

# بررسی میزان کامل شماری ثبت سرطان مری در داده‌های ثبت سرطان مبتنی بر جمعیت در استان اردبیل

محمود خدادوست<sup>۱</sup>، پروین یآوری<sup>۲</sup>، مسعود بابایی<sup>۳</sup>، فاطمه سروی<sup>۴</sup>، سید سعید هاشمی نظری<sup>۵</sup>

<sup>۱</sup> دانشجوی دکترای تخصصی اپیدمیولوژی، مرکز تحقیقات کبد و گوارش، دانشگاه علوم پزشکی بقیه الله، تهران، ایران

<sup>۲</sup> استاد اپیدمیولوژی، گروه بهداشت و پزشکی پیشگیری، دانشکده پزشکی، دانشگاه علوم پزشکی شهید بهشتی، مرکز تحقیقات سرطان دانشگاه علوم پزشکی شهید بهشتی

<sup>۳</sup> انجمن علمی اپیدمیولوژیست‌های ایران

<sup>۴</sup> پزشک عمومی، معاونت سلامت، دانشگاه علوم پزشکی اردبیل

<sup>۵</sup> مرکز تحقیقات سرطان آلمان (DKFZ)

<sup>۶</sup> دانشجوی دکترای تخصصی آمار زیستی، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی همدان، همدان، ایران

<sup>۷</sup> مرکز تحقیقات پیشگیری از مصلوبیت، ها و ارتقای ایمنی، دانشگاه علوم پزشکی شهید بهشتی

<sup>۸</sup> استادیار اپیدمیولوژی، گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شهید بهشتی

نویسنده رابط: پروین یآوری، نشانی: تهران، اوین، گروه بهداشت و پزشکی پیشگیری، دانشکده پزشکی، دانشگاه علوم پزشکی شهید بهشتی، تلفن: ۲۲۴۳۹۹۳۶

پست الکترونیک: p.yavari-grc@sbmu.ac.ir

تاریخ دریافت: ۱۳۹۳/۰۸/۲۵؛ پذیرش: ۱۳۹۴/۰۳/۰۲

**مقدمه و اهداف:** سیاستگذاران حوزه‌ی بهداشت لازم است بدانند داده‌هایی که مبنای تصمیم‌گیری آنان را تشکیل می‌دهد تا چه اندازه قابل اعتماد است. میزان کامل شماری ثبت به‌عنوان معیاری از کیفیت ثبت سرطان مورد استفاده قرار می‌گیرد که عبارت است از مقداری از موارد بروز سرطان قابل گزارش در جمعیت مورد بررسی که واقعاً در سامانه ثبت شده‌اند.

**روش کار:** پس از حذف موارد تکراری، تعداد کل ۴۷۱ مورد جدید سرطان مری که توسط داده‌های پاتولوژی، مستندات بیمارستانی و گواهی‌های مرگ در طی سال‌های ۱۳۸۵ و ۱۳۸۷ به مرکز ثبت سرطان اردبیل گزارش شدند؛ وارد مطالعه گردید. میزان بروز با استفاده از روش صید-بازصید و با استفاده از مدل‌های لگاریتم خطی برآورد شد. به‌منظور انتخاب مدلی که بیش‌ترین برازش را دارد؛ از معیارهای G2.BIC و معیارهای اطلاعات آکائیک و بی‌زین استفاده شد.

**یافته‌ها:** در این مطالعه، مدلی دارای ارتباط بین دو منبع پاتولوژی و مستندات بیمارستانی و هم‌چنین منبع گواهی‌های مرگ به‌طور مستقل از این دو منبع به‌عنوان مدلی که بهترین برازش را دارد؛ انتخاب شد. برآورد میزان کلی کامل بودن ثبت برای سرطان مری برای سال‌های ۱۳۸۵ و ۱۳۸۷ برابر با ۳۶ درصد بود. منبع پاتولوژی بیش‌ترین میزان کامل‌شماری ثبت را برای سرطان مری با کامل‌شماری ۲۱/۱۷ درصد داشت. میزان برآورد شده بروز محاسبه شده توسط روش لگاریتم خطی به ترتیب ۴۹/۷۱ و ۵۳/۸۷ در صدهزار نفر محاسبه شد.

**نتیجه‌گیری:** میزان پایین کامل‌شماری ثبت در اردبیل نیاز به اعمال تغییراتی در بیماریابی و خلاصه‌سازی مشخصاتی مانند استفاده از کد ملی و پرونده الکترونیک سلامت افراد به منظور ثبت دقیق‌تر سرطان است.

**واژگان کلیدی:** صید-بازصید، برآورد بروز، لگ خطی، سرطان مری، کامل‌شماری

## مقدمه

دهه آینده به دو برابر افزایش یابد. بنابراین نظام مراقبت موارد سرطانی به عنوان یکی از اولویت‌های اصلی سامانه بهداشتی-درمانی مطرح است (۱). فعالیت ثبت سرطان به عنوان جزء اساسی برنامه‌های کنترل سرطان مطرح است؛ چراکه داده‌های حاصل از آن در حوزه‌های گوناگون از جمله بررسی‌های سبب‌شناسی، برنامه‌ریزی‌های بهداشتی برای پیشگیری در سطح‌های اول و دوم و درمان‌های مراقبتی بیماران دارای کاربرد

در بین بیماری‌های غیر واگیر، سرطان‌ها بار بیماری عمده‌ای را به جامعه تحمیل می‌کنند. در حال حاضر سرطان عامل ۱۲ درصد مرگ‌ومیر در جهان است، و در سال‌های آینده روند رو به رشدی خواهد داشت. در ایران، سرطان‌ها سومین عامل مرگ‌ومیر هستند و سالانه بیش از ۳۰ هزار نفر در ایران جان خود را در اثر ابتلا به این بیماری‌ها از دست می‌دهند. از طرفی با افزایش امید به زندگی و افزایش درصد سالمندی انتظار می‌رود میزان بروز سرطان در

بیماری‌ها مورد استفاده قرار گیرد، به طوری که تعداد کل بیمارانی که توسط منابع مختلف به صورت ناقص ثبت می‌شوند؛ با استفاده از میزان موارد مشترک ثبت شده بین منابع مختلف برآورد می‌شود. امروزه روش صید- بازصید در اپیدمیولوژی به عنوان یک ابزار برای برآورد کامل بودن ثبت بیماری‌ها عمومیت یافته است (۱۰). کاربرد مهم دیگر این روش در اپیدمیولوژی به منظور برآورد شیوع یک بیماری خاص و همچنین برآورد میزان کامل‌شماری ثبت بیماری است، اما روش صید- بازصید می‌تواند به ویژه در شرایطی ۲ و یا بیش از ۲ فهرست ناکامل از بیماران وجود داشته باشد و مورد استفاده قرار گیرد (۱۱). در هنگام استفاده از روش صید- بازصید ساده دو فرض باید مورد توجه قرار گیرند: اول این که منابع مستقل از یکدیگر باشند، دوم این که تمامی افراد درون یک منبع مشابه، شانس برابری برای حضور داشته باشند (۱۲، ۱۳). استفاده از روش صید- بازصید برای کاهش هزینه‌های ثبت بیماری بسیار کارآیی دارد و همچنین در کاهش سوگرایی در برآورد میزان بروز و در موارد مقایسه زیرگروه‌های جمعیت کاربرد دارد (۱۴). ثبت سرطان مبتنی بر داده‌های پاتولوژی از سال ۱۳۷۸ و همچنین ثبت سرطان مبتنی بر جمعیت در سال ۱۳۸۲ برای نخستین بار در ایران در اردبیل راه‌اندازی شد. فعالیت‌های مرکز ثبت سرطان اردبیل شامل جمع‌آوری داده‌های رخداد سرطان از منابع پاتولوژی، بیمارستانی و گواهی‌های مرگ است. هدف اصلی مرکز ثبت سرطان اردبیل، اندازه‌گیری میزان بروز و مرگ و میر سرطان در ساکنان استان اردبیل است (۱۵). بسیاری از ناحیه‌های شمالی ایران، بیش‌ترین میزان بروز سرطان مری را دارند و استان گلستان واقع در شمال شرقی و استان اردبیل در شمال غربی ایران از استان‌های با میزان بروز بالای سرطان مری محسوب می‌شوند، و با پیشروی از شمال به سمت جنوب کشور، از میزان بروز سرطان مری کاسته می‌شود (۱۴، ۱۶). این مطالعه با هدف برآورد میزان بروز سرطان مری در استان اردبیل با مقایسه ۳ منبع ثبت داده‌ها (شامل داده‌های پاتولوژی، گواهی مرگ و مستندات بیمارستانی) و استفاده از روش صید- بازصید انجام شده است.

