

ارتباط شاخص‌های اجتماعی - اقتصادی با بیماری سل با استفاده از مدل‌های رگرسیونی دوجمله‌ای منفی و پواسون

فاطمه سروی^۱، دکتر یداله محرابی^{۲*}، دکتر علیرضا ابدی^۳، دکتر مهشید ناصحی^۴، ابوالفضل پاینده^۵

۱- کارشناس ارشد آمار زیستی، گروه آمار زیستی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهیدبهشتی. ۲- استاد، گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شهیدبهشتی. ۳- دانشیار، گروه پزشکی اجتماعی، دانشکده پزشکی، دانشگاه علوم پزشکی شهیدبهشتی. ۴- استادیار، گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی ایران. ۵- دانشجوی دکتری آمار زیستی، گروه آمار زیستی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهیدبهشتی.

چکیده

زمینه و هدف: بیماری سل بزرگ‌ترین علت مرگ ناشی از بیماری‌های عفونی تک عاملی است. از دلایل اصلی افزایش بار جهانی سل، فقر و اختلاف شدید طبقاتی جوامع مختلف است. این مطالعه به منظور تعیین ارتباط شاخص‌های اجتماعی - اقتصادی با بیماری سل با استفاده از مدل‌های رگرسیونی دوجمله‌ای منفی و پواسون انجام شد.

روش بررسی: این مطالعه توصیفی - تحلیلی روی پرونده ۱۱۳۲۰ مسلول سال ۱۳۸۹ از شهرهای مختلف انجام شد. داده‌ها از اداره کنترل بیماری‌های سل و جدام معاونت بهداشت، درمان و آموزش پزشکی و شاخص‌های اجتماعی - اقتصادی از مرکز آمار ایران دریافت شد. شاخص‌های اجتماعی - اقتصادی شامل میزان بیکاری، میزان بیسودی، تراکم خانوار در هر واحد مسکونی، متوسط درآمد، تعداد پزشکان به تعداد نقاط جمعیتی، میزان مهاجرت و ضریب شهرنشینی بود. برازش مدل‌ها با استفاده از معیار اطلاع آکائیک و بیزی مقایسه شد. **یافته‌ها:** با توجه به مدل رگرسیونی پواسون بین شاخص‌های اجتماعی - اقتصادی و میزان ابتلا به سل ارتباط آماری معنی‌داری یافت شد ($P < 0/05$). با توجه به مدل رگرسیونی دوجمله‌ای منفی بین میزان بیکاری، میزان بیسودی، میزان مهاجرت و ضریب شهرنشینی با ابتلا به بیماری سل ارتباط آماری معنی‌دار بود ($P < 0/05$) و برای تعداد پزشکان به تعداد نقاط جمعیتی و تراکم خانوار به واحد مسکونی و متوسط درآمد سالانه خانوار و ابتلا به بیماری سل ارتباط آماری معنی‌داری به دست نیامد. به دلیل وجود بیش پراکنش در مدل رگرسیونی پواسون، مدل رگرسیونی دوجمله‌ای منفی مناسب‌تر ارزیابی شد.

نتیجه‌گیری: شاخص‌های اجتماعی - اقتصادی با موارد ابتلا به بیماری سل مرتبط است. همچنین استفاده از مدل رگرسیونی دوجمله‌ای منفی برای داده‌های شمارشی در مقایسه با مدل رگرسیونی پواسون (دارای بیش پراکنش) ارجحیت داشت.

کلید واژه‌ها: سل، شاخص‌های اجتماعی - اقتصادی، مدل رگرسیونی دوجمله‌ای منفی، مدل رگرسیونی پواسون

* نویسنده مسؤول: دکتر یداله محرابی، پست الکترونیکی mehrabi@sbmu.ac.ir

نشانی: تهران، ولنجک، دانشگاه علوم پزشکی شهیدبهشتی، دانشکده بهداشت، گروه اپیدمیولوژی، تلفن و نمابر ۰۲۱-۲۲۷۰۷۳۴۷

