

# تحلیل روند میزان بروز سل با استفاده از مدل رگرسیون قطعه‌ای خطی در ایران طی سال‌های ۸۷-۱۳۴۳

شهرام ارسنگ<sup>۱</sup>، انوشیروان کاظم‌نژاد<sup>۲</sup>، فیروز امانی<sup>۳</sup>

<sup>۱</sup>کارشناس ارشد آمار زیستی دانشگاه تربیت مدرس تهران، ایران

<sup>۲</sup>استاد گروه آمار زیستی دانشگاه تربیت مدرس تهران، ایران

<sup>۳</sup>استادیار گروه آمار زیستی دانشکده پزشکی دانشگاه علوم پزشکی اردبیل، ایران

نویسنده مسئول: انوشیروان کاظم‌نژاد، نشانی: تهران، پل نصر، دانشگاه تربیت مدرس تهران، تلفن: ۴۴۳۳۰۰۴، پست الکترونیک: kazem\_an@modares.ac.ir

تاریخ دریافت: ۸۹/۶/۲۸؛ پذیرش: ۹۰/۴/۴

**مقدمه و اهداف:** بررسی روند تغییرات میزان‌های مشاهده شده، اطلاعات ارزشمندی را جهت نیازسنجی، طرح، بازنگری برنامه‌ها و شاخص‌های توسعه هر کشور فراهم می‌سازد. هدف از مطالعه حاضر، به کار بردن مدل رگرسیون قطعه‌ای در تعیین روند تغییرات میزان بروز سل در ایران بوده است.

**روش کار:** در این مطالعه روند تغییرات الگوی میزان بروز سل در ایران طی ۴۴ سال گذشته (۸۷-۱۳۴۳) با استفاده از مدل رگرسیون قطعه‌ای خطی مورد بررسی و تحلیل قرار گرفت. برای برآورد پارامترهای رگرسیونی از روش‌های کمترین مربعات و آزمون جایگشت و از معیار اطلاع بیز (BIC) نیز برای تصمیم‌گیری در انتخاب بهترین مدل از بین دو مدل رگرسیون پواسون و قطعه‌ای، استفاده گردید. تحلیل داده‌ها در نرم‌افزارهای آماری SAS و Joinpoint3.4 صورت گرفت.

**نتایج:** با توجه به مقدار  $p$  به دست آمده از آزمون جایگشت، فرض وجود دو نقطه تغییر طی سال‌های ۱۳۷۲ و ۱۳۵۶ پذیرفته شد ( $p=0/108$ ). میزان بروز بیماری سل طی ۱۱ سال اول بررسی سیر نزولی با  $1/10$ ٪ کاهش سالانه، برای قطعه دوم طی سال‌های ۷۲-۱۳۵۶ سیر صعودی با  $4/3$  درصد افزایش در میزان بروز سالانه و برای قطعه سوم سیر نزولی با  $4/5$  درصد کاهش در میزان بروز سالانه بوده است. همچنین برای حداقل ۱۰ سال آخر به طور میانگین سالانه  $4/5$  درصد از بروز بیماری سل کاسته شده است.

**نتیجه‌گیری:** با توجه به نتایج بدست آمده از مدل رگرسیون قطعه‌ای می‌توان گفت، الگوی میزان بروز بیماری سل از سال ۱۳۷۱ به بعد روند نزولی داشته، که این کاهش از سرعت بالاتری در مقایسه با میانگین جهانی برآورد شده دارد، که نشان دهنده ارتقاء برنامه‌های کنترل بیماری سل و بهبود وضعیت بهداشتی در کشور است.

**واژگان کلیدی:** رگرسیون قطعه‌ای، نقطه‌ی شکست، آزمون جایگشت، درصد تغییرات سالانه (APC)

## مقدمه

یکی از مولفه‌های مهم در برنامه‌ریزی‌های مربوط به سلامت هر جامعه، تعیین روند بروز بیماری‌ها در آن جامعه است. دانستن الگوی تغییرات میزان بروز بیماری‌ها در هر کشوری می‌تواند از اهمیت بالایی برای برنامه‌ریزی‌های کشوری برخوردار باشد. سازمان‌های بهداشت عمومی بر این باورند که بررسی یا نظارت بر روند میزان بروز بیماری‌ها، مرگ و میر و عوامل خطر اجتماعی، رفتاری و سلامتی ممکن است در رخدادهای نامطلوب سلامتی نقش داشته باشند. بررسی روند تغییرات میزان‌های بروز یا شیوع مشاهده شده، اطلاعات ارزشمندی را برای نیازسنجی، طرح، بازنگری برنامه‌ها و شاخص‌های توسعه هر کشور فراهم می‌نماید. بررسی داده‌ها در طی زمان نیز می‌تواند در پیش‌بینی میزان و

فراوانی رخدادهای آینده نقش داشته باشد.