## روش کار

این مطالعه، یک مطالعه توصیفی با استفاده از روش صید- بازصید برای برآورد میزان کامل بودن ثبت سرطان است. در این مطالعه تمامی موارد سرطان مری که توسط مرکز ثبت سرطان مبتنی بر جمعیت استان اردبیل در سال‌های ۱۳۸۵ و ۱۳۸۷ ثبت

فراگیر می‌باشد (۲). آگاهی از اطلاعات میزان دقیق بروز سرطان برای برنامه‌ریزی، پایش و ارزشیابی برنامه‌های ملی و منطقه‌ای کنترل سرطان ضروری است (۳). یکی از اهداف ثبت سرطان مبتنی بر جمعیت، برآورد بار ناشی از سرطان در جمعیت تحت پوشش به منظور مشاهده روند و تفاوت‌های منطقه‌ای بیماری و همچنین تهیه یک بانک اطلاعاتی به منظور پژوهش‌های اپیدمیولوژیک است (۴). تصمیم‌گیران حوزه بهداشت نیاز دارند بدانند که چگونه به اطلاعاتی که مبنای سیاست‌ها و برنامه‌ریزی‌های آنان را تشکیل می‌دهد؛ اعتماد کنند (۴، ۵). دو فاکتور مهم تعیین‌کننده کیفیت داده‌ها در سامانه ثبت، کامل بودن ثبت موارد-یا به نوعی تعیین میزان کم‌شماری- و درستی جزئیات- عدم وجود بیش‌شماری و دوباره‌نویسی) می‌باشد (۲، ۶). میزان کامل بودن<sup>۱</sup> سامانه ثبت شامل نسبتی از موارد بروز یافته در جمعیت می‌باشد که در سامانه ثبت به ثبت رسیده‌اند. کامل بودن ثبت تا جایی که امکان دارد باید نزدیک به ۱۰۰ درصد باشد. مقایسه میزان‌های بروز بین سامانه‌های ثبت، تفاوت‌های واقعی خطر ابتلا به سرطان را نشان می‌دهد (۲، ۷). در متون علمی، چندین روش مختلف به منظور ارزیابی میزان کامل‌شماری بیان شده است که به دو گروه کلی روش‌های کمی و کیفی تقسیم می‌شوند. روش‌های کیفی مورد استفاده شامل نسبت مرگ‌ومیر به بروز و نسبت موارد تأیید میکروسکوپی، و از بین روش‌های کمی، آن‌هایی که مورد استفاده قرار گرفته‌اند؛ شامل روش صید- بازصید<sup>۲</sup>، گواهی‌های مرگ و نسبت میزان مرگ‌ومیر به میزان بروز، و روش جریان‌ی<sup>۳</sup> است (۸). از آنجایی که بیش‌تر سامانه‌های ثبت سرطان از بیش از یک منبع اطلاعات برای تشخیص موارد بیماری استفاده می‌کنند، روش‌های صید-بازصید ممکن است برای برآورد تعداد موارد رخداد در جمعیت مورد استفاده قرار گیرند. از این رو، به منظور بررسی میزان کامل بودن ثبت موارد بیماری مورد استفاده قرار می‌گیرد (۹). روش صید- بازصید به طور عمده در علوم حیات وحش برای برآورد کل حیواناتی که آزادانه در طبیعت زندگی می‌کنند؛ استفاده می‌شود. در این روش، نمونه‌های پی‌درپی از حیوانات شکار شده و علامت‌گذاری می‌شوند و تعداد کل جمعیت این حیوانات از روی تعداد موارد مشترک بین نمونه‌های صید- بازصید برآورد می‌شوند. یک رویکرد مشابه ممکن است برای برآورد تعداد کل بیماران در پایش اپیدمیولوژیک

<sup>۱</sup> Completeness

<sup>۲</sup> Capture recapture

<sup>۳</sup> Flow method

لگاریتم خطی، سوگرایی ناشی از برقرار نبودن این پیش‌فرض کنترل شد. بر حسب نوع و تعداد اثرهای متقابل، ۸ مدل مختلف لگاریتم خطی به وجود آمد:

یک مدل مستقل؛ ۳ مدل دارای یک اثر متقابل؛ ۳ مدل دارای دو اثر متقابل و یک مدل دارای ۳ اثر متقابل (مدل اشباع).

برای ۳ منبع ثبت، A با i سطح، B با j سطح، C با k سطح، لگاریتم طبیعی (Ln یا log e) فراوانی مورد انتظار  $F_{ijk}$  برای خانه  $i, j, k$ ، به صورت زیر نمایان شد:

$$\ln F_{ijk} = \theta + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

که در آن  $\theta$  پارامتر مشترک،  $\lambda_i^A, \lambda_j^B, \lambda_k^C$  پارامترهای اثر اصلی،  $\lambda_{ij}^{AB}, \lambda_{ik}^{AC}, \lambda_{jk}^{BC}$  پارامترهای برهم‌کنش‌های ۲ طرفه و  $\lambda_{ijk}^{ABC}$  پارامتر برهم‌کنش ۳ طرفه (بالاترین اثر مورد نظر) می‌باشد. مقدار برهم‌کنش سه طرفه را نمی‌توان با استفاده از داده‌های مطالعه آزمون کرد و معمولاً مقدار آن صفر فرض می‌شود، اما مقدارهای سایر برهم‌کنش‌ها قابل آزمون هستند. در این مطالعه، ۳ نوع مدل لگ خطی وجود دارد، یک «مدل مستقل» که فرض بر این است که هر یک از سه منبع ثبت سرطان نسبت به هم مستقل هستند. دوم مدلی است که در آن فقط ۲ منبع ثبت مستقل هستند و در نهایت نوع سوم «مدل اشباع»<sup>۲</sup> نامیده می‌شود که در آن تمامی برهم‌کنش‌های ممکن بین ۳ منبع ثبت را شامل می‌شود.