وصول مقاله: ۹۲/۹/۱۱، اصلاح نهایی: ۹۳/۱/۳۱، پذیرش مقاله: ۹۳/۱/۳۱

مقدمه

توسعه رخ می‌دهد. کشورهایی که ۷۵ درصد موارد بیماری در آنها به فعال‌ترین گروه سنی (۶۴-۱۵ سال) از لحاظ اقتصادی تعلق دارد (۲). در مطالعه ارسنگ و همکاران بیشترین موارد بیماری سل در افراد بالای ۶۵ سال مشاهده شد (۳). سل به‌جز لطامات اقتصادی، اثرات منفی غیرمستقیم دیگری نیز بر کیفیت زندگی بیماران یا افراد خانواده آنها دارد و از این نظر دارای مرتبه دهم در بار جهانی بیماری‌هاست. پیش‌بینی شده این بیماری تا سال ۲۰۲۰ میلادی همچنان جایگاه کنونی خود را حفظ نموده و یا از رتبه دهم به رتبه هفتم خواهد رسید (۱).

فقر و اختلاف شدید طبقاتی در جوامع مختلف، نه فقط در

بیماری سل که بزرگ‌ترین علت مرگ ناشی از بیماری‌های عفونی تک‌عاملی است؛ به‌وسیله مایکوباکتریوم توبرکلوزیس ایجاد می‌شود. در سل ریوی فقط ریه‌ها مبتلا می‌شوند و در سل خارج ریوی، اندام‌های دیگر بدن مبتلا می‌شوند (۱). برطبق گزارش سازمان بهداشت جهانی بهداشت تقریباً یک‌سوم جمعیت جهان به مایکوباکتریوم توبرکلوزیس یا آلوده هستند و یا در خطر ابتلا به آن قرار دارند. هر ساله حدود ۹ میلیون نفر به سل فعال مبتلا شده و ۱/۵ تا ۲ میلیون نفر در اثر این بیماری جان می‌سپارند. بیش از ۹۰ درصد موارد بیماری و مرگ ناشی از سل در کشورهای در حال

مهاجران آسیایی، آفریقای و اطراف اقیانوس آرام نسبت به مهاجران بقیه کشورها بیشتر بود. در اولین ۶ سال مهاجرت، در هر ۱۰۰ هزار نفر ۱۰۰ تا ۱۵۰ نفر مبتلا وجود داشت (۱۱).

این مطالعه به منظور تعیین ارتباط شاخص‌های اجتماعی - اقتصادی با بیماری سل با استفاده از مدل‌های رگرسیونی دوجمله‌ای منفی و پواسون انجام شد.

روش بررسی

این مطالعه توصیفی - تحلیلی گذشته نگر روی پرونده ۱۱۳۲۰ مسلول توسط گروه آمار زیستی دانشکده پیراپزشکی دانشگاه علوم پزشکی شهیدبهشتی طی سال‌های ۹۳-۱۳۹۲ انجام شد.

داده‌های مسلولین سال ۱۳۸۹ به تفکیک شهرستان (۳۵۹ شهرستان از ۴۲۰ شهرستان)، از اداره کنترل بیماری‌های سل و جذام معاونت بهداشت، درمان و آموزش پزشکی و شاخص‌های اجتماعی - اقتصادی از مرکز آمار ایران دریافت شد.

شاخص‌های اجتماعی - اقتصادی شامل میزان بیکاری، میزان بیسوادی، تراکم خانوار در هر واحد مسکونی، متوسط درآمد، تعداد پزشکان به تعداد نقاط جمعیتی، میزان مهاجرت و ضریب شهرنشینی بود. از آنجایی که شاخص تعداد پزشکان به تعداد نقاط جمعیتی با شاخص مجموع مراکز بهداشتی به تعداد نقاط جمعیتی هم‌خطی داشت؛ شاخص مجموع مراکز بهداشتی به تعداد نقاط جمعیتی در مطالعه وارد نشد.

برای مدل‌بندی متغیر شمارشی تعداد افراد مسلول از مدل خطی تعمیم‌یافته پواسون استفاده شد. یکی از مفروضات اساسی مدل رگرسیون پواسون مساوی بودن میانگین و واریانس متغیر پاسخ است؛ ولی معمولاً در عمل، پراکندگی داده‌ها از میانگین آنها بیشتر است. به این موضوع بیش پراکنش اطلاق می‌شود. دلیل بیش پراکنش معمولاً ناهمگنی زیاد بین پاسخ‌ها است (۱۲ و ۱۳).