سل یک بیماری باکتریال مزمن است که در اثر مجموعه‌ای از مایکوباکتریوم‌های سلی ایجاد می‌شود که در  $85$ ٪ موارد به شکل ریوی و  $15$ ٪ به شکل غیر ریوی تظاهر می‌نماید. بیماری سل امروزه جزو یکی از بزرگترین مسائل بهداشتی جهان به شمار می‌رود، حدس زده می‌شود که از هر سه نفر جمعیت جهان یک نفر به باسیل سل آلوده بوده و در هر ثانیه یک نفر به تعداد آنان افزوده می‌شود. براساس آخرین گزارش سازمان جهانی بهداشت در سال ۲۰۰۷ تعداد کل مبتلایان به بیماری سل  $13/7$  میلیون برآورد شده و حدود  $0/5$  میلیون نفر از این افراد، به باسیل سل مقاوم به چند دارو (MDR-TB) آلوده هستند. هر ساله حدود  $9/3$

پارامترهای مدل رگرسیون قطعه‌ای با فرض نقاط شکست معلوم استفاده نموده است. رابیسون<sup>۸</sup> (۱۹۶۴) برای برآورد محل اتصال دو تابع رگرسیونی با فرض وجود N1 مشاهده درقطعه اول و N2 مشاهده درقطعه دوم خط رگرسیونی برفرض معلوم یا نامعلوم بودن نقطه شکست به ترتیب از روش برآورد ماکزیمم درست‌نمایی و درست‌نمایی شرطی استفاده نموده است (۵،۶). برمن<sup>۹</sup> (۱۹۹۶) از روش حداقل مربعات غیرخطی برای تعیین نقاط انتقال<sup>۱۰</sup> سطوح قند خون، تستسترون و ارتباط کورتیزول با سن استفاده نموده است (۷). لرمن<sup>۱۱</sup> (۱۹۸۰) از روش‌های جستجوی شبکه‌ای برای برازش کمترین توان‌های دوم مدل‌های قطعه‌ای خطی و هادسون (۱۹۶۶) از الگوریتم برازش پیوسته برای برآورد پارامترهای رگرسیون قطعه‌ای استفاده نمودند (۸،۹).

در مطالعه انجام شده توسط لارا و باتیستا در سال ۲۰۰۵ در زمینه پایداری باسیل ترنژین سیس در برگه ذرت با غلظت آن‌ها از مدل رگرسیون تکه‌ای برای برآورد پارامترهای مدل استفاده کردند و مقادیر پارامترهای رگرسیونی را برای هر تکه برآورد کردند (۱۰). مطالعه‌ای توسط بودما در سال ۲۰۱۰ روی عفونت ناشی از تنفس مصنوعی در بیمارستان انجام شده و از تحلیل رگرسیون تکه‌ای برای ارزیابی تاثیر برنامه مداخله طراحی شده روی کنترل عفونت استفاده کردند (۱۱). مطالعه‌ای توسط لائو و گانگه به سال ۲۰۰۳ در راستای بررسی روند کاهش تعداد لنفوسیت‌ها و غلظت هموگلوبین به عنوان پیشینی برای ایدز در مردان مبتلا به عفونت HIV انجام شده که از رگرسیون تکه‌ای برای تحلیل ارزیابی روند فوق استفاده کردند (۱۲). رگرسیون قطعه‌ای کاربردهای زیادی در علوم پزشکی، زیست‌شناسی، علوم جانوری و گیاهی و ژنتیک داشته است و مساله برآورد تعداد نقاط تغییر در مدل رگرسیون تکه‌ای توسط کیم در سال ۲۰۰۰ بیان شده است. کیم در سال ۲۰۰۰ یک سری از آزمون‌های جایگشتی را برای تعیین تعداد نقاط تغییر نامعلوم در رگرسیون تکه‌ای به کار برد (۱۳). رگرسیون قطعه‌ای خطی یک ابزار مفید برای توصیف روند تغییرات داده‌ها به ویژه داده‌های مربوط به میزان بروز بیماری یا مرگ و میر است. از آنجا که روند تغییرات الگوی بیماری در کشور بر اساس بروز سالیانه تغییر می‌کند لذا برای بررسی این تغییرات روند، استفاده از مدل‌های ریاضی اغلب اوقات می‌تواند مشکل‌گشا بوده و دید

میلیون نفر جدید به سل فعال مبتلا شده و حدود ۱/۷ میلیون نفر در اثر این بیماری می‌میرند. بیماری سل که بزرگترین علت ناشی از بیماری‌های عفونی تک‌عاملی در جهان است (حتی بیشتر از ایدز، مالاریا و سرخک)، دارای مرتبه دهم در بار جهانی بیماری‌هاست و پیش‌بینی می‌شود تا سال ۲۰۲۰ همچنان جایگاه کنونی خود را حفظ و گاهی تا رتبه هفتم هم بالا رود. در حال حاضر میزان بروز سل ریوی اسمیر مثبت اصلی‌ترین شاخص اندازه‌گیری وضعیت سل در جامعه بوده و عبارت است از تعداد افراد مبتلا به سل ریوی اسمیر مثبت در طول یک سال به ازای هر یکصد هزار نفر است (۱،۲،۳).

رگرسیون قطعه‌ای خطی که به نام‌های رگرسیون تکه‌ای<sup>۱</sup>، رگرسیون دو یا چندحالتی پیوسته<sup>۲</sup>، کمترین توان‌های دوم دو مرحله‌ای و رگرسیون خط شکسته<sup>۳</sup> معروف است، یکی از روش‌های تحلیل رگرسیونی است که در آن متغیر مستقل در فواصل تقسیم و برای هر فاصله یک خط رگرسیونی جداگانه برازش داده شده و مرز بین قطعات نقاط شکست نامیده می‌شود. از رگرسیون قطعه‌ای خطی برای بیان نقاط شکست پیوسته در مرگ و میر بیماری‌ها و میزان‌های بروز استفاده می‌شود (۴). به نقاط شکست، نقاط اتصال، نقطه تغییر<sup>۴</sup> و گره<sup>۵</sup> (در اسپلاین رگرسیون) نیز گفته می‌شود. هدف اصلی این مقاله، استنباط در مورد وجود یا عدم وجود نقطه شکست، برآورد مکان و تعداد نقاط شکست، برآورد پارامترهای قطعه‌های رگرسیونی و مقایسه پارامترهای به دست آمده از توابع رگرسیونی است.