برای ارزیابی چگونگی برازش مدل‌های لگ خطی با داده‌ها (برازش مدل)<sup>۳</sup> و انتخاب بهترین مدل، از لگاریتم نسبت درست‌نمایی (که به‌عنوان آماره‌ی  $G^2$  نیز شناخته می‌شود)، معیار اطلاعات آکائیک (AIC)<sup>۴</sup> و معیار اطلاعات بی‌زیب (BIC)<sup>۵</sup> استفاده شد که چگونگی محاسبه‌ی آن‌ها به‌صورت زیر است:

$$G^2 = -2 \sum \text{Obs}_j \ln[\text{Obs}_j / \text{Exp}_j]$$

که در آن  $\text{Obs}_j$  تعداد افراد مشاهده شده در هر خانه  $j$ ، و  $\text{Exp}_j$  تعداد موارد مورد انتظار افراد در هر خانه  $j$  تحت مدل  $i$  است که هر چقدر مقدار آماره‌ی  $G^2$  کم‌تر باشد، مدل برازش شده مناسب‌تر خواهد بود.

$$AIC = G^2 - 2[df]$$

شده بودند؛ وارد مطالعه شدند. از این جمعیت هیچ‌گونه نمونه‌گیری انجام نشد و تمامی موارد گزارش شده سرطان مری از منابع مختلف مستندات بیمارستانی، گزارش‌های پاتولوژی و گواهی‌های مرگ در این استان وارد مطالعه شدند. از آنجایی که داده‌های ثبت شده در سال ۱۳۸۶ ناقص بوده و از کیفیت مناسب برخوردار نبودند، از مطالعه خارج شدند. ابتدا موارد تکراری در هر منبع با استفاده از نرم‌افزار اکسل نسخه ۲۰۰۷ شناسایی و حذف شدند. سپس موارد مشترک بین ۳ منبع با استفاده از مشخصات بیماران مانند نام، نام‌خانوادگی، نام پدر و کد ICD مربوط به نوع سرطان آن‌ها و در صورت نیاز، تاریخ تولد و محل تولد و سکونت آن‌ها در نرم‌افزار اکسل شناسایی شدند. موارد مشترک بین ۳ منبع با مقایسه دوه‌دوی آن‌ها و پس از ردیف کردن اسامی به ترتیب‌های مختلف (به ترتیب نام‌خانوادگی، نام، نام پدر و شهر محل سکونت) با دستور Sort و مقایسه هر مورد با افراد ثبت شده در منبع دیگر شناسایی شدند. میزان بروز سرطان مری توسط روش صید- بازصید و مدل لگاریتم خطی برآورد شد. در این روش، دو فرض اصلی باید برقرار باشد. نخست این‌که منابع اطلاعات باید مستقل باشند و دوم این‌که همه افرادی که در هر منبع اطلاعاتی وجود دارند؛ باید شانس برابری برای حضور در مطالعه داشته باشند (۱۲، ۱۳)، اما در بیشتر مطالعات جوامع انسانی و حوزه علوم پزشکی معمولاً این فرض‌ها به‌طور کامل برقرار نیستند و منابع مختلف مستقل از هم نیستند. تجزیه و تحلیل به‌صورت ۳ منبعی و با استفاده از مدل لگ خطی<sup>۱</sup> با نرم‌افزار Stata 12 نسخه ۱۲ (StataCorp, Texas, USA) انجام شد. با ۳ منبع ثبت، ۸ ترکیب احتمالی از این منابع وجود داشت که در آن احتمال حضور یا عدم حضور موارد در این منابع وجود دارد. به‌منظور تجزیه و تحلیل ابتدا یک جدول توافقی با ۸ خانه ( $2^3 = 8$ ) که ۷ خانه‌ی آن مشخص بودند، تشکیل شد (نمودار شماره ۱). به‌منظور برآورد فراوانی خانه هشتم در جدول توافقی- افرادی که در هیچ‌کدام از ۳ منبع ثبت نشده‌اند-، ۸ مدل لگ خطی مختلف روی این هفت خانه جدول برازنده شدند و با استفاده از هر یک از این مدل‌ها، فراوانی منتظره‌ی خانه هشتم برآورد گردید. در این مطالعه، از آنجایی که از ۳ منبع ثبت مورد استفاده قرار گرفت و فرض استقلال منابع به‌طور کامل برقرار نمی‌باشد، با وارد کردن اثر متقابل بین منابع مختلف در مدل

<sup>۲</sup> Saturated model

<sup>۳</sup> Model fitting

<sup>۴</sup> Akaike's Information Criterion; AIC

<sup>۵</sup> Bayesian Information Criterion; BIC

<sup>۱</sup> Loglinear

شامل نام و نام خانوادگی و شماره تلفن و نشانی و سایر مشخصات آن‌ها به صورت محرمانه باقی ماند، و پس از انجام اتصال اطلاعات به منظور یافتن موارد مشترک بین منابع ثبت، تمامی مشخصات آن‌ها حذف شد و از سامانه کدگذاری افراد استفاده شد و تمامی تجزیه و تحلیل‌ها بر اساس کدهای اختصاص یافته به هر فرد صورت گرفت. به منظور انتخاب مدل و برآورد میزان بروز از نرم‌افزار Stata نسخه ۱۲ استفاده شد.

### یافته‌ها

پس از حذف موارد تکراری سرطان مری بین ۳ منبع ثبت سرطان، در مجموع ۴۷۱ مورد جدید سرطان مری در سال‌های ۱۳۸۵ و ۱۳۸۷ به مرکز ثبت سرطان اردبیل گزارش شد که ۲۷۷ مورد توسط منبع پاتولوژی، ۱۹۳ مورد توسط مستندات بیمارستانی و ۱۵۲ مورد توسط گواهی‌های مرگ گزارش شده بودند. از این تعداد ۲۶۶ مورد (۵۶/۵۰ درصد) آن‌ها مرد و ۲۰۵ مورد (۴۳/۵۰ درصد) مورد زن بودند. نسبت مرد به زن برابر ۱/۲۹ بود. میانگین کلی سن افراد  $66/68 \pm 13/23$  سال و هم‌چنین برای مردان و زنان به ترتیب  $67/60 \pm 12/91$  سال و  $65/48 \pm 13/58$  سال بود، و تفاوت آماری معنی‌داری بین سن مردان و زنان مشاهده نشد ( $P=0/08$ ).

نمودار شماره ۱ تعداد موارد مشترک سرطان مری بین ۳ منبع پاتولوژی، مستندات بیمارستانی و ثبت مرگ را نشان می‌دهد، که کل موارد گزارش شده (با یک بار احتساب موارد مشترک) ۴۷۱ نفر می‌باشد. منابع پاتولوژی و مستندات بیمارستانی بیش‌ترین موارد اشتراک را داشتند. جدول شماره ۱ مقایسه برخی ویژگی‌های دموگرافیک بیماران که توسط سه منبع ثبت سرطان گزارش شده‌اند (پس از حذف موارد تکراری) را طی سال‌های ۱۳۸۵ و ۱۳۸۷ نشان می‌دهد. ملاحظه می‌شود که توزیع سنی و جنسی افراد در سال‌های ۱۳۸۵ و ۱۳۸۷ تفاوت آماری معنی‌داری را نشان نداد. بیش‌ترین موارد بیماری در مردان با ۲۶۶ مورد (۵۶/۵۰ درصد) و گروه سنی بالای ۶۵ سال با ۲۹۳ مورد (۶۲/۲۰ درصد) بود.

نتایج تجزیه و تحلیل ۳ منبعی توسط مدل لگ خطی در جدول شماره ۲ آمده است. از بین ۸ مدل برازش شده بر اساس آماره AIC، مدل دوم شامل یک اثر متقابل بین منابع پاتولوژی و مستندات بیمارستانی و یک اثر مستقل منبع ثبت مرگ که بیش‌ترین برازندگی را داشت؛ انتخاب گردید. در صورتی که بر مبنای آماره‌ی BIC نیز همین مدل دارای بیش‌ترین میزان

اصطلاح اول " $G^2$ " معیاری است برای این‌که چگونه مدل متناسب با داده‌ها برازش شده است و اصطلاح دوم " $2 [df]$ " به منظور پارامترهای اضافه و هم‌چنین پیچیدگی مدل است.