میانگین تعداد مسلولین هر شهرستان ۳۱/۵۳ و واریانس آن ۶۰۵۴/۲۱ حاصل شد که نشان دهنده بیش پراکنش بود. یکی از راه کارها برای از بین بردن بیش پراکنش استفاده از مدل رگرسیون دوجمله‌ای منفی به عنوان جایگزینی برای مدل پواسون است. زیرا این مدل یک پارامتر جداگانه برای واریانس منظور می‌کند و با این مدل برآورد پارامترها دقیق‌تر است.

زمانی که متغیر پاسخ شمارشی وابسته به متغیر دیگری مانند t باشد؛ می‌توان میزان مورد انتظار را محاسبه کرد و مدل رگرسیونی پواسون و همچنین مدل دوجمله‌ای منفی برای میزان مورد نظر را برآش داد. به دلیل وابستگی تعداد مسلولین هر شهرستان به جمعیت شهرستان، ابتدا از تقسیم تعداد افراد به جمعیت هر شهرستان، میزان مورد نظر به دست آمد و سپس مدل‌های مورد نظر برآش داده شد (۱۳).

کشورهای در حال رشد بلکه حتی در کشورهای پیشرفته نیز یکی از دلایل اصلی افزایش بار جهانی سل است. در کشور ما به دلیل محرومیت اجتماعی و اقتصادی بعضی شهرستان‌ها، پیدا کردن ارتباط این عوامل با ابتلا به بیماری سل برای سیاست‌گذاری‌های سلامت اهمیت عمده‌ای دارد (۱). سل و فقر ارتباط دوسویه دارند. به طوری که افزایش فقر و شرایط نامناسب زندگی میزان بروز سل را می‌افزاید. از طرفی بیماری سل و هزینه درمان بالا به زایش فقر کمک می‌کند. احتمال آلودگی به بیماری سل و فعال شدن این بیماری در جوامع با وضعیت اجتماعی - اقتصادی پایین‌تر، به دلیل زندگی در نواحی پرجمعیت با فرزندان بیشتر و داشتن شرایط زیست محیطی نامطلوب، بیشتر است (۴).

در مطالعه Vendramini و همکاران در برزیل شاخص‌های اجتماعی - اقتصادی مرتبط با اپیدمی بیماری سل، تحصیل، درآمد و تعداد افراد در هر خانوار تعیین شد. سل با شرایط زندگی افراد ارتباط نزدیکی داشت و شیوع بیماری سل برای نواحی فقیرتر بالاتر بود (۵). در مطالعه Hawker و همکاران به تفاوت نژادی و ارتباط آن با بیماری سل و فقر پرداخته شد. در بین افراد با نژاد سفید و افراد مهاجر از کشورهای آسیای جنوبی که در ۳۹ منطقه محروم بیرمینگهام ساکن بودند؛ ۱۵۱۶ مسلول شناسایی گردید. بین نسبت ابتلا به سل و محرومیت در افراد آسیای جنوبی ارتباط معنی‌داری دیده نشد؛ اما این ارتباط با نژاد سفید مرتبط بود (۶).

در مطالعه Sánchez-Pérez و همکاران در شهر Chiapas مکزیک از بین ۲۲۷ مسلول، افراد با سطح فقر بالاتری به بیماری سل مبتلا بودند (۷). در مطالعه Bhunu و همکاران مدل جبری برای ارزیابی اثر شرایط اجتماعی - اقتصادی در انتقال بیماری سل استفاده شد و نتایج نشان داد که انتقال بیماری سل با شرایط زندگی نامناسب و فقر افزایش نشان داده است (۸).

با افزایش شهرنشینی میزان ابتلا به بیماری سل به خاطر تراکم زیاد جمعیت و تماس زیاد افراد، افزایش می‌یابد (۹). از طرفی میزان بالاتری از بیماری سل در بین گروه‌های مهاجر وجود دارد (۱۰).