مدل‌های نقطه شکست در بسیاری از زمینه‌های تحقیقات آماری از قبیل مدل‌های خطی تعمیم یافته، مدل‌های تابع خطر، سری‌های زمانی، روش‌های ناپارامتری و مطالعات طولی مورد استفاده قرار می‌گیرند. کیم و فای<sup>۶</sup> در سال ۲۰۰۴ برای مقایسه میزان بروز سرطان در دو گروه جنسی و دو منطقه جغرافیایی مختلف، از مدل رگرسیونی قطعه‌ای خطی و روش کمترین توان‌های دوم برای برآورد پارامترهای رگرسیونی استفاده نمودند (۴). اسپرنت<sup>۷</sup> (۱۹۶۱) از روش کمترین توان‌های دوم برای برآورد

<sup>۱</sup> Piecewise Regression

<sup>۲</sup> Continuse two-or multiphase

<sup>۳</sup> Break line

<sup>۴</sup> Change point

<sup>۵</sup> Knot

<sup>۶</sup> Kim and Fay

<sup>۷</sup> Sprent

<sup>۸</sup> Robison

<sup>۹</sup> Berman

<sup>۱۰</sup> Transition point

<sup>۱۱</sup> Lerman

وزارت بهداشت و درمان استفاده نمودیم که بر اساس برنامه کشوری هر ساله تمام موارد بیماری سل توسط مراکز بهداشتی تمام شهرستان‌های استان‌ها جمع‌آوری و تحلیل می‌شوند. روش نمونه‌گیری به صورت استفاده از داده‌های کل مسلولین ثبت شده به تعداد ۶۱۶۵۱۶ نفر به تفکیک سال و میزان بروز سالانه طی سال‌های ۸۷-۱۳۴۳ بود.

### ساختار مدل رگرسیون قطعه‌ای

مدل رگرسیون قطعه‌ای روشی برای قطعه‌بندی مدل رگرسیونی غیر خطی به تکه‌های خطی جداگانه است که این قطعه‌ها توسط نقاط شکست از هم جدا می‌شوند. در این حالت برای هر قطعه یک تابع  $F_i(X)$  رگرسیون خطی با پارامترهای متفاوت خواهیم داشت. منحنی رگرسیون قطعه‌ای برای  $r$  قطعه به صورت مدل کلی زیر بیان می‌شود:

$$f(x) = E[y|x] = \begin{cases} f_1(x; \beta_1) & x \leq \tau_1 \\ f_2(x; \beta_2) & \tau_1 < x < \tau_2 \\ \vdots & \\ f_r(x; \beta_r) & \tau_{r-1} < x \end{cases}$$

که  $\tau_i$  نشان دهنده نقاط شکست و  $f_i(x; \beta_i)$  توابع رگرسیونی برای هر قطعه می‌باشد. مواردی که باید در رگرسیون قطعه‌ای مورد توجه قرار گیرد، شامل تعداد نقاط شکست، مکان نقاط شکست و معلوم یا نامعلوم بودن نقاط شکست می‌باشند.

مدل رگرسیونی در این حالت به صورت زیر بدست می‌آید:

$$E(y|x) = \beta_0 + \beta_1 x + \delta_1 (x - \tau_1)^+ + \dots + \delta_k (x - \tau_k)^+$$

که  $k$  تعداد نقاط تغییر نامعلوم،  $\tau_i$  برای  $i=(1, \dots, k)$  مکان‌های

نقاط تغییر نامعلوم و  $\delta_i$  ضرایب رگرسیونی قطعات هستند.

همچنین از مقادیر درصد تغییرات سالانه (APC) و متوسط درصد تغییرات سالانه (AAPC) برای مقایسه روند کاهش یا افزایش میزان بروز بیماری سل استفاده کردیم. در این بررسی متغیر پاسخ برای تحلیل داده‌های میزان بروز، لگاریتم طبیعی میزان بروز سالانه بیماری سل و متغیر مستقل سال‌های مورد مطالعه (۸۷-۱۳۴۳) می‌باشد. از روی رسم نمودار گرافیکی برای میزان بروز و تحلیل اولیه داده‌ها معلوم شد که حداکثر سه نقطه تغییر ( $k_{\max}=3$ ) برای داده‌های میزان بروز بیماری سل داریم که نتایج را با در نظر گرفتن سه نقطه تغییر تحلیل کردیم. در مطالعه‌ی حاضر روند تغییرات میزان بروز بیماری سل طی ۴۴