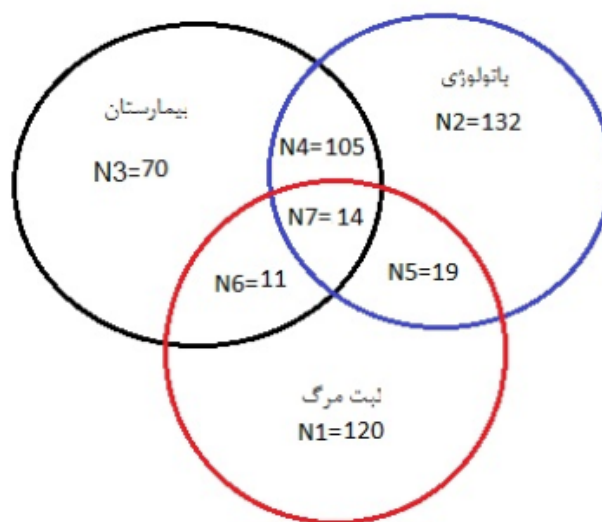
$$BIC = G^2 - [\ln N_{obs}] [df]$$

که در آن  $N_{obs}$  برابر است با تعداد کل افراد مشاهده شده. نسبت به AIC به طور سنگین‌تری تحت تأثیر پیچیدگی مدل‌ها قرار می‌گیرد (۱۷). شایع‌ترین معیار برای انتخاب مدل مناسب که بیش‌تر توسط پژوهشگران مورد استفاده قرار می‌گیرد؛ AIC است (۲۰-۱۸). در این مطالعه، برای انتخاب یکی از ۸ مدل برازش شده از معیارهای AIC، BIC و  $G^2$  استفاده شد و در نهایت بر اساس آماره‌ی AIC که بیش‌ترین کاربرد را توسط پژوهشگران دارد، مدلی که شامل یک اثر متقابل بین منابع پاتولوژی و مستندات بیمارستانی و یک اثر مستقل منبع ثبت مرگ بود (مدل PC/D) انتخاب شد. پس از انتخاب مدلی که بیش‌ترین برازش ممکن را داشت، تجزیه و تحلیل‌ها در زیرگروه‌های سنی، جنسی و هم‌چنین بر اساس سال‌های تقویمی به صورت جداگانه انجام شد. در صید- بازصید ۳ منبعی، اگر یکی از خانه‌های جدول توافقی فراوانی صفر داشته باشد، برخی از مدل‌های لگاریتم خطی منجر به برآورد بی‌نهایت خواهد شد، که این حالت برای برخی از خانه‌های جدول توافقی، این مطالعه فراوانی صفر داشتند. برای اصلاح این مسأله روش افزودن  $\frac{1}{2} (k-1)$  (k: تعداد منابع) به همه خانه‌های جدول توافقی پیشنهاد می‌شود که در این مطالعه از این اصلاح برای خانه‌های جدول توافقی سرطان ریه استفاده شد (۲۲-۲۰). در این مطالعه برای کنترل کردن سوگرایی ناشی از عدم استقلال منابع از روش لگاریتم خطی با استفاده از ۳ منبع استفاده شد. مزیت استفاده از این روش آن است که با وارد کردن اثر متقابل بین منابع می‌توان وابستگی بین منابع مختلف را کنترل نمود. از آنجایی که ممکن است افراد بیمار به علت متفاوت بودن شدت و مرحله بیماری، شانس برابری برای حضور در منبع ثبت مرگ در مقایسه با منابع ثبت پاتولوژی و مستندات بیمارستانی را نداشته باشند، در این مطالعه سرطان مری که میزان کشندگی بالایی دارد؛ انتخاب شد تا شانس حضور افراد در منبع ثبت بیش‌تر باشد، و بتوان منبع ثبت مرگ را نیز مورد استفاده قرار داد. هم‌چنین در این مطالعه، رخداد بیماری بر اساس لایه‌های مختلف متغیرهای تعدیل‌کننده مانند سن و گروه‌های سنی به طور جداگانه محاسبه شد. در تمامی مراحل آماده‌سازی، تجزیه و تحلیل و انتشار نتایج این مطالعه، مشخصات افراد تحت مطالعه

با ۳۴/۳۶ درصد محاسبه گردید. در هر دو زیر گروه سال‌های تقویمی، میزان کامل بودن ثبت در زنان بیش‌تر از مردان بود که در سال‌های ۸۵ و ۸۷ برای مردان به ترتیب برابر با ۳۰/۹۰ و ۳۳/۹۳ درصد و برای زنان به ترتیب برابر با ۴۹/۲۴ و ۳۵/۲۹ درصد بود. بیش‌ترین موارد رخداد برآورد شده در بین گروه‌های سنی مربوط به افراد ۶۵ سال و بالاتر بود که برای سال ۱۳۸۵ برابر با ۴۰.۸ (حدود اطمینان ۹۵ درصد: ۶۳۲/۶۶-۲۸۷/۲۶) با میزان کامل بودن ۳۶/۰۲ درصد و برای سال ۱۳۸۷ برابر با ۳۳۷/۸۱ (حدود اطمینان ۹۵ درصد: ۵۰۷/۸۴-۲۴۷/۶۸) با میزان کامل بودن ۴۳/۲۱ درصد بود. میزان بروز سرطان مری در استان اردبیل بر اساس تعداد موارد برآورد شده بیماری در سال‌های ۱۳۸۵ و ۱۳۸۷ و همچنین بر اساس جمعیت کل ساکن استان اردبیل در این سال‌ها، یک‌بار به‌صورت کلی و بار دیگر در زیرگروه‌های جنسی مختلف به‌صورت مجزا محاسبه شد (جدول شماره ۴). میزان بروز برآورد شده‌ی سرطان مری در سال ۱۳۸۵ برابر با ۴۹/۷۱ در صد هزار نفر (حدود اطمینان ۹۵ درصد: ۶۹/۱۲-۳۷/۷۹) و در سال ۱۳۸۷ برابر با ۵۳/۸۷ در صد هزار نفر (حدود اطمینان ۹۵ درصد: ۷۹/۶۴-۳۸/۹۷) محاسبه شد.

برازندگی مدل بود. از میان معیارهای برازندگی AIC بیش از  $G^2$  و BIC توسط پژوهشگران استفاده و توصیه می‌شود. بنابراین با استفاده از AIC که برآوردهای محافظه‌کارانه‌تری ارائه می‌دهد، تعداد کل موارد بیماری که توسط هیچ‌یک از منابع ثبت نشده‌اند (میزان کم‌شماری) ۸۳۷ نفر برآورد شد. در نتیجه برآورد تعداد کل موارد سرطان مری در سال‌های ۱۳۸۵ و ۱۳۸۷ برابر ۱۳۰۸/۲۷ (حدود اطمینان ۹۵ درصد: ۱۶۷۴/۸۰-۱۰۵۳/۳۴) نفر شد. بر همین اساس، میزان کامل بودن ثبت برای هر سه منبع پس از حذف موارد تکراری برابر است با ۳۶ درصد (۴۷۱ مورد) و برای هر یک از منابع پاتولوژی، مستندات بیمارستانی و گواهی‌های مرگ به ترتیب ۲۱/۱۷ درصد (۲۷۷ مورد)، ۱۴/۷۵ درصد (۱۹۳ مورد) و ۱۱/۶۲ درصد (۱۵۲ مورد) برآورد شد.