در مطالعه Gilbert و همکاران انجام شده روی مهاجرت و میزان ابتلا به بیماری سل در انگلستان در مقایسه با ۲۱ کشور اروپایی نشان داد که میزان ابتلا به سل در سه کشور نروژ، سوئد و انگلستان روند افزایشی داشته است. در هر سه کشور تقریباً ۰/۷۵ درصد مبتلایان به سل، در خارج از کشور و عموماً در کشورهایی با میزان بالای بروز بیماری سل متولد شده بودند (۱۰).

در مطالعه Denholm و McBryde میزان خطر بروز بیماری سل فعال در مهاجران ایالت ویکتوریای استرالیا در سال ۲۰۱۲ بررسی شد. بسته به این که فرد مهاجر از چه کشوری و از چه ناحیه‌ای مهاجرت کرده بود؛ میزان خطر بروز سل نوسان داشت. این میزان در

جدول ۱: میانگین و انحراف معیار شاخص‌های اجتماعی - اقتصادی مسلولین در سال ۱۳۸۹

میانگین و انحراف معیار	حداکثر مقدار	حداقل مقدار	
۶/۷۵±۰/۰۷	۳۳/۷۰	-۱۷/۸۰	میزان مهاجرت
۰/۲۰±۰/۰۵۲	۰/۹۹	۰/۰۸	ضریب شهرنشینی
۹/۷۸±۱۴/۵۲	۵۹/۵۰	۰/۱۵	میزان بیکاری
۶/۱۸±۱۹/۳۷	۴۷/۴۶	۲/۸۱	میزان بیسوادی
۹۶/۷۵±۱۲/۵۵	۱۳۲/۲۸	۷۴/۴۷	متوسط درآمد سالانه برای هر خانوار شهرستانی
۴/۱۰۱±۵۷/۳۵	۱۰۸۱/۲۰	۰/۰۰	نسبت پزشک به تعداد نقاط جمعیتی
۰/۱۶۵±۱/۱۳	۳/۱۵	۱/۱۲	تراکم خانوار در هر واحد مسکونی

جدول ۲: نتایج مدل‌های رگرسیونی پواسون و رگرسیونی دوجمله‌ای منفی با شاخص‌های اجتماعی - اقتصادی و تعداد مسلولین سال ۱۳۸۹

p-value	فاصله اطمینان ۹۵ درصد exp (β)	خطای استاندارد	exp (β)	β	برآورد ضریب β	
۰/۰۰۱	۰/۰۰۶ و ۰/۰۱۱	۰/۰۰۱۴	۱/۰۱۹	۰/۰۱۹	میزان مهاجرت	
۰/۰۰۱	۱۲/۳۱ و ۱۵/۹۵	۰/۰۶۶	۱۴/۰۱	۲/۶۴	ضریب شهرنشینی	
۰/۰۰۱	۱/۰۱۳ و ۱/۰۱۸	۰/۰۰۱۱	۱/۰۱۵	۰/۰۱۵	میزان بیکاری	مدل
۰/۰۰۱	۱/۰۳۲ و ۱/۰۴۱	۰/۰۰۲۲	۱/۰۰۴	۰/۰۳۶	میزان بیسوادی	رگرسیونی
۰/۰۰۱	۱/۰۰۲ و ۱/۰۰۵	۰/۰۰۰۸	۱/۰۰۳	۰/۰۰۳	متوسط درآمد سالانه برای هر خانوار شهرستانی	پواسون
۰/۰۰۱	۱/۰۰۱ و ۱/۰۰۱	۰/۰۰۰۰۴۷	۱/۰۰۱	۰/۰۰۱	نسبت پزشک به تعداد نقاط جمعیتی	
۰/۰۰۱	۱/۰۶۱ و ۱/۴۱۸	۰/۰۷۴	۱/۲۲۷	۰/۲۰۴	تراکم خانوار در هر واحد مسکونی	
۰/۰۰۷	۱/۰۰۷ و ۱/۰۴۵	۰/۰۰۹۵	۱/۰۲۶	۰/۰۲۵	میزان مهاجرت	
۰/۰۰۱	۳/۳۵ و ۱۳/۴۸	۰/۳۵۵۳	۶/۷۲	۱/۹۰۵	ضریب شهرنشینی	مدل
۰/۰۰۸	۱/۰۰۵ و ۱/۰۳۳	۰/۰۰۷	۱/۰۱۹	۰/۰۱۹	میزان بیکاری	رگرسیونی
۰/۰۰۹	۱/۰۰۹ و ۱/۰۶۱	۰/۰۱۲۹	۱/۰۳۵	۰/۰۳۴	میزان بیسوادی	دوجمله‌ای
۰/۳۴۵	۰/۹۹۵ و ۱/۰۱۵	۰/۰۰۵	۱/۰۰۵	۰/۰۰۵	متوسط درآمد سالانه برای هر خانوار شهرستانی	منفی
۰/۱۴۵	۱/۰۰ و ۱/۰۰۳	۰/۰۰۱	۱/۰۰۱	۰/۰۰۱	نسبت پزشک به تعداد نقاط جمعیتی	
۰/۵۷۴	۰/۵۴۵ و ۲/۹۹	۰/۴۳۴۴	۱/۲۸	۰/۲۴۴	تراکم خانوار در هر واحد مسکونی	