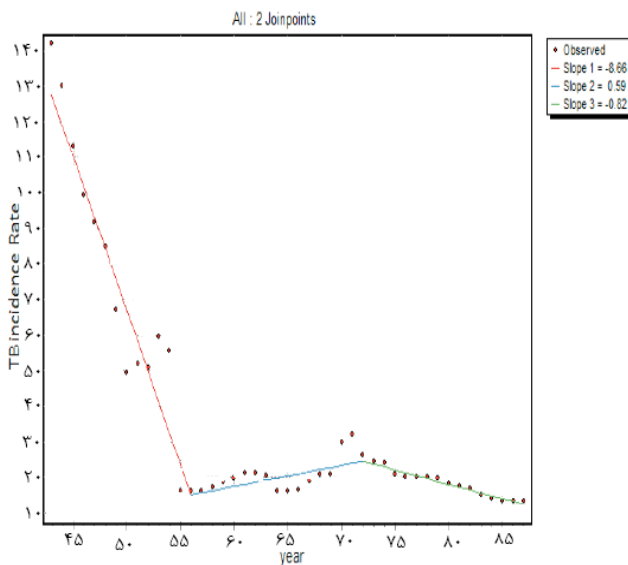
جدیدی را در زمینه بررسی مسایل پزشکی در اختیار ما قرار دهد. در این مطالعه از یک مدل خطی رگرسیونی تحت نام رگرسیون قطعه‌ای استفاده نمودیم تا بتوانیم مکان‌هایی را که احتمال تغییر در الگوی بیماری سل در کشور وجود دارد را شناسایی و پارامترهای مورد نیاز برای تفسیر مدل‌های رگرسیونی را برآورد نماییم.

### روش کار

در راستای اجرای برنامه کشوری DOTS برای شناسایی و گزارش دهی موارد سل، از افراد مراجعه کننده به آزمایشگاه‌های مراکز بهداشتی که به مدت دو هفته و بیشتر سرفه داشتند در ۳ مرحله نمونه‌گیری اسمیر خلط مثبت به عمل می‌آید و اگر در ۲ بار از این نمونه‌گیری‌ها باسیل سل مشاهده شود، شخص مسلول شناخته شده و به مراکز بهداشتی مربوطه جهت درمان و کنترل معرفی می‌گردد. آمار بیماران مسلول شناسایی شده به صورت فصلی (سه ماه یکبار) جمع‌آوری و در اختیار معاونت بهداشتی و وزارت بهداشت جهت جمع‌آوری سالانه و تصمیم‌گیری‌های سیستم بهداشت و درمان قرار می‌گیرند. لازم به ذکر است برنامه ملی کشوری DOTS یک برنامه عملیاتی است که تمام موارد از قبیل وضعیت کلی اپیدمیولوژی سل در ایران و جهان، تعاریف استاندارد بروز و شیوع سل و شاخص‌های استاندارد پایش برنامه سل و برنامه‌های کلی مبارزه با سل را در بر می‌گیرد. آزمایشگاه‌های مخصوص سل در هر شهرستان مسئول گزارش موارد اسمیر مثبت هستند. در ضمن برای تایید تشخیص، علاوه باینی به همراه گزارش آزمایشگاه و تایید پزشک مخصوص سل شهرستان استفاده می‌شود و در این تعریف از CXR فقط برای پیگیری بیماران و بررسی عوارض استفاده می‌شود و هیچ نقشی در تشخیص سل ندارد.

با توجه به اینکه از شاخص‌های مهم در مسایل اپیدمیولوژیک بروز و شیوع بیماری در شهر، استان و استان‌های همجوار و کشورهای مجاور ایران است و از طرفی با توجه به هدف کلی طرح که تحلیل روند بروز بیماری سل می‌باشد، لذا روند تحلیل داده‌های سل بر اساس بروز سالانه و از دید اپیدمیولوژیک در این مقاله در قسمت بحث آورده و بحث شده است.

در این مطالعه از سالنامه‌های آماری کشور برای برآورد آمار جمعیت کشور (مورد نیاز برای بدست آوردن میزان بروز سالانه بیماری سل در کشور طی سال‌های مورد مطالعه) و تعداد بیماران مسلول ثبت شده در واحد بیماریه‌های واگیردار معاونت سلامت



نمودار شماره ۱ - روند تغییرات میزان بروز سالانه سل در ایران طی سالهای ۸۷-۱۳۴۳

با مقایسه میزان بروز سل ریوی اسمیر مثبت طی سالهای ۸۷ و ۸۰ در سطح استان‌های کشور مشاهده شد که این میزان در تمامی استان‌ها به جز استان‌های همدان، کردستان و مرکزی کاسته شده است. بیشترین و کمترین کاهش با ۷۹٪ و ۳٪ به ترتیب مربوط به استان کهگیلویه و بویر احمد و دو استان فارس و خوزستان به طور همزمان بوده است.

با توجه به مقادیر بدست آمده برای آماره‌ی معیار اطلاع بیزی (BIC) که در جدول (۱) نشان داده شده است، از بین دو مدل رگرسیون قطعه‌ای و رگرسیون پواسون، مدل رگرسیون قطعه‌ای با دو نقطه شکست را بدلیل دارا بودن کمترین مقدار آماره به عنوان بهترین مدل انتخاب کردیم.

جدول شماره ۱ - معیار اطلاع بیزی برای برازش مدل‌های رگرسیون پواسون و رگرسیون قطعه‌ای

تعداد نقاط شکست	معیار اطلاع شوارتز (BIC)	تغییرات معیار اطلاع شوارتز $\Delta BIC$
بدون نقطه‌ی شکست	۲/۴۱۵۳	
یک نقطه شکست	۰/۵۳۶۰۱	-۱/۸۷
دو نقطه شکست	۰/۳۵۷۶۸	-۰/۱۷

نتایج حاصل از برآورد پارامترها با استفاده از روش حداقل مربعات و بهره‌گیری از دو تکنیک لرنن و هادسون نشان داد که مقادیر برآورد شده از هر دو روش تقریباً مشابه می‌باشند (جدول شماره ۲).

