مقاله پژوهشی
(Original paper)

مقایسه برآورد نسبت شانس حاصل از دو مدل رگرسیون لجستیک و رگرسیون لجستیک شرطی در مطالعه مورد- شاهدی تعیین عوامل خطر مسمومیت‌های غیر عمد کودکان در تهران

کامیار منصوری[۱]، شیوا منصوری هانیس[۲*]، حمید سوری[۱]، فاطمه خسروی شادمانی[۳]، فرهاد مرادپور[۴]، عبدالحلیم رجبی[۴]، علی غلامی[۴]

۱- گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شهیدبهشتی،تهران، ایران.

۲- کمیته تحقیقات دانشجویی، دانشگاه علوم پزشکی کردستان، سنندج، ایران

(مولف مسول) ایمیل: mansorishiva@gmail.com

۳- گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی کرمانشاه ،کرمانشاه ، ایران.

۴ - گروه اپیدمیولوژی، دانشکده بهداشت ، دانشگاه علوم پزشکی ایران، تهران، ایران.

چکیده

زمینه و هدف: شناخت عوامل خطر موثر در بروز مسمومیت یک اقدام اساسی برای پیشگیری از مسمومیت در هر جامعه‌ای است، در این میان یافتن معادله‌ای برای تعیین اثر این عوامل و شدت ارتباط آنها با ابتلا به مسمومیت، دارای اهمیت فراوانی است هدف مطالعه حاضر مقایسه عملکرد دو مدل رگرسیون لجستیک معمولی و رگرسیون لجستیک شرطی در تعیین عوامل خطر موثر بر ایتلا به مسمومیت‌های غیر عمد کودکان است.

روش بررسی: در مطالعه حاضر از داده‌های یک مطالعه مطالعه مورد- شاهدی که در سال ۱۳۹۲ با هدف تعیین عوامل خطر مسمومیت‌های خطر عمد کودکان در تهران انجام گرفته بود استفاده شد. داده‌های مربوطه پس از جمع‌آوری در نرم افزار stata وارد و برای تعیین عوامل خطر موثر بر ایتلا به مسمومیت از دو مدل رگرسیون لجستیک شرطی (Conditional Logistic Regression) و رگرسیون لجستیک معمولی (Logistic Regression) استفاده شد. سپس نسبت شانس (OR) با فاصله اطمینان (CI) ۹۵٪ بصورت خام و تعدیل شده همراه با حساسیت، ویژگی و سطح زیر منحنی راک برای هر کدام از دو مدل فوق برآورد گردید، سپس به مقایسه آنها پرداخته شد.

یافته‌ها: در مدل چندگانه رگرسیون لجستیک شرطی اشتغال مادر، مسمومیت قبلی، اعتیاد در خانواده و دور از دسترس بودن مواد مسمومیت زا، بر ابتلا به مسمومیت تاثیر قابل توجهی نشان دارد اما در مدل چندگانه رگرسیون لجستیک معمولی علاوه بر موارد ذکر شده، بیماری روحی/روانی در خانواده، نیز رابطه معنی‌داری را با پیامد (مسمومیت) نشان داد. مقدار سطح زیر منحنی ROC برای مدل رگرسیون لجستیک شرطی (۰/۹۷ - ۰/۹۴ :CI ۹۵٪) ۹۵/۸۲٪ با حساست ۹۲/۹ درصد و ویژگی ۸۴/۲ درصد و برای مدل رگرسیون لجستیک معمولی (۰/۹۳- ۰/۸۶ :CI ۹۵٪) ۸۹/۷۵ ٪ با حساسیت ۸۳/۶ درصد و ویژگی ۸۳/۱ درصد بدست آمد.

نتیجه‌گیری: اگرچه فاصله اطمینان‌های مربوط به مدل رگرسیون لجستیک شرطی در مقایسه با رگرسیون لجستیک معمولی دامنه کمتری را در بر می‌گیرد و از سویی دیگر بالاتر بودن سطح زیر منحنی ROC، نشان دهنده خوب بودن دقت کلی این مدل در تشخیص مسمومیت است. اما باید توجه داشت استفاده از هرکدام از این مدل‌ها به خصوصیات داده‌های مورد بررسی (همسان شده و نشده) وابسته است.