به سل افزایش یافت ($P < 0.001$) (جدول ۲).

با توجه به ضرایب برآورده شده مدل رگرسیونی دوجمله‌ای منفی، با افزایش یک واحد در میزان مهاجرت ($P < 0.007$)، ضریب شهرنشینی ($P < 0.001$)، میزان بیکاری ($P < 0.008$) و میزان بیسوادی ($P < 0.009$) میزان ابتلا به سل افزایش آماری معنی داری نشان داد. بین شاخص‌های تعداد پزشکان به تعداد نقاط جمعیتی و تراکم خانوار در هر واحد مسکونی و متوسط درآمد سالانه برای هر خانوار شهرستانی و میزان ابتلا به سل ارتباط آماری معنی دار نبود (جدول ۲).

جدول ۳: نتایج معیارهای انتخاب مدل

بر اساس معیار اطلاع بیزی و معیار اطلاع آکائیک	
BIC	AIC
۱۰۷۸۰/۴۷	۱۰۷۴۹/۴۸
۲۸۸۲/۱۲۹	۲۸۵۱/۰۶۳

BIC: معیار اطلاع بیزی، AIC: معیار اطلاع آکائیک

با توجه به مقایسه انجام شده مدل رگرسیونی دوجمله‌ای منفی با مدل رگرسیونی پواسون از نظر وجود بیش پراکنش با استفاده از آزمون میزان درست‌نمایی؛ مدل رگرسیونی پواسون دارای بیش پراکنش بود و برازش مدل رگرسیونی دوجمله‌ای منفی بسیار

برای تشخیص مناسب تر بودن مدل، معیارهای انتخاب مدل، شامل معیار اطلاع بیزی (Bayesian Information Criterion: BIC) و معیار اطلاع آکائیک (Akaike Information Criterion: AIC) استفاده شد (۱۴). این معیارها نشان می‌دهند که استفاده از یک مدل به چه میزان باعث از دست رفتن اطلاعات می‌شود. با توجه به داده‌ها، گاهی می‌توان چند مدل را برازش داد. مدلی بهتر است که در نتیجه برازش آن به داده‌ها اطلاعات کمتری از دست برود. یعنی میزان AIC و BIC کمتری داشته باشد.

برای تجزیه و تحلیل داده‌ها از نرم‌افزار R (3.0.1) و بسته MASS استفاده گردید. سطح معنی‌داری کمتر از ۰/۰۵ در نظر گرفته شد.

یافته‌ها

۵۴۳۳ نفر (۴۸ درصد) مرد و ۵۸۸۷ نفر (۵۲ درصد) زن بودند. میانگین سنی مسلولین ۴۷/۷۳±۱۴/۲۵ سال بود.

میزان مهاجرت، ضریب شهرنشینی، میزان بیکاری، میزان بیسوادی، متوسط درآمد سالانه برای هر خانوار شهرستانی، نسبت پزشک به تعداد نقاط جمعیتی و تراکم خانوار در هر واحد مسکونی در جدول یک آمده است.