سال گذشته (۸۷-۱۳۴۳) با استفاده از مدل رگرسیون قطعه‌ای مورد بررسی و تحلیل قرار گرفت. برای برازش مدل رگرسیون قطعه‌ای از روش حداقل مربعات (هادسون، لرنن و آزمون جایگشت) و به دلیل شمارشی بودن ماهیت داده‌های جمع‌آوری شده (تعداد بیماران مسلول) از مدل پواسن و برآوردگرهای کمترین توان‌های دوم برای برآورد پارامترهای رگرسیونی قطعات و در تجزیه و تحلیل اطلاعات از نرم افزارهای Joinpoint 3.4 و SAS استفاده شد (۱۴). همچنین تعداد و مکان‌های نقاط تغییر در روند بیماری سل و همچنین برآورد پارامترهای رگرسیونی قطعات با استفاده از تکنیک‌های لرنن، هادسون و آزمون جایگشت برآورد شده و از معیار اطلاع بیزی برای انتخاب بهترین مدل از بین مدل‌های رگرسیون قطعه‌ای و مدل پواسون استفاده گردید.

## یافته‌ها

برازش داده‌های میزان بروز سالانه بیماری سل با سه نقطه تغییر نشان داد که تنها دو نقطه شکست معنی دار با مقدار  $p=0/0108$  در ۴۴۹۹ تکرار آزمون جایگشت به ترتیب در سال‌های ۱۳۵۶ با فاصله اطمینان (۵۷-۱۳۵۴) و ۱۳۷۲ با فاصله اطمینان (۸۰-۱۳۶۷) وجود داشته است.

بنابراین تابع پاسخ میانگین رگرسیون قطعه‌ای برای میزان بروز سل برابر

$$E(y|x) = 136.29 - 8.655x + 9.24(x-1356)^+ - 1.40(x-1372)^+$$

بدست آمد. درصد تغییرات سالانه (APC)<sup>۱</sup> میزان بروز بیماری سل برای قطعه اول (۵۵-۱۳۴۳) برابر ۱۰/۱-٪ و با فاصله اطمینان (۱۱/۴-، -۸/۸)، برای قطعه‌ی دوم (۷۱-۱۳۵۶) برابر ۴/۳٪ با فاصله اطمینان (۶/۶، ۲)، و برای قطعه سوم (سال‌های ۸۷-۱۳۷۲) برابر ۴/۵-٪ با فاصله اطمینان (۲/۸-، -۶/۱) بدست آمد. همچنین میانگین درصد تغییرات سالانه (AAPC)<sup>۲</sup> برای پنج و ده سال اخیر برابر ۴/۵-٪ محاسبه شد. میزان بروز بیماری سل از ۱۴۲ مورد در یکصد هزار نفر جمعیت در سال ۱۳۴۳ به ۱۳/۴ در یکصد هزار نفر جمعیت در سال ۱۳۸۷ کاهش یافته است (بیش از ده برابرکاهش). بیشترین و کمترین میزان بروز بیماری سل از سال ۱۳۵۵ به بعد با ۳۲/۱ و ۱۳/۲ در یکصد هزار نفر جمعیت به ترتیب مربوط به سال‌های ۱۳۷۱ و ۱۳۸۶ بوده است (نمودار شماره ۱).

<sup>۱</sup> Annual percent change

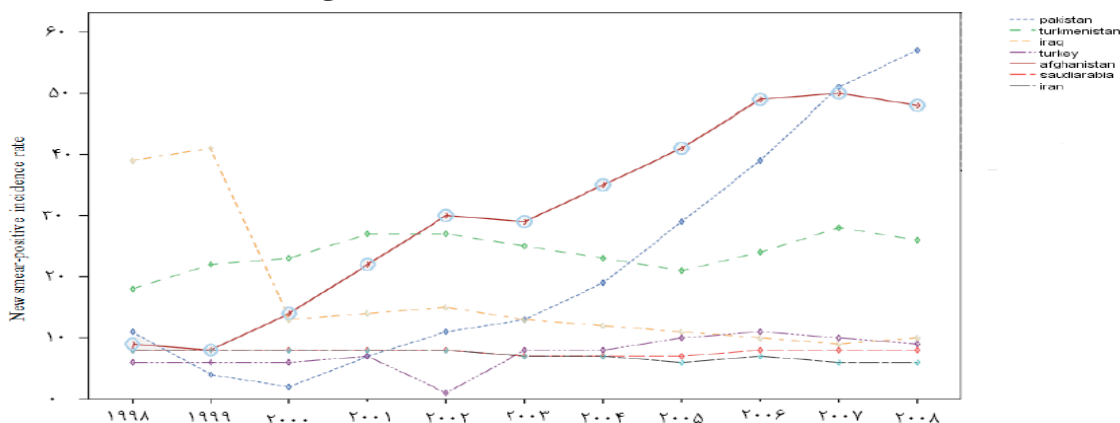
<sup>۲</sup> Average annual percent change