میزان کامل بودن ثبت سرطان مری در استان اردبیل بر اساس تعداد موارد برآورد شده بیماری (با استفاده از مدل مناسب انتخاب شده) در سال‌های ۱۳۸۵ و ۱۳۸۷ یک بار به‌صورت کلی و بار دیگر در زیرگروه‌های سنی و جنسی مختلف به‌صورت مجزا محاسبه شد (جدول شماره ۳). میزان کامل بودن ثبت سرطان مری در سال ۱۳۸۵ برابر با ۳۷/۷۶ درصد و در سال ۱۳۸۷ برابر



نمودار شماره ۱- نمودار ون تعداد موارد مشترک سرطان مری در استان اردبیل که در ۳ منبع پاتولوژی، مستندات بیمارستانی و گواهی مرگ ثبت شدند.



جدول شماره ۳- برآورد موارد بروز و میزان کامل‌شماری سرطان مری با استفاده از روش لگاریتم خطی بر مبنای جمعیت استان اردبیل در سال‌های ۱۳۸۵ و ۱۳۸۷

سال	زیرگروه‌ها	* موارد بروز گزارش شده ( $I_1$ )	موارد بروز برآورد شده ( $I_2$ )	95% CI for $I_2$	**میزان کامل بودن (درصد)	
۱۳۸۵	جنسیت	مرد	۴۴۰	(۲۹۱/۳۴-۷۳۰/۸۹)	۳۰/۹۰	
		زن	۱۰۰	(۱۴۹/۲۶-۳۱۵/۶۵)	۴۹/۲۴	
	گروه‌های سنی	≤۵۰	۲۵	۸۸	(۳۷/۴۷-۳۴۳/۱۱)	۲۸/۴۰
		۵۱-۶۴	۶۵	۱۳۰/۳۳	(۹۱/۸۲-۲۲۴/۱۴)	۴۹/۸۷
		≥۶۵	۱۴۷	۴۰۸	(۲۸۷/۲۶-۶۳۲/۶۶)	۳۶/۰۲
		کل	۲۳۶	۶۲۴/۹۶	(۴۷۵/۰۳-۸۶۸/۹۰)	۳۷/۷۶
۱۳۸۷	جنس	مرد	۳۸۳/۰۹	(۲۵۳/۶۱-۶۴۸/۱۶)	۳۳/۹۳	
		زن	۱۰۵	(۱۸۷/۲۹-۵۵۵/۲۹)	۳۵/۲۹	
	گروه‌های سنی	≤۵۰	۳۳	۱۱۷	(۴۹/۵۹-۴۵۸/۱۴)	۲۸/۲۰
		۵۱-۶۴	۵۵	۶۲۹	(۱۲۸/۷۶-۴۵۲/۱۶۰)	۸/۷۴
		≥۶۵	۱۴۶	۳۳۷/۸۱	(۲۴۷/۶۸-۵۰۷/۸۴)	۴۳/۲۱
		کل	۲۳۵	۶۸۳/۸۴	(۴۹۴/۶۷-۱۰۱۰/۸۲)	۳۴/۳۶

\* بروز گزارش شده توسط سه منبع پاتولوژی، مستندات بیمارستانی و گواهی‌های مرگ پس از حذف موارد تکراری

\*\* تعداد موارد بروز ثبت شده تقسیم بر تعداد موارد بروز برآورد شده ضربدر ۱۰۰

جدول شماره ۴- میزان بروز برآورد شده سرطان مری در استان اردبیل در سال‌های ۱۳۸۵ و ۱۳۸۷

نوع سرطان	سال	جنس	* میزان بروز گزارش شده ( $I_1$ )	میزان بروز برآورد شده ( $I_2$ )	95% CI for $I_2$
۱۳۸۵		مرد	۲۱/۰۸	۶۸/۲۲	(۱۱۳/۳۲-۴۵/۱۷)
		زن	۱۶/۳۳	۳۳/۱۷	(۵۱/۵۷-۲۴/۳۸)
		کل	۱۸/۷۷	۴۹/۷۱	(۶۹/۱۲-۳۷/۷۹)
۱۳۸۷		مرد	۱۹/۹۵	۵۸/۸۱	(۹۹/۵۰-۳۸/۹۳)
		زن	۱۶/۹۹	۴۸/۱۵	(۸۹/۸۸-۳۰/۳۱)
		کل	۱۸/۵۱	۵۳/۸۷	(۷۹/۶۴-۳۸/۹۷)

\* میزان بروز گزارش شده (در صد هزار نفر) توسط سه منبع پاتولوژی، مستندات بیمارستانی و گواهی‌های مرگ پس از حذف موارد تکراری

## بحث

در این مطالعه، میزان بروز سرطان مری توسط روش صید-بازصید و مدل لگاریتم خطی برآورد شد. میانگین  $\pm$  انحراف معیار سنی برای سرطان مری  $۱۳/۲۳ \pm ۶۶/۶۸$  سال (مردان  $۱۲/۹۱ \pm ۶۷/۶۰$  سال و زنان  $۱۳/۵۸ \pm ۶۵/۴۸$  سال) بود و تفاوت آماری معنی‌داری از نظر توزیع سنی بروز بیماری بین مردان و زنان مشاهده نشد ( $P=۰/۰۸$ ) که نسبت به مطالعه‌های مشابه انجام شده در ایران که میانگین سنی افراد برای سرطان مری ۶۵-۵۸ سال گزارش شده، کمی بالاتر است (۲۶-۲۳)، اما مطالعه مشابه در تهران که یک مرکز ارجاعی از سایر استان‌ها در ایران محسوب می‌شود، میانگین سنی برای سرطان مری  $۷۰/۴۰$  (مردان  $۷۰/۵۰$  و زنان  $۷۰/۳۰$ ) سال بوده است که نسبت به این مطالعه، حاضر اندکی بیش‌تر است (۲۷). نسبت بروز مرد به زن  $۱/۲۹$  بود که نشان دهنده‌ی موارد بروز متفاوت در زیرگروه‌های جنسی در اردبیل است؛ در حالی‌که در فنلاند و اتحادیه اروپا این مقدار برای سرطان مری  $۱/۶۶$  بوده است، که با نتایج این مطالعه هم‌خوانی دارد (۲۸). نسبت جنسی در این مطالعه با گزارش کشوری سال‌های ۱۳۸۴ و ۱۳۸۵ (برای سرطان مری  $۱/۱۳$  و  $۱/۲۴$ ) (۲)، و هم‌چنین با نسبت جنسی گزارش شده در استان اردبیل در سال‌های ۸۵-۱۳۸۳ هم‌خوانی دارد (۲۹)، که نشان دهنده عدم تغییر رخداد جنسی سرطان در طی سال‌های ۱۳۸۵ و ۱۳۸۷ در این استان است، اما در مطالعه مشابه انجام شده در تهران در سال‌های ۸۵-۱۳۸۲ نسبت جنسی مرد به زن برای سرطان مری  $۱/۳۶$  گزارش شده است (۱۴،۲۷) که با یافته‌های این مطالعه هم‌خوانی دارد. دو منبع پاتولوژی و مستندات بیمارستانی بیش‌ترین موارد مشترک را با  $۲۳/۳$  درصد (۱۰۵ مورد) داشته‌اند که نشان دهنده‌ی وابستگی بیش‌تر این دو منبع به یکدیگر می‌باشد. هم‌چنین دو منبع ثبت مرگ و مستندات بیمارستانی کم‌ترین موارد اشتراک را با  $۲/۳$  درصد (۱۱ مورد) داشته‌اند. میزان اشتراک منابع در این مطالعه با مطالعه مشابه انجام شده در تهران که بیش‌ترین مقدار اشتراک بین منابع پاتولوژی و مستندات بیمارستانی و برای سرطان مری با  $۱۷/۱$  درصد (۲۴۰ مورد) بوده است؛ کاملاً هم‌خوانی دارد، که نشان دهنده یکسان بودن ثبت سرطان و الگوی این سرطان در استان‌های تهران و اردبیل می‌باشد (۲). میزان کامل‌شماری در سال ۱۳۸۵ با  $۳۷/۷۶$  درصد نسبت به سال ۱۳۸۷ با  $۳۴/۳۶$  درصد و در میان زیر گروه‌های سنی، در سال ۱۳۸۵ گروه سنی ۶۴-۵۱