با توجه به ضرایب برآورد شده مدل رگرسیونی پواسون به ازای یک واحد افزایش در شاخص‌های اجتماعی - اقتصادی، میزان ابتلا

وجود شهرنشینی، تماس افراد زیاد بوده و از طرفی تراکم جمعیت زیاد شده و ممکن است از طرفی مهاجرت به شهرها از مناطقی با میزان بالای سل صورت گرفته باشد که زمینه را برای افزایش میزان ابتلا به بیماری سل افزایش می‌افزاید. در مطالعه Kwan و Ernst نیز یکی از دلایل افزایش ابتلا به بیماری سل، افزایش ضریب شهرنشینی ذکر گردید (۹).

در مطالعه ما با افزایش یک واحد در میزان بیکاری و میزان بیسوادی به ترتیب میزان ابتلا به بیماری سل ۰/۰۱۹ و ۰/۰۳۵ افزایش نشان داد. چرا که این افراد اطلاعات کمتری در مورد واکسینه شدن، بهداشت و راه‌های پیشگیری و مبارزه دارند و با عوامل خطر بیشتری مواجه هستند. در مطالعه Vendramini و همکاران بیسوادی (۵) و در مطالعه Coker و همکاران بیکاری (۱۸) باعث افزایش موارد ابتلا به بیماری سل شده بود. در مطالعه Vendramini و همکاران با افزایش درآمد میزان ابتلا به بیماری سل کاهش نشان داد (۵). همچنین در مطالعه Bhunu و همکاران انتقال بیماری سل در جوامع فقیرتر بالاتر گزارش شد (۸).

در مطالعه حاضر شاخص درآمد و ابتلا به بیماری سل با مدل رگرسیونی دوجمله‌ای منفی معنی‌دار نبود. دلیل آن ممکن است به خاطر اثر دیگر شاخص‌ها بوده باشد. از طرفی در شهرستان‌هایی که درآمد به نسبت بالا است؛ هزینه‌ها هم به نسبت بالا بوده و تفاوت بین هزینه و درآمد در شهرستان‌ها تقریباً مشابه است. همچنین بین تعداد پزشکان به تعداد نقاط جمعیتی و ابتلا به بیماری سل رابطه معنی‌دار نبود. ممکن است برای بعضی از شهرستان‌ها تعداد نقاط جمعیتی مثل روستاها بیشتر بوده و این شاخص کمتر شده باشد. از طرفی ممکن است با افزایش تعداد پزشک، امکان تشخیص بهتر بیماران مبتلا به سل فراهم آمده و تعداد افراد مبتلا افزایش نیافته است. در مطالعه Vendramini و همکاران یکی از عوامل کاهش ابتلا به بیماری سل دسترسی به پزشک معرفی شد (۵).

نتیجه‌گیری

نتایج این مطالعه نشان داد که شاخص‌های اجتماعی - اقتصادی با موارد ابتلا به بیماری سل مرتبط است. همچنین استفاده از مدل رگرسیونی دوجمله‌ای منفی برای داده‌های شمارشی در مقایسه با مدل رگرسیونی پواسون (دارای بیش پراکنش) ارجحیت داشت.

تشکر و قدردانی

این مقاله حاصل پایان‌نامه خانم فاطمه سروی برای اخذ درجه کارشناسی ارشد در رشته آمار زیستی از دانشکده پیراپزشکی دانشگاه علوم پزشکی شهیدبهشتی بود. نویسندگان مقاله از همکاری آقای یوسف بشیری تشکر می‌نمایند.

مناسب‌تر ارزیابی شد ($P < 0/001$). همچنین هر دو شاخص AIC و BIC نشان داد که مدل دوجمله‌ای منفی مناسب‌تر است. هر دو شاخص برای مدل دوجمله‌ای منفی کمتر بودند که نشان داد برآزش این مدل سبب از دست رفتن اطلاعات کمتری می‌شود (جدول ۳).