واژه‌های کلیدی: مسمومیت در کودکان، رگرسیون لجستیک شرطی، رگرسیون لجستیک معمولی، مورد شاهدی

[Downloaded from zanko.muk.ac.ir on 2025-11-19]

مقدمه

مسمومیت یکی از مشکلات اساسی بهداشت عمومی است و هنوز یکی از علل مهم پذیرش در بخش‌های اورژانس به حساب می‌آید (۱). بطورکلی مسمومیت حاد عامل مرگ سالانه حدود ۴۵۰۰۰ نفر از کودکان و نوجوانان زیر ۲۰ سال است. میزان جهانی مرگ ناشی از مسمومیت برای کودکان زیر۲۰ سال ۱/۸ در صد هزار نفر است که برای کشورهای با درآمد بالا این میزان ۰/۵ در صد هزار نفر است در حالی که برای کشورهای با درآمد کم و متوسط ۴ برابر بیشتر است (۲ در صدهزار) (۲). گرچه بسیاری از این مسمومیت‌ها توسط کادر پزشکی- درمانی کنترل شده و به مرگ منتهی نمی‌شود ولیکن خسارات جبران ناپذیری را به پیکره اقتصاد، سلامت و بهداشت جسمی و روانی جامعه و خانواده وارد خواهد نمود و نیز بار عاطفی و اضطراب خانوادگی ناشی از آن نیز سنگین بوده و باید هنگام ارزیابی ابعاد این مساله مورد توجه قرار گیرد (۳).

بنابراین شناخت عوامل خطر موثر در بروز مسمومیت یک اقدام اساسی برای برنامه‌های پیشگیری از مسمومیت در هر جامعه‌ای است، چرا که کاهش دادن این عوامل خطر باعث کاهش نرخ بروز مسمومیت خواهد شد، در این میان یافتن معادله‌ای برای تعیین اثر عوامل خطر و شدت ارتباط آنها با ابتلا به مسمومیت، دارای اهمیت فراوانی است (۴). مطالعات مورد شاهدی یکی از متداول‌ترین طرح‌های مطالعاتی در اپیدمیولوژی تحلیلی است که برای چنین مواقعی بکار برده می‌شود (۵). هدف چنین مطالعه‌ای بررسی عوامل بالقوه موثر در ایجاد بیماری است. در این نوع مطالعه مشاهده‌ای برای تعیین ارتباط یک مواجهه (عامل بیماری) و یک بیماری خاص، یک گروه از افراد مبتلا به آن بیماری (مورد) را در نظر گرفته و یک گروه از افرادی که به آن بیماری مبتلا نیستند (شاهد) را با آنها مقایسه می‌کنند (۶).

بطور کلی در مطالعات اپیدمیولوژیک تعیین تأثیر یک مواجهه بر خطر ابتلا به بیماری معمولاً توسط خطر نسبی (Relative Risk) بیان می‌گردد (۷). در مطالعات مورد شاهدی هم نسبت شانس (Odds Ratio) یک اندازه خطر است که نشان دهنده آن است که چقدر شانس کسی که با یک عامل خطر مواجهه داشته از کسی که مواجهه نداشته بیشتر است. نسبت شانس (Odds Ratio) برای متغیر پاسخ هم از طریق جداول توافقی هم روش مدلسازی قابل محاسبه است. در این مقاله روش مدلسازی مد نظر است (۶ و۸). از بین کل متغیرهای این مطالعه، متغیر پاسخ، مسمومیت در نظر گرفته شده و با مدلسازی، هدف تعیین میزان تأثیر سایر متغیرها بر مسمومیت بود.

در مطالعه حاضر به مقایسه برآورد نسبت شانس (Odds Ratio)، حساسیت، ویژگی و سطح زیر منحنی ROC دو مدل رگرسیون لجستیک و رگرسیون لجستیک شرطی توسط داده‌های حاصل از یک مطالعه مورد- شاهدی که با هدف تعیین عوامل خطر مسمومیت‌های غیرعمد کودکان در تهران انجام گرفته بود (۱) پرداخته خواهد شد. لازم به ذکر است که مدل رگرسیون لجستیک شرطی همانند مدل لجستیک معمولی است. تنها تفاوت این دو روش این است که رگرسیون لجستیک شرطی برای داده‌های همسان شده فردی به کار می‌رود. به عبارتی دیگر هنگامی که در یک مطالعه مورد-شاهدی هریک از افراد گروه شاهد با نظیر خود در گروه مورد از نظر یک یا چند متغیر همسان شده باشد باید برای تجزیه و تحلیل آماری از رگرسیون لجستیک شرطی استفاده شود.

[Downloaded from zanko.muk.ac.ir on 2025-11-19]

روش بررسی

در این مطالعه برای پیش‌بینی مسمومیت در کودکان، مدل رگرسیون لجستیک و رگرسیون لجستیک شرطی اعمال شد و همچنین شاخص‌هایی مانند حساسیت و ویژگی و سطح زیر منحنی ROC برای سنجش میزان قدرت و دقت مدل محاسبه گردید. در این قسمت هر کدام از این مدل‌ها به اختصار توضیح داده می شوند.