سال با  $۴۹/۸۷$  درصد و در سال ۱۳۸۷ گروه سنی ۶۵ سال و بالاتر با  $۴۳/۲۱$  درصد بیش‌تر بوده است. در زیر گروه‌های جنسی، جنس زن با میزان کامل‌شماری ثبت برای سال‌های ۱۳۸۵ و ۱۳۸۷ به ترتیب  $۴۹/۲۴$  درصد و  $۳۵/۲۹$  درصد بیش‌ترین میزان کامل‌شماری ثبت را داشته‌اند. برآورد میزان کلی کامل بودن ثبت برابر با  $۳۶$  درصد بود که این مقدار در مقایسه با سایر کشورها بسیار پایین‌تر است. در مطالعه مشابه در ایتالیا، میزان کامل بودن ثبت سرطان به‌طور کلی  $۹۷/۴$  درصد بود (۳۰). هم‌چنین در مطالعه انجام شده در هند نیز میزان کامل بودن ثبت سرطان  $۹۶$  درصد بود، که در مقایسه با مطالعه حاضر بسیار بالاتر است (۳۱). در مطالعه مشابه در کانادا میزان کامل بودن ثبت سرطان‌های دستگاه گوارش  $۹۵/۸۳$  درصد بوده است (۹)، که با نتایج سایر کشورها هم‌خوانی دارد و همگی آن‌ها بسیار بالاتر از نتایج این مطالعه هستند. اگرچه میزان کامل بودن ثبت سرطان در این مطالعه نسبت به سایر کشورها بسیار پایین‌تر است، اما با یافته‌های مطالعه مشابه انجام شده در تهران هم‌خوانی دارد. میزان کامل بودن ثبت سرطان مری در تهران  $۲۷/۳$  درصد بود که نسبت به این مطالعه کم‌تر است، که می‌تواند ناشی از موارد ارجاعی از سایر استان‌ها به استان تهران باشد (۲). یافته‌های این مطالعه به‌عنوان دومین مطالعه صورت گرفته در ایران نشان دهنده آن است که وضع ثبت سرطان در استان‌های مختلف کشور تقریباً مشابه و در مقایسه با سایر کشورها وضع نامطلوبی را دارند. میزان کامل‌شماری ثبت برای منابع مختلف ثبت سرطان نیز به‌طور جداگانه محاسبه شد که از بین منابع مختلف ثبت سرطان، منبع پاتولوژی با  $۲۱/۱۷$  درصد بیش‌ترین میزان و ثبت مرگ با  $۱۱/۶۱$  درصد کم‌ترین میزان کامل‌شماری ثبت را داشته‌اند که در مطالعه انجام شده در تهران نیز منبع پاتولوژی بیش‌ترین میزان کامل‌شماری ثبت را برای این سرطان داشته است، که بانتهای این مطالعه هم‌خوانی دارد و مستندات بیمارستانی کم‌ترین میزان کامل‌شماری ثبت را داشته است (۲).

میزان بروز گزارش شده برای سرطان مری بر اساس ۳ منبع پاتولوژی، مستندات بیمارستانی و ثبت مرگ و پس از حذف موارد تکراری برای سال‌های ۱۳۸۵ و ۱۳۸۷ به ترتیب برابر با  $۱۸/۷۷$  و  $۱۸/۵۱$  در صد هزار نفر بود. با استفاده از مدل لگاریتم خطی، میزان بروز سرطان مری برای این سال‌ها به ترتیب  $۴۹/۷۱$  و  $۵۳/۸۷$  در صد هزار نفر برآورد شد که این اختلاف به دلیل موارد کم‌شماری ثبت سرطان مری در این استان است که توسط هیچ‌یک از منابع ثبت نشدند. میزان بروز استاندارد شده سرطان



شده در این مطالعه با روش صید-بازصید و بر اساس ۳ منبع ثبت سرطان صورت گرفته است، که به نظر می‌رسد به واقعیت نزدیک‌تر باشد و در مقایسه با میزان بروز گزارش شده از ثبت سرطان مبتنی بر جمعیت استان اردبیل دقیق‌تر است، و وضع سرطان مری در استان اردبیل را بهتر نشان می‌دهد. باید توجه داشت که در این مطالعه با توجه به این که ممکن است تعدادی از موارد مرگ ثبت شده که در منبع پاتولوژی همان سال حضور ندارند، توسط پاتولوژی سال‌های قبل ثبت شده باشند، و همچنین تعدادی از افرادی که در منبع پاتولوژی حضور دارند و در ثبت مرگ همان سال حضور ندارند، ممکن است توسط منبع ثبت مرگ سال‌های بعد ثبت شده باشند، و با توجه به محدودیت مطالعه در دسترسی به داده‌های سال‌های پیش و پس از مطالعه، ممکن است موارد بروز برآورد شده دارای مقداری بیش‌برآورد<sup>۲</sup> از مقدار واقعی باشد (به‌خاطر زیاد شدن تعداد افراد مربوط به خانه‌ای از جدول که در ثبت مرگ هستند، اما در پاتولوژی و مستندات بیمارستانی ثبت نشده‌اند). علاوه بر آن، یکی از عوامل مهمی که در این مطالعه باعث میزان کم‌شماری ثبت شده است، مربوط به منبع ثبت مرگ و اشکالاتی مانند ثبت شدن علت اصلی مرگ و عدم توجه به علل زمینه‌ای منجر به مرگ که ممکن است سرطان یکی از این عوامل زمینه‌ای باشد، همچنین موارد مرگ ناشی از متاستاز سرطان به سایر اعضای بدن که ممکن است باعث ثبت علت مرگ از سرطان دیگری غیر از سرطان اولیه ثبت شده در پاتولوژی و مستندات بیمارستانی می‌شود. از طرفی ممکن است افرادی که مراحل تشخیصی را در استان اردبیل انجام داده‌اند و در منابع پاتولوژی و مستندات بیمارستانی استان اردبیل ثبت شده‌اند، به‌منظور دریافت خدمات درمانی تخصصی به استان‌های مجاور جمله تهران عزیمت کرده باشند و در نهایت در استان‌های مجاور فوت کرده و ثبت شده باشند و در نتیجه در این مطالعه به‌عنوان کم‌شماری ثبت محسوب شده باشند. یکی از محدودیت‌های این مطالعه، عدم دسترسی به داده‌های ثبت سایر استان‌های مجاور به‌منظور شناسایی موارد ثبت شده در استان‌های مجاور بود، اما حتی با در نظر گرفتن محدودیت‌های بیان شده و احتمال مقداری بیش‌برآورد موارد بروز سرطان از مقدار واقعی که ممکن است باعث کاهش میزان کامل‌شماری برآورد شده باشد، باز هم این مقدار اختلاف بسیار زیادی با ثبت سرطان سایر کشورها دارد و با