بحث

نتایج این مطالعه نشان داد که شاخص‌های اجتماعی - اقتصادی با تعداد ابتلای افراد به بیماری سل ارتباط آماری معنی‌داری دارد. در کاربرد مدل رگرسیونی دوجمله‌ای منفی بین شاخص‌های متوسط درآمد سالانه برای هر خانوار شهرستانی، تعداد پزشکان به تعداد نقاط جمعیتی و تراکم خانوار در هر واحد مسکونی رابطه آماری معنی‌داری وجود نداشت. همچنین واریانس ضرایب در مدل رگرسیونی دوجمله‌ای منفی بیشتر بود. دلیل این اختلاف هم می‌تواند به خاطر وجود بیش پراکنش در مدل رگرسیونی پواسون باشد. لذا استفاده از مدل رگرسیونی دوجمله‌ای منفی برای داده‌های شمارشی در مقایسه با مدل رگرسیونی پواسون (دارای بیش پراکنش) ارجحیت دارد. بیش پراکنش ممکن است منجر به معنی‌دار شدن ارتباط‌هایی شود که در واقع معنی‌دار نیستند. این یافته با سایر مطالعات همراستا است (۱۶ و ۱۵).

در مطالعه Lambert مدل پواسون و دوجمله‌ای منفی درباره تعداد نقص در خط تولید یک کارخانه مقایسه شد. تعداد خط‌های بدون نقص در مدل پواسون ۷۱ درصد و در مدل دوجمله‌ای منفی ۷۸ درصد پیش‌بینی شد که با توجه به مقدار واقعی یعنی ۸۱ درصد، مدل دوجمله‌ای منفی پیش‌بینی بهتری نشان داد (۱۵). در مطالعه Greene برای داده‌های مربوط به تعداد تخلفات در پرداخت با کارت‌های اعتباری در شصت روز گذشته از مدل‌های پواسون، دوجمله‌ای منفی و ZIP استفاده شد و مدل دوجمله‌ای منفی به دلیل وجود بیش پراکنش در داده‌ها، برآزش بهتری نشان داد (۱۶).

در مطالعه حاضر نتایج مدل رگرسیونی دوجمله‌ای منفی نشان داد که به ازای یک واحد افزایش در میزان مهاجرت، میزان ابتلا به سل به میزان ۰/۰۲۶ افزایش می‌یابد. مهاجرت افراد معمولاً به مناطق شهری بیشتر است و در این مناطق تراکم جمعیتی زیاد بوده و افراد مهاجر در ابتدای امر، معمولاً در مناطق حاشیه‌ای که از لحاظ امکانات بهداشتی نامناسب است؛ ساکن می‌شوند. همچنین مهاجرت بیشتر از کشورهای افغانستان و پاکستان با میزان بروز بالای سل (۱۷) به ایران انجام می‌شود. در مطالعات Gilbert و همکاران (۱۰)، Talbot و همکاران (۱۳) و Coker و همکاران (۱۸) نیز افزایش میزان مهاجرت سبب افزایش در میزان ابتلا به بیماری سل شده بود. در مطالعه حاضر به ازای یک واحد افزایش در ضریب شهرنشینی، میزان ابتلا به بیماری سل به میزان ۰/۹ افزایش یافت. با