جدول شماره ۲- پارامترهای برآورد شده مدل رگرسیون قطعه‌ای

روش	پارامتر	برآورد پارامترها	خطای استاندارد	P
جستجو شبکه لرمین	b0	۱۳۶/۲۹	۵/۶۹۳	<۰/۰۰۰۱
	b1	-۸/۶۵۵	۰/۵۹۰	<۰/۰۰۰۱
	b1 - b2	۹/۲۴۳	۰/۶۵۳	<۰/۰۰۰۱
	b2 - b3	-۱/۴۰۵	۰/۳۷۹	۰/۰۰۰۶
الگوریتم پیوسته هادسون	b0	۱۳۷/۷۸	۵/۶۹۳	<۰/۰۰۰۱
	b1	-۸/۸۸	۰/۵۸۴	<۰/۰۰۰۱
	b1 - b2	۹/۴۲۷	۰/۶۳۶	<۰/۰۰۰۱
	b2 - b3	-۱/۳۵۱	۰/۳۵۷	۰/۰۰۰۵

سایر مناطق بروز سل ریوی اسمیر مثبت از روند ثابتی برخوردار بوده است. تحلیل داده‌های دو استان سیستان و بلوچستان و گلستان با بیشترین میزان بروز سل ریوی اسمیر مثبت با استفاده از رگرسیون قطعه‌ای نشان داد که میانگین تغییرات میزان بروز سالانه پنج سال اخیر (سال‌های ۱۳۸۳ تا ۸۷) برای استان سیستان و بلوچستان برابر ۱/۷٪ با فاصله اطمینان (۵/۹،۹/۹-) معنی‌دار نشد ولی طی سال‌های ۸۵-۸۰ برابر ۱۰/۱- درصد و معنی‌دار بود و برای استان گلستان برابر ۱/۷-٪ با فاصله اطمینان (۴/۱،۵/۴-) معنی‌دار نشد.

تحلیل روند تغییرات میزان بروز سل ریوی اسمیر مثبت به تفکیک دانشگاه‌های کشور طی سال‌های ۸۷-۱۳۸۰ نشان می‌دهد که بروز سالانه بیماری در رفسنجان، لرستان، هرمزگان، اردبیل، بیرجند، آذربایجان غربی، آذربایجان شرقی، کرمان، اصفهان، چهار محال بختیاری، مازندران، سمنان، زابل و خراسان کاهش یافته است، بیشترین کاهش در روند بروز طی سال‌های ۸۷-۱۳۸۳ مربوط به رفسنجان با ۲۰/۳ درصد و کمترین کاهش مربوط به خراسان با ۳/۶ درصد در سال می‌باشد. بروز بیماری در شهرستان فسا با APC برابر ۱۹/۹+ روند صعودی داشته و برای



نمودار شماره ۲ - روند میزان بروز بیماری سل ریوی اسمیر مثبت در کشورهای منطقه، ۱۹۹۸-۲۰۰۸

میزان بروز بیماری سل کشور برای قطعه اول تا حدودی دور از انتظار است، که می‌تواند به دلیل نقص در ثبت داده‌ها یا دقیق نبودن روش‌های تشخیص بیماری باشد، بنابراین تمرکز بر داده‌های بدست آمده از قطعه دوم و سوم می‌تواند مفیدتر واقع گردد. میزان بروز سل در ایران طی ۱۰ سال اخیر روند نزولی داشته و با توجه به نتایج به دست آمده در این بررسی می‌توان گفت میزان بروز

## بحث

با توجه به نتایج درصد تغییرات سالانه (APC) و میانگین درصد تغییرات سالانه (AAPC) می‌توان گفت که روند بروز بیماری سل تا سال ۵۵ سیر نزولی و از سال ۱۳۵۵ تا ۷۱ سیر صعودی و از سال ۷۱ به بعد سیر نزولی داشته است. روند مشاهده شده در

ریوی اسمیر مثبت کل کشور ۵۰/۱۷٪ موارد زن بودند که بیانگر بالا بودن شیوع بیماری سل در زنان نسبت به مردان در استان اردبیل طی سال‌های مطالعه بوده است. در مطالعه ۱۳ ساله در سالهای ۸۰-۱۳۷۵ از کل بیماران ۴۶/۶٪ شهری و ۵۳/۴٪ روستایی بودند (۱۸).

بر اساس گزارش معاونت سلامت، از تعداد ۹۷۴۲ مورد مبتلا به سل گزارش شده در کشور در سال ۱۳۸۷، ۵۰٪ موارد را زنان و ۱۳٪ موارد را بیماران افغانی تشکیل داده بودند و بیشترین میزان بروز سل مربوط به گروه سنی ۶۵ سال به بالا بوده است (۱۶). همچنین بین سال‌های ۸۷-۱۳۵۶، بیشترین تعداد مسلولین شناخته شده مربوط به سال ۷۱ با ۱۵۰۲۱ مورد و کمترین مربوط به سال ۸۶ با ۹۴۲۳ مورد بود.

با توجه به اینکه سل بیماری فقر، کمبود اقتصادی و کمبود دانش می‌باشد از این نقطه به نظر می‌رسد سیر کاهش یافته در منحنی بروز مربوط به افزایش آگاهی‌های عمومی و افزایش درآمدهای ناخالص ملی طی سال‌های ۷۷-۱۹۶۴ باشد از طرفی افزایش دوباره بروز طی سال‌های ۹۳-۱۹۷۸ می‌تواند منعکس کننده تغییرات سیاسی و اجتماعی در داخل و اوضاع همسایگان مخصوصاً افغانستان باشد که با ورود بسیار زیاد مهاجران افغان به کشور ایران در طی سال‌های اشغال کشور افغانستان باشد.