مدل رگرسیون لجستیک:

تشخیص الگوها و طبقه‌بندی یکی از مهمترین کاربردهای روشهای آماری در علوم مختلف است. از جمله اهداف عمده مدل سازی در علوم آمار، پیش‌بینی بر اساس شواهد، متغیرها و داده‌های موجود از یک موضوع خاص است. این امر در علوم آماری توسط روشهایی مانند، رگرسیون، تحلیل ممیزی (جداسازی)، سری‌های زمانی، رده‌بندی، رگرسیون درختی و سایر روشها انجام می‌شود. در نظر گرفتن یک توزیع پیش فرض مانند توزیع نرمال برای متغیرهای پاسخ، خطی بودن رابطه پیشنهادی، یکسان بودن واریانس خطاها،... از جمله محدودیت‌های بعضی روشهای کلاسیک هستند که هنگام استفاده علمی از این روشها، اگر داده-های واقعی شرایط مفروض مدل را نداشته باشند امکان‌پذیر نبوده یا با خطای قابل توجه همراه است (۹). در بین این روشها رگرسیون لجستیک فرضیات زیادی را لازم ندارد و از جمله روشهای آماری چند متغیره‌ای است که می‌تواند برای ارزیابی ارتباط بین متغیرهای مستقل هر چند مخدوش کننده و متغیر وابسته (پیشامد طبقه‌ای) و پیش‌بینی ابتلا به بیماری بر اساس متغیرهای پیشگو در مدل مورد استفاده قرار می‌گیرد (۱۰). رگرسیون لجستیک زمانی مورد استفاده قرار می‌گیرد که متغیر وابسته به صورت دوتایی، اسمی یا ترتیبی باشد و برای متغیرهای مستقل هیچ محدودیتی وجود ندارد. در علوم پزشکی متغیر پیشامد به طور معمول حضور یا عدم آن از یک وضعیت بیان شده یا یک بیماری می‌باشد. مفهوم اصلی ریاضی که پایه و اساس رگرسیون لجستیک است، لجیت یعنی لگاریتم طبیعی نسبت شانس (Odds Ratio). رگرسیون لجستیک بر اساس یک فرضیه که شامل یک ارتباط لجستیکی موجود بین احتمال عضویت گروه و یک یا چند متغیر پیش‌بینی‌کننده می‌باشد بنا نهاده شده است (۱, ۱۱, ۱۲). از آنجایی که احتمال پیش‌بینی شده باید بین اعداد صفر و یک قرار گیرد، تکنیک‌های رگرسیون خطی برای دستیابی به آن کفایت نمی‌کند، به این دلیل که آنها به متغیرهای وابسته اجازه داده اند که از محدودیت ها گذشته، نتایج ناسازگار تولید کنند. با تعریف P_0 به عنوان احتمال تعلق یک مشاهده به گروه غیر بیمار و P_1 احتمال تعلق یک مشاهده به گروه بیمار، مدل رگرسیون لجستیک بصورت زیر است:

$$Z_i = \log\left(\frac{P_{i1}}{P_{i0}}\right) = b_0 + b_1 X_{i1} + b_2 X_{i2} + \cdots + b_k X_{ik}$$

که $\frac{P_{i1}}{P_{i0}}$ نسبت شانس نامگذاری شده است. b_j مقدار j امین ضریب را که j= 1,....K و X_{ij} مقدار i امین مشاهده از j امین متغیر پیش‌بینی‌کننده است. پارامترهای b_0 تا b_k از مدل لجستیک با استفاده از روش ماکزیمم درستنمایی برآورد می‌شوند. در معادله ذکر شده تبدیل لجیت جهت مرتبط ساختن احتمالات عضویت گروه به یک تابع خطی از متغیرهای پیش-بینی‌کننده مورد استفاده قرار می گیرد. (۱۰-۱۲). مقادیر P_0 و P_1 احتمال اختصاص مشاهده به گروههای کنترل و مسموم بود و در نهایت جهت ساختن آماره Z تحت الگاریتم طبیعی توسط نرم افزار بدست می‌آید. از مزایای استفاده ازمدل رگرسیون

[Downloaded from zanko.muk.ac.ir on 2025-11-19]

لجستیک علاوه بر مدلسازی مشاهده‌ها، امکان پیش‌بینی احتمال تعلق هر فرد به هر یک از سطوح متغیر وابسته و همچنین امکان محاسبه مستقیم نسبت شانس (Odds Ratio) با استفاده از ضرایب مدل است (۹).