References

- Nasehi M, Mirhaghaani L. [Guideline for control TB]. 2nd. Tehran: Andishmand Publication. 2009; pp: 5-17. [Persian]
- World Health Organization. Addressing Poverty in TB Control. Options for National TB Control Programmes (WHO/HTM/TB/2005.352)
- Arsang Sh, Kazemnejad A, Amani F. [Epidemiology of tuberculosis in Iran (2001-08)]. J Gorgan Uni Med Sci. 2001; 13(3): 78-86. [Article in Persian]
- Figuroa-Munoz JI, Ramon-Pardo P. Tuberculosis control in vulnerable groups. Bull World Health Organ. Sep 2008; 86(9): 733-5.
- Vendramini SH, Santos ML, Gazetta CE, Chiaravalloti-Neto F, Ruffino-Netto A, Villa TC. Tuberculosis risks and socio-economic level: a case study of a city in the Brazilian south-east, 1998-2004. Int J Tuberc Lung Dis. 2006 Nov;10(11):1231-5.
- Hawker JI, Bakhshi SS, Ali S, Farrington CP. Ecological analysis of ethnic differences in relation between tuberculosis and poverty. BMJ. 1999 Oct;319(7216):1031-4.
- Sánchez-Pérez H, Flores-Hernández J, Jansá J, Caylá J, Martín-Mateo M. Pulmonary tuberculosis and associated factors in areas of high levels of poverty in Chiapas, Mexico. Int J Epidemiol. 2001 Apr;30(2):386-93.
- Bhunu CP, Mushayabasa S, Smith RJ. Assessing the effects of poverty in tuberculosis transmission dynamics. Appl Math Model. 2012; 36(9):4173-85.
- Kwan CK, Ernst JD. HIV and tuberculosis: a deadly human syndemic. Clin Microbiol Rev. 2011 Apr;24(2):351-76.
- Gilbert RL, Antoine D, French CE, Abubakar I, Watson JM, Jones JA. The impact of immigration on tuberculosis rates in the United Kingdom compared with other European countries. Int J Tuberc Lung Dis. 2009 May;13(5):645-51.
- McBryde ES, Denholm JT. Risk of active tuberculosis in immigrants: effects of age, region of origin and time since arrival in a low-exposure setting. Med J Aust. 2012 Oct;197(8):458-61.
- Hilbe JM. Negative Binomial Regression. 1st. Cambridge: Cambridge University Press. 2007; pp: 142-6.
- Talbot EA, Moore M, McCray E, Binkin NJ. Tuberculosis among foreign-born persons in the United States, 1993-1998. JAMA. 2000 Dec;284(22):2894-900.
- Kutner M, Nachtsheim C, Neter J, Li W. Applied Linear Statistical Models. 5th. New York: McGraw-Hill/Irwin. 2004; pp: 359-60.
- Lambert D. Zero-inflated poisson regression, with an application to defects in manufacturing. Technometrics. 1992; 34(1):1-14.
- Greene WH. Accounting for excess zeros and sample selection in poisson and negative binomial regression models. New York: New York University Stern School of Business. 1994; no EC-94-10. Available at SSRN:<http://ssrn.com/abstract=1293115>
- Rezaeian M, Dunn G, St Leger S, Appleby L. Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. J Epidemiol Community Health. 2007 Feb;61(2):98-102.
- Coker RJ, McKee M, Atun R, Rechel B, Dodonova E, Kuznetsov S, et al. Risk factors for pulmonary tuberculosis in Russia: a case-control study. British Medical Journal (BMJ). 2006; 332(7533): 85-7.

Original Paper

Relationship between socio-economic factors and tuberculosis using negative binomial and Poisson regression models

Sarvi F (M.Sc)¹, Mehrabi Y (Ph.D)*², Abadi AR (Ph.D)³
Nasehi M (Ph.D)⁴, Payandeh A (Ph.D)⁵

¹M.Sc in Biostatistics, Department of Biostatistics and Epidemiology, Paramedical Sciences Faculty, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ²Professor, Department of Epidemiology, School of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ³Associate Professor, Department of Health and Community Medicine, School of Medicine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ⁴Assistant Professor, Department of Epidemiology, School of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ⁵Ph.D Candidate in Biostatistics, Department of Biostatistics and Epidemiology, Paramedical Sciences Faculty, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Abstract

Background and Objective: Tuberculosis (TB) is the most important cause of death worldwide. The main reason for the increasing global burden of TB are severe poverty and class distinctions between rich and poor population groups in various communities. This study was performed to determine the relationship between socio-economic factors and TB using negative binomial and Poisson regression models.

Method: This descriptive - analytic study was conducted on 11320 TB affected patients in Iran during 2010. Data was gathered from the Iranian Ministry of Health and Medical Education. The relationship between the numbers of cases with socio-economic indicators was determined using negative binomial and Poisson models. Fitting models were compared using AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion).

Results: The Poisson regression model showed a significant relationship between the TB mortality rate and socio-economic factors ($P<0.05$). Negative binomial regression model showed a significant relationship between TB and unemployment, illiterate, immigration and urban residency ($P<0.05$). Negative binomial regression model showed no relationship between TB and family size, physicians' ratio to the number of population centers and annual average income.

Conclusion: There is a significant impact of socio-economic factors with the number of TB cases. Negative binomial regression model is suitable for accountable data in comparison with Poisson regression model.

Keywords: Tuberculosis, Socio-economic indicators, Negative binomial regression model, Poisson regression model

* Corresponding Author: Mehrabi Y (Ph.D), E-mail: mehrabi@sbmu.ac.ir

Received 2 Dec 2013

Revised 20 Apr 2014

Accepted 20 Apr 2014