کاهش در قطعه سوم طی سال‌های ۲۰۰۸-۱۹۹۴ نیز احتمالاً برآیند کنترل فاکتورهای مستعد کننده‌ای نظیر افزایش مراقبت‌های داخلی PHC و کاهش مهاجران و پایداری وضعیت درمان سل در کشور باشد. همچنین MDR مشکل اصلی سیستم ایران نیست و موارد اسمیر منفی نیز در انتقال بیماری موثر نیستند چون تشخیص قطعی نبوده و به صورت کلینیکی تشخیص داده شده‌اند لذا در این مطالعه آورده نشده است. در این بررسی تنها روند میزان بروز بیماری سل مورد بررسی قرار گرفت و تحلیل در سطح گروه‌های سنی و همچنین به تفکیک شهری و روستایی و بروز بر اساس استان طی سال‌های مورد مطالعه بدلیل وجود داده‌های مربوطه امکان‌پذیر نبود.

### نتیجه‌گیری

با توجه به نتایج بدست آمده از مدل رگرسیون قطعه‌ای می‌توان گفت، الگوی میزان بروز بیماری سل از سال ۱۳۷۱ به بعد روند نزولی داشته، که این کاهش از سرعت بالاتری در مقایسه با میانگین جهانی برآورد شده دارد، که نشان دهنده ارتقاء برنامه‌های کنترل بیماری سل، چشم‌گیر بودن سیر نزولی بیماری، بهبود

بیماری سل برای استان‌های سیستان و بلوچستان و گلستان که بیشترین موارد گزارش بیماری سل را به خود اختصاص داده‌اند، در سال ۸۷ نسبت به سال ۸۰ کاهش یافته است.

مطالعه انجام شده بین سال‌های ۱۹۹۷ تا ۲۰۰۶ بر روی ۱۳۴ کشور برای بررسی روند بروز بیماری سل، نشان داد که میزان بروز سالانه بیماری سل در این کشورها با دامنه  $\pm 10$  تغییر نموده و میانگین میزان تغییرات سالانه برابر ۱/۹-٪ بود و می‌شد گفت که به طور متوسط، کاهش در کشورهای با درآمد بالا سریعتر بوده است (سالانه ۳/۹-٪) (۱۵). کشور ایران با AAPC برابر ۴/۵٪، از نظر کاهش در بروز سل بالاتر از میانگین گزارش شده برای کشورهای با درآمد بالا با ۳/۹٪ قرار داشت. مقایسه میانگین درصد میزان بروز سالانه سل به تفکیک دانشگاه‌های کشور نشان داد که رفسنجان و خراسان با AAPC برابر ۲۰/۳-٪ و ۳/۶-٪ به ترتیب بیشترین و کمترین سرعت را در کاهش بروز سل داشتند. با توجه به اینکه مهاجرت نقش مهمی را در انتقال بیماری دارد، و همچنین با توجه به اینکه دو کشور افغانستان و پاکستان که جزو ۲۲ کشور دارای بار بالای بیماری در دنیا هستند و همچنین وجود عراق و کشورهای تازه استقلال یافته در شمال کشور ایران ضرورت توجه بیش از پیش به این بیماری را معلوم می‌سازد (نمودار شماره ۲).

بنابر اعلام معاونت سلامت وزارت بهداشت حدود ۱۳٪ موارد بیماری مربوط به مهاجران افغانی بوده به طوریکه بیماری سل ریوی اسمیر مثبت در کشورهای افغانستان و پاکستان دارای روند صعودی به ترتیب با میزان بروز ۲۳/۱٪ و ۳۳/۴٪ در سال ۲۰۰۸ حدود ۷ و ۹ برابر ایران با ۳/۶-٪ بودند، می‌توان این تصور را از داده‌ها داشت که خود مهاجرین از کشورهای همسایه منطقه نیز به نوبه خود می‌توانند در تغییرات روند بروز بیماری سل کشور ایران موثر واقع گردند و بنابراین به نظر می‌رسد کنترل مهاجرین از نظر ابتلا به بیماری سل و تدوین برنامه‌های منسجم جهت پیگیری و درمان این افراد، می‌تواند نقش کلیدی در ریشه‌کنی این بیماری داشته باشد و از طرفی با توجه به بیشترین میزان مشاهده شده در سال ۸۷ در شهرستان زابل با ۱۰۳/۹ مورد در ۱۰۰۰۰ نفر، به نظر می‌رسد که همسایگی ایران با پاکستان و افغانستان در این میزان نقش داشته است (۱۶).

در مطالعه انجام شده در سال‌های ۸۴-۱۳۸۱ در شهرستان اردبیل از ۸۶ بیمار اسمیر ریوی مثبت (۳۷/۱٪ کل موارد)، ۵۹/۳٪ موارد زن با میانگین سنی ۴۲ سال و انحراف معیار ۱۹/۴۵ سال بودند (۱۷). در حالی که در سال‌های ۸۴-۱۳۸۱ از ۲۰۹۸۰ بیمار

مطالعاتی در زمینه تحلیل روند میزان‌های بروز بیماری سل و همچنین مدل رگرسیون قطعه‌ای خطی در ایران برای مقایسه نتایج حاضر اشاره کرد و پیشنهاد می‌شود که در آینده مطالعات اپیدمیولوژیکی دیگری روی بیماری سل با استفاده از سایر مدل‌های خطی انجام شود.