مری در جهان توسط IARC<sup>۱</sup> ۵/۹ در صد هزار نفر برآورد شده است (۳۲)، که میزان بروز گزارش شده و هم‌چنین میزان بروز برآورد شده در این مطالعه، بسیار بالاتر از آمارهای جهانی است. میزان بروز برآورد شده سرطان مری در سال ۱۳۸۵ برای مردان و زنان به ترتیب ۶۸/۲۲ و ۳۳/۱۷ در صد هزار نفر و در سال ۱۳۸۷ به ترتیب ۵۸/۸۱ و ۴۸/۱۵ در صد هزار نفر بود که این میزان در کشورهای توسعه یافته برای مردان و زنان به ترتیب ۶/۵ و ۱/۲ در صد هزار نفر و در کشورهای توسعه نیافته برای مردان و زنان به ترتیب ۱۱/۸ و ۵/۷ در صد هزار نفر بوده است (۳۳) و نتایج این مطالعه در هر دو جنس بیش‌تر از میانگین جهانی بوده است. هم‌چنین میزان بروز گزارش شده سرطان مری در این مطالعه نسبت به میزان بروز خام گزارش شده در ژاپن که برای مردان و زنان به ترتیب ۲۲ و ۳/۹ در صد هزار نفر بوده، بیش‌تر است (۳۴). بر اساس گزارش GLOBOCAN در سال ۲۰۱۲ میلادی، میزان بروز سرطان مری در ایران ۸/۶ در صد هزار نفر برآورد شده است که خیلی کم‌تر از میزان برآورد شده در این مطالعه است (۳۵)، اما باید توجه داشت که این میزان گزارش شده مربوط به کل جمعیت ایران است، در صورتی که بسیاری از نواحی شمالی ایران بیش‌ترین میزان بروز سرطان مری را دارند و استان گلستان واقع در شمال‌شرقی و استان اردبیل واقع در شمال‌غربی ایران از استان‌های با میزان بروز بالای سرطان مری محسوب می‌شوند و با پیشروی از شمال به سمت جنوب کشور، از میزان بروز سرطان مری کاسته می‌شود (۱۶، ۱۴). بابایی و همکاران میزان بروز خام و میزان بروز استاندارد شده سنی را برای سرطان مری در استان اردبیل در طی سال‌های ۸۵-۱۳۸۳ به ترتیب ۱۴/۲ و ۱۹/۵ در صد هزار نفر گزارش کردند (۲۹) که به میزان بروز گزارش شده در این مطالعه در سال‌های ۱۳۸۵ و ۱۳۸۷ نزدیک است، اما میزان بروز برآورد شده در این مطالعه با روش لگ خطی بسیار بالاتر بود که این اختلاف به دلیل مقداری کم‌شماری ثبت سرطان در استان اردبیل است که توسط هیچ‌یک از منابع ثبت نشده‌اند. در مطالعه مشابه انجام شده در تهران در سال‌های ۸۵-۱۳۸۱، میزان بروز سرطان مری ۱۰/۵ در صد هزار نفر گزارش شده است (۱۴)، که این اختلاف با توجه به این که استان اردبیل از استان‌های شمال‌غربی کشور و با خطر بالای سرطان مری و تهران از استان‌های با خطر متوسط محسوب می‌شوند؛ توجیه‌پذیر است. موارد بروز برآورد

<sup>۲</sup> Over estimate

<sup>۱</sup> International Agency for Research on Cancer; IARC

سایر علل زمینه‌ای مرگ از سرطان و هم‌چنین سایر اقداماتی که به افزایش دقت و کامل‌شماری ثبت سرطان می‌انجامد را می‌رساند. نتایج این مطالعه، لزوم استفاده از روش صید- بازصید برای اصلاح کم‌شماری میزان بروز سرطان‌ها و به‌دست آوردن برآورد دقیق‌تری از میزان بروز که به واقعیت نزدیک‌تر باشد را نشان می‌دهد.

### تشکر و قدردانی

این مقاله برگرفته از پایان‌نامه آقای محمود خدادوست برای اخذ درجه‌ی کارشناسی ارشد اپیدمیولوژی از دانشکده بهداشت دانشگاه علوم پزشکی شهید بهشتی می‌باشد. نویسندگان این مقاله مراتب تشکر و قدردانی خود را از ریاست و کارکنان محترم مرکز ثبت سرطان اردبیل که نگارندگان را در انجام این مطالعه یاری نمودند؛ اعلام می‌دارند.

توجه به نزدیک بودن یافته‌های این مطالعه با مطالعه انجام شده در تهران، نشان دهنده دقت و کیفیت پایین ثبت سرطان در استان اردبیل و ایران می‌باشد. بنابراین تدابیر برای افزایش دقت ثبت سرطان در سامانه ثبت سرطان ایران مانند آموزش دادن کارشناسان ثبت مرگ و هم‌چنین ثبت علت اصلی و علت زمینه‌ای مرگ و هم‌چنین ثبت افراد با استفاده از کد ملی و هم‌چنین راه‌اندازی پرونده الکترونیک سلامت و یکسان‌سازی ثبت سرطان کل کشور توصیه می‌شود.

### نتیجه‌گیری

برآورد میزان کامل بودن ثبت سرطان در مرکز ثبت سرطان اردبیل در مقایسه با ثبت سرطان در سایر کشورها بسیار پایین‌تر است (۹۰،۳۰،۳۱). بنابراین لزوم توجه و دقت بیشتر در زمینه اقدامات تشخیصی، ثبت موارد بیماری، بررسی گواهی‌های فوت از

### منابع

- Rafati M, Gotbi M, Ahmadnia M. Principles of Prevention and Surveillance of Disease. Tehran: Sepid Barge Baghe Ketab; 2008; 108-18, 1<sup>st</sup> ed.
- Aghaei A. Estimating completeness of stomach and esophagus cancer's registration, based on Tehran metropolitan population, using capture- recapture method, between 21 march 2002 to 21 march 2007. Tehran: Shahid Beheshti University of Medical Sciences; 2011.
- Kamo K-i, Kaneko S, Satoh K, Yanagihara H, Mizuno S, Sobue T. A mathematical estimation of true cancer incidence using data from population-based cancer registries. Japanese Journal of Clinical Oncology. 2007; 37: 150-5.
- Schmidtman I. Estimating Completeness in Cancer Registries—Comparing Capture Recapture Methods in a Simulation Study. Biometrical Journal. 2008; 50: 1077-92.
- Parkin D, Chen V, Ferlay J, Galceran J, Storm H, Whelan S. Comparability and Quality Control in Cancer Registration: IARC Technical Report No19. International Agency for Research on Cancer, Lyon, France. 1994.
- Parkin DM, Bray F, Ferlay J, Pisani P. Estimating the world cancer burden: Globocan 2000. International Journal of Cancer. 2001; 94: 153-6.
- Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, et al. Cancer incidence in five continents. IARC Scientific Publications. 2008, No.160, Lyon, France.
- Castro CSFCd. Estimating Completeness of Cancer Registration-Na application to gastric cancer in the North Region Registry of Portugal 2011.
- Robles SC, Marrett LD, Aileen Clarke E, Risch HA. An application of capture-recapture methods to the estimation of completeness of cancer registration. Journal of Clinical Epidemiology. 1988; 41: 495-501.
- Brenner H, Stegmaier C, Ziegler H. Estimating completeness of cancer registration in Saarland/Germany with capture-recapture methods. European Journal of Cancer. 1994; 30: 1659-63.
- Poorolajal J, Haghdoost AA, Mahmoodi M, Majdzadeh R, Nasser-Moghaddam S, Fotouhi A. Capture- recapture method for assessing publication bias. Journal of Research in Medical Sciences: the official Journal of Isfahan University of Medical Sciences. 2010; 15: 107.
- Suwanrungruang K, Sriplung H, Attasara P, Temiyasathit S, Buasom R, Waisri N, et al. Quality of Case Ascertainment in Cancer Registries: A Proposal for a Virtual Three-source Capture-recapture Technique. Asian Pacific Journal of Cancer Prevention. 2011; 12: 173-8.
- Parkin DM, Bray F. Evaluation of data quality in the cancer registry: principles and methods Part II. Completeness. European Journal of Cancer. 2009; 45: 756-64.
- Mosavi-Jarrahi A, Ahmadi-Jouibari T, Najafi F, Mehrabi Y, Aghaei A. Estimation of Esophageal Cancer Incidence in Tehran by Log-linear Method using Population-based Cancer Registry Data. Asian Pac J Cancer Prev. 2013; 14: 5367-70.
- Babaei M, Pourfarzi F, Yazdanbod A, Chiniforush MM, Derakhshan MH, Mousavi SM, et al. Gastric cancer in Ardabil, Iran--a review and update on cancer registry data. Asian Pac J Cancer Prev. 2010; 11: 595-9.
- Radmard AR. Five common cancers in Iran. Archives of Iranian medicine. 2010; 13: 143-6.
- van Hest NAH. Capture-recapture methods in surveillance of tuberculosis and other infectious diseases: Erasmus University Rotterdam; 2007.
- Motevalian A, Holakoei naeini K, Mahmoodi M, Majdzadeh R, Akbari M. Estimating deaths due to traffic accidents in Kerman using capture\_recapture method. Journal of Health Faculty and Health Research. 2007; 5: 61-72.
- Hook EB, Regal RR. Validity of methods for model selection, weighting for model uncertainty, and small sample adjustment in capture-recapture estimation. American Journal of Epidemiology. 1997; 145: 1138-44.
- Hook EB, Regal RR. Accuracy of alternative approaches to capture-recapture estimates of disease frequency: internal validity analysis of data from five sources. American Journal of Epidemiology. 2000; 152: 771-9.
- Motevalian S, Mahmoodi M, Majdzadeh R, Akbari M. Estimation of death due to road traffic injuries in Kerman district: application of capture-recapture method. Journal of School of Public Health and Institute of Public Health Research. 2007; 5: 61-72.

