1[i]+b2*AFF2[i]+b3*AFF3[i]+b4*AFF4[i]+b5*AFF5[i]
# Relative Risk
RR_adj[i]<-exp(v[i]+u[i])

# Prior distribution for the uncorrelated heterogeneity
v[i]~dnorm(0,tau.v)

}

}

eps<-1.0E-6

# CAR prior distribution for spatial correlated heterogeneity
u[1:m]~car.normal(adj[,weights[,num[,tau.u)

# Weights
for(k in 1:sumNumNeig)
{
weights[k]<-1
}

# Improper prior distribution for the mean relative risk in the study region
alpha~dflat()
mean<-exp(alpha)

# Prior on regression coefficients
b1~dnorm(0.0,1.0E-5)
b2~dnorm(0.0,1.0E-5)
b3~dnorm(0.0,1.0E-5)
b4~dnorm(0.0,1.0E-5)
b5~dnorm(0.0,1.0E-5)

# Hyperprior distributions on inverse variance parameter of random effects
tau.u~dgamma(0.5,0.0005)
tau.v~dgamma(0.5,0.0005)

}

```

**OpenBUGS code for log-normal model**

```

{
for (i in 1:m)
{
# Poisson likelihood for observed counts
y[i]~dpois(mu[i])
log(mu[i])<-log(e[i])+alpha+v[i]+b1*AFF1[i]+b2*AFF2[i]+b3*AFF3[i]+b4*AFF4[i]+b5*AFF5[i]

# RR_adj
RR_adj[i]<-exp(v[i])

# Prior distribution for the uncorrelated heterogeneity
v[i]~dnorm(0,tau.v)

}

}

for(k in 1:sumNumNeig)
{
weights[k]<-1
}
alpha~dflat()
mean<-exp(alpha)
# Prior on regression coefficients

```

```

b1~dnorm(0.0,1.0E-5)
b2~dnorm(0.0,1.0E-5)
b3~dnorm(0.0,1.0E-5)
b4~dnorm(0.0,1.0E-5)
b5~dnorm(0.0,1.0E-5)

# Hyperprior distributions on inverse variance parameter of random effects
tau.v~dgamma(0.5,0.0005)

```

**OpenBUGS code for gamma-poisson model**

```

{
for (i in 1:m)
{
# Poisson likelihood for observed counts
y[i]~dpois(mu[i])
log(mu[i])<-log(e[i])+alpha+b1*AFF1[i]+b2*AFF2[i]+b3*AFF3[i]+b4*AFF4[i]+b5*AFF5[i]
# Relative Risk decomposition
RR_exp[i]<-exp(b1*AFF1[i]+b2*AFF2[i]+b3*AFF3[i]+b4*AFF4[i]+b5*AFF5[i])
}

# Weights
for(k in 1:sumNumNeig)
{
weights[k]<-1
}

# Improper prior distribution for the mean relative risk in the study region
alpha~dflat()
mean<-exp(alpha)

# Prior on regression coefficients
b1~dnorm(0.0,1.0E-5)
b2~dnorm(0.0,1.0E-5)
b3~dnorm(0.0,1.0E-5)
b4~dnorm(0.0,1.0E-5)
b5~dnorm(0.0,1.0E-5)
}

```