وضعیت بهداشتی و ارتقای شاخص‌های بهداشتی در کشور ایران است. به طور کلی می‌توان گفت، روند بروز سل در ایران نزولی و به صورت غیر خطی بوده و در دو مقطع زمانی شکست‌هایی در خط مربوطه وجود داشته است، همچنین این کاهش بروز بیماری در مقایسه با کشورهای همجوار و وضعیت جهانی سرعت بیشتری داشته است. از محدودیت‌های مطالعه حاضر می‌توان به عدم وجود

## منابع

- 1- Azizi F, Janghorbani M, Hatami H. Epidemiology and control of prevalent diseases in Iran. 2th ed, Tehran, Khosravi press; 2004.
- 2- Nasehi M, Mirhaghaani L. Guideline for control TB. 1 th ed, Tehran, Andishmand press; 2009:5-21.
- 3- [http://www.who.int/tb/publications/global\\_report/2009/update/en/index.html](http://www.who.int/tb/publications/global_report/2009/update/en/index.html)
- 4- Kim HJ, Fay MP. Comparability of Segmented Line Regression Models. *Biometrics* 2004; 1005-14.
- 5- Robison DE. Estimates for the points of intersection of two polynomial regressions. *Journal of the American Statistical Association* 1964; 59: 214-24.
- 6- Sprent P. Some hypothesis concerning two phase regression lines. *International Biometric Society* 1961 Dec; 17: 634-45.
- 7- Berman NG. Applications of segmented regression models for biomedical studies. *American Journal of Physiology* 1996; 270: 723-73.
- 8- Lerman PM. Fitting segmented regression models by grid search. *Applied Statistics* 1980; 29: 77-84.
- 9- Hudson DJ. Fitting segmented curves whose join points have to be estimated. *Journal of the American Statistical Association* 1966; 61: 1097-129.
- 10- H L M, Polanczyk R, A Batista. Field persistence of bacillus thuringiensis on maize leaves. *Brazilian Journal of Microbiology* 2005; 36: 309-314.
- 11- Bouadma L, Deslandes E, Lolom I, Corre BL. Long Term Impact of a Multifaceted Prevention Program on Ventilator Associated Pneumonia in a Medical Intensive Care Unit. *Clinical Infectious Diseases* 2010; 51: 1115-22.
- 12- Lava B, Gangea S, Phair J, Riddler S, Detels R. Rapid declines in total lymphocyte counts and hemoglobin concentration prior to AIDS among HIV-1-infected men. *AIDS* 2003; 17: 2035-44.
- 13- Kim HJ, Yu B, Feuer EJ. Selecting the number of change points in segmented line regression. *Statistica Sinica* 2009; 19: 597-609.
- 14- Joinpoint Regression Program, Version 3.4 - April 2009; Statistical Methodology and Applications Branch and Data Modeling Branch, Surveillance Research Program National Cancer Institute. Available from: <http://surveillance.cancer.gov/joinpoint/>
- 15- C Dye, K Lonnroth, E Jaramillo, BG Williams & M Raviglione. Trends in Tuberculosis Incidence and their determinants in 134 countries. *Bull World Health Organ* 2009; 87: 683-91.
- 16- Hasanloo GH, Nasehi F, Mahmoodi E. Periodical news of Communicational center unit of ministry of health and education 2009; 1:1-4.
- 17- Amani F, Bashiri J, Garousee B, Nahanmoghadam N. epidemiologic study of TB in Ardabil city in years 2002-2005. *J of Ardabil university medical science* 2007; 7: 236-41.
- 18- Rajabi A, Abazari R. Study epidemiology of TB in Bam city in years 1996-2002. *Iranian Journal of Infectious Diseases and Tropical Medicine* 2003; 8: 41-6.

Iranian Journal of Epidemiology 2011; 7(3): 6-12.

**Original Article**

# Applying Segmented Regression Model to Analysis the Trend of Tuberculosis Incidence Rate in Iran Between 1964 –2008

Arsang SH<sup>1</sup>, Kazemnejad A<sup>1</sup>, Amani F<sup>2</sup>

1- Department of Biostatistics, Faculty of Medicine, Tarbiat Modares University, Tehran, Iran

2- Department of Biostatistics, Faculty of Medicine, Ardabil University of Medical Science, Ardabil, Iran

**Corresponding author:** Kazemnejad A., kazem\_an@modares.ac.ir

**Background & Objectives:** Study trend of observed rates changes provide valuable information for need assessment, plan, reload programs and develop indicators of each country. The main objective of this paper is to determine the changes in tuberculosis incidence rate trend in Iran by applying segmented regression model.

**Methods:** In this study, segmented Linear Regression employed to analyze the trend of changes in pattern of Tuberculosis incidence rate during past 44 years (1964-2008) in Iran. We used least square method and permutation test and Bayesian Information Criteria to decide which of the two segment regression model and poisson regression would be better. Data analyzed by Joinpoint3.4 and SAS9.1 software.

**Results:** According the permutation test, it was detected that there were two breakpoints over 1977 and 1993 years ( $p=0.0108$ ). Incidence rate of tuberculosis during the first 11 years of review had declined with annual percentage change = -10.1%, for second segment it rose upward with 4.3% increase in per year and for end segment TB incidence rate again declined with annually 4.5%. The average annual change of Tuberculosis incidence rate in Iran for at least 10 years has been estimated -4.5 percentages.

**Conclusion:** The findings of this study have shown that the incidence rate of Tuberculosis decreased after 1992 that interestingly this decline seems faster than estimated by international TB control program. This indicates that preventive and treatment of Tuberculosis programs have been successful in Iran.

**Keywords:** Segmented regression, Break point, Permutation test, Annual percent change (APC)