22. Evans MA, Bonett DG. Bias reduction for multiple-recapture estimators of closed population size. *Biometrics*. 1994;388-95.
23. Rajaiefard A, Moghimi B, Tabatabaie SH, Safaie A, Tabeie SZ. Epidemiological and clinical features of gastric cancer: descriptive study of cancer registry cases of Fars province (2001-2006). *ISMJ*. 2011; 14: 114-21.
24. Mehrabian A, Esna-Ashari F, Zham H, Hadizadeh M, Bohlooli M, Khayamzadeh M, et al. Gastric cancer prevalence, according to survival data in Iran (National Study-2007). *Iranian Journal of Public Health*. 2010; 39.
25. Biglarian A, Hajizadeh E, Kazemnejad A, Zali M. Survival analysis of gastric cancer patients using Cox model: a five year study. *Tehran University Medical Journal*. 2009; 67: 317-325.
26. Khodabakhshi R, Gouhari M, Biglarian A, Hajizadeh E. Survival analysis of patients with Gastric adenocarcinoma. *Kosar Medical Journal*. 2009; 12: 345-55.
27. Aghaei A, Ahmadi-Jouibari T, Baiki O, Mosavi-Jarrahi A. Estimation of the gastric cancer incidence in Tehran by two-source capture- recapture. *Asian Pac J Cancer Prev*. 2013; 14: 673-7.
28. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh J, Comber H, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *European Journal of Cancer*. 2013; 49: 1374-1403.
29. Babaei M, Jaafarzadeh H, Sadjadi A, Samadi F, Yazdanbod A, Fallah M, et al. Cancer incidence and mortality in Ardabil: Report of an ongoing population-based cancer registry in Iran, 2004-2006. *Iranian Journal of Public Health*. 2009; 38: 35-45.
30. Crocetti E, Miccinesi G, Paci E, Zappa M. An application of the two-source capture-recapture method to estimate the completeness of the Tuscany Cancer Registry, Italy. *European journal of cancer prevention*. 2001; 10: 417-23.
31. Gajalakshmi V, Swaminathan R, Shanta V. An independent survey to assess completeness of registration: population based cancer registry, Chennai, India. *Asian Pac J Cancer Prev*. 2001; 2: 179-83.
32. Estimated cancer mortality and prevalence worldwide in 2012. *International Agency for Research on Cancer*; 2012 [cited 2014 6/8/2014]; Available from: [http://globocan.iarc.fr/Pages/fact\\_sheets\\_population.aspx](http://globocan.iarc.fr/Pages/fact_sheets_population.aspx).
33. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA: A Cancer Journal for Clinicians*. 2011; 61: 69-90.
34. Matsuda T, Marugame T, Kamo K-i, Katanoda K, Ajiki W, Sobue T. Cancer incidence and incidence rates in Japan in 2002: based on data from 11 population-based cancer registries. *Japanese Journal of Clinical Oncology*. 2008; 38: 641-8.
35. GLOBOCAN. Estimated cancer incidence, mortality and prevalence worldwide in 2012. 2012; Available from: [http://globocan.iarc.fr/Pages/fact\\_sheets\\_population.aspx](http://globocan.iarc.fr/Pages/fact_sheets_population.aspx).

# Evaluation Completeness of Esophagus Cancer Registry in Ardabil Using Log-Linear Model

Khodadost M<sup>1</sup>, Yavari P<sup>2,3</sup>, Babaei M<sup>4,5</sup>, Sarvi F<sup>6</sup>, Hashemi Nazari SS<sup>7,8</sup>

1- Ph.D Student in Epidemiology, Gastroenterology and Liver Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran

2- Professor of Epidemiology, Dept. of Health & Community Medicine, School of Medicine, Shahid Beheshti University of Medical Sciences, Evin, Tehran, Iran

3- Iranian Epidemiological Association

4- Health Deputy, Ardabil University of Medical Sciences, Ardabil, Iran

5- German Cancer Research Center (DKFZ)

6- Ph.D Student in Biostatistics, Department of Biostatistics and Epidemiology, School of Public Health, Hamadan University of Medical Sciences, Hamadan, Iran

7- Safety Promotion and Injury Prevention Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

8- Assistant Professor of Epidemiology, Department of Epidemiology, School of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran

**Corresponding author:** Yavari P, p.yavari-grc@sbmu.ac.ir

**Background and Objectives:** completeness of registration is used as one of the measures of the quality of a cancer registry, which is the degree to which reportable incident cases of cancer in the population of interest is actually recorded in the registry.

**Methods:** After removing the duplicates, a total of 471 new cases of esophagus cancer reported by three sources of pathology reports, medical records, and death certificates to Ardabil Province Cancer Registry Center in 2006 and 2008 were enrolled in the study. The incidence rate was estimated based on the capture-recapture method and the use of the log-linear models. BIC, G<sub>2</sub> and Akaike statistics were used to select the best-fit model.

**Results:** In this study, a model with linkage between pathology reports and medical records and a model with death certificates alone, independent of the previous two sources, was the best fitted model. The estimated total completeness of esophagus cancer in 2006 and 2008 was 36%. The source that had the most completeness for esophagus cancers was pathology reports with 21.17%. The estimated incidence rate calculated by the log-linear method for the years 2006 and 2008 was 49.71 and 53.87 per 100,000 population, respectively.

**Conclusion:** Based on the obtained results, it can be concluded that the low degree of completeness in Ardabil Province requires some changes in data abstracting and case finding such as the use of personal national code and electronic health records to create a more accurate cancer registry.

**Keywords:** Capture-recapture, Incidence estimation, Log-linear, Esophagus cancer, Cancer registration, Completeness