مدل رگرسیون لجستیک شرطی:

همسان‌سازی (matching) یکی از روشهای کنترل اثر متغیرهای مخدوش‌کننده در طراحی مطالعه به شمار می‌آید. معمولا در مطالعات مورد شاهدی برای اطمینان از قابلیت مقایسه کیس و کنترل‌ها از همسان سازی (matching) استفاده می‌شود. در این فرایند گروه شاهد بنحوی انتخاب می‌شود که از نظر برخی متغیرها مشابه گروه مورد باشد. همسان سازی معمولاً با دو هدف صورت صورت گیرد یکی افزایش دقت مطالعه (precision) و دیگری کنترل اثر متغیر یا متغیرهای مخدوش‌کننده‌ای(confounding Variables) که به آسانی نمی‌توان آنها را تعیین کرد. همسان سازی به دو صورت امکان‌پذیر است که شامل همسان‌سازی فردی (individual matching) و همسان‌سازی گروهی (frequency matching) می‌باشد (۱۳). هنگامی که در مطالعات مورد شاهدی همسان‌سازی فردی (individual matching) صورت گیرد یعنی به ازای هر فرد گروه مورد یک یا چند فرد از گروه شاهد که از نظر متغیر(های) مخدوش‌کننده مشابه باشد انتخاب شود از رگرسیون لجستیک شرطی (conditional logistic regression) استفاده می‌شود. چرا که در این موارد که با تعداد زیادی از طبقات و به تبع آن با حجم نمونه کم در طبقات روبه رو هستیم و با افزایش حجم نمونه تعداد پارامترها زیاد می‌شود برآورد حاصل از برازش مدل رگرسیون لجستیک استاندارد دارای اریبی خواهد بود. از طرفی برآورد پارامتر مربوط به طبقات در مطالعات همسان سازی شده اهمیت چندانی ندارد لذا با بکارگیری مدل رگرسیون لجستیک شرطی بجای مدل رگرسیون لجستیک استاندارد با حذف پارامتر طبقات (پارامترهای مزاحم) برآورد نااریبی از پارامترهای مهم که ارتباط عوامل خطر و پاسخ را نشان می‌دهد خواهیم داشت. شکل کلی تابع لجستیک شرطی برای مطالعات مورد شاهدی همسان‌سازی فردی بصورت زیر است:

$$L_{con} = \prod_{i=1}^{k}\left(\frac{\prod_{j=1}^{n1i}\exp(\beta x\acute{i}_J)}{\sum_{li}\prod_{j=1}^{n1i}\exp(\beta x\acute{i}_J)}\right)$$

n_1 تعداد موردها، β بردار پارامترها (vector of parameters) و X بردار متغیرهای مستقل (vector of covariates)می‌باشد و k نیز تعداد مجموعه‌های همسان‌سازی شده (numbers of match sets) یا همان strata است (۱۴, ۱۵).

داده‌ها:

در مطالعه حاضر از داده های یک مطالعه مطالعه مورد- شاهدی که در سال ۱۳۹۲ با هدف تعیین عوامل خطر مسمومیت های خطر عمد کودکان در تهران انجام گرفته بود استفاده شد. که در آن مطالعه، موردها از میان کودکان با تشخیص مسمومیت‌های غیرعمد که بطور متوالی از تاریخ بیستم اسفند ۱۳۹۱ تا بیست و پنجم تیر ۱۳۹۲ به بخش مسمومیت اطفال بیمارستان لقمان تهران مراجعه کرده بودند انتخاب شده بودند. سپس به ازای هر مورد، ۲ شاهد از میان کودکانی که به دلایل دیگری غیر مسمومیت از قبیل تنفسی، جراحی و گوارشی به سایر بخش‌های این بیمارستان مراجعه کرده بودند انتخاب شده بودند. بیماریهای نورولوژیک، بیماریهای متابولیک، عفونت‌های مغزی و تروماهای مغزی به عنوان معیارهای خروج برای شاهدها در نظر گرفته شده شدند. سپس همسان سازی فردی صورت گرفته بود و شاهدها از لحاظ سن (۶±

[Downloaded from zanko.muk.ac.ir on 2025-11-19]

ماه) و جنس با موردها همسان شده بودند. به منظور جمع‌آوری اطلاعات همه والدین کودکان توسط محقق از طریق یک پرسشنامه که شامل خصوصیات جمعیت شناختی، رفتاری و عوامل خطر مسمومیت‌های غیرعمد از جمله اشتغال مادر، تحصیلات والدین، سیگاری بودن والدین، مسمومیت قبلی، اعتیاد در خانواده، بیماری روحی/روانی در خانواده، دور از دسترس بودن مواد مسمومیت‌زا و بعد خانوار بود مورد مصاحبه قرار گرفتنه بودند. برای موردها اطلاعاتی درباره نوع وشرایط مسمومیت نیز جمع‌آوری شده بود (۴).

داده‌های مربوطه پس از جمع‌آوری در نرم افزار stata وارد سپس تعیین عوامل خطر به دو روش صورت گرفت: در روش اول از آنالیزهای تک متغیره و چند متغیره رگرسیون لجستیک شرطی (Conditional Logistic Regression) با فرض همسان سازی فردی و در روش دوم نیز از آنالیزهای تک متغیره و چند متغیره رگرسیون لجستیک معمولی (Logistic Regression) با فرض عدم همسان‌سازی فردی استفاده شد. در نهایت نسبت شانس (OR) با فاصله اطمینان (CI) ۹۵٪ بصورت خام و تعدیل شده همراه با حساسیت، ویژگی و سطح زیر منحنی راک برای هر کدام از دو مدل فوق برآورد گردید، سپس به مقایسه آنها پرداخته شد. لازم به ذکر است که در این مطالعه درست است که در ابتدا داده‌ها از نظر سن و جنس همسان سازی شده بودند اما خود نرم افزار stata می‌تواند داده‌های همسان سازی شده را ناهمسان فرض نماید و نسبت شانس (OR) را برای این حالت (unmatch) نیز برآورد نماید.

یافته‌ها

در مطالعه‌ای که ذکر شد ۱۴۰ کودک مسموم با ۲۸۰ کنترل مورد بررسی قرار گرفته بودند. جدول ۱ توزیع خصوصیات مربوط به خانواده ۱۴۰ کودک مسموم مراجعه‌کننده به بیمارستان لقمان حکیم تهران و ۲۸۰ شاهد را که از لحاظ سن (۶±ماه) و جنس همسان سازی شده بودند را نشان می‌دهد.

در ابتدا به منظور شناسایی مهم‌ترین عوامل تاثیرگذار بر مسمومیت به ترتیب اقدام به آنالیز تک متغیره رگرسیون لجستیک شرطی و رگرسیون لجستیک گردید و برای اینکه هیچ متغیر مهمی نادیده گرفته نشود، در ابتد سطح معنی‌داری $\alpha=۰/۱$ در نظر گرفته شد. نتایج این هر دو این آنالیزها نشان داد که اشتغال مادر، تحصیلات پدر، تحصیلات مادر، سیگاری بودن پدر، سیگاری بودن مادر، مسمومیت قبلی، اعتیاد در خانواده، بیماری روحی/روانی در خانواده، دور از دسترس بودن مواد مسمومیت زا و بعد خانوار بر ابتلا به مسمومیت تاثیر قابل توجهی دارد ($P<۰/۱$). (جدول۱).

سپس به منظور بررسی اثر همزمان متغیرها و حذف متغیرهای احتمالی مخدوش‌کننده، متغیرهای با سطح معنی‌داری کمتر از ۰/۱ در هر کدام از مدل های فوق بطور همزمان به ترتیب در مدل چندگانه رگرسیون لجستیک شرطی و رگرسیون لجستیک وارد کردیم و با استفاده از روش گام به گام پسرو (backward Stepwise) به حذف مخدوش کنندها و تعدیل یافته-های معنی‌دار پرداخته شد. نتایج حاصل از آنالیز مدل چندگانه رگرسیون لجستیک شرطی نشان داد که اشتغال مادر، مسمومیت قبلی، اعتیاد در خانواده و دور از دسترس بودن مواد مسمومیت‌زا بر ابتلا به مسمومیت تاثیر قابل توجهی دارد ($P<۰/۰۵$) (جدول ۲). اما نتایج حاصل از مدل چندگانه رگرسیون لجستیک نشان داد

[Downloaded from zanko.muk.ac.ir on 2025-11-19]

که علاوه بر موارد اثرگذار بدست آمده از مدل چندگانه رگرسیون لجستیک شرطی که با ابتلا به مسمومیت رابطه معنی‌داری داشتند، بیماری روحی/روانی در خانواده کودک نیز رابطه معنی‌داری را با پیامد (مسمومیت) نشان داد (P<۰/۰۵) (جدول ۳).

از تحلیل ROC به عنوان مقیاس اندازه‌گیری توانایی جداسازی یک مدل با بیشترین ناحیه که نشان‌دهنده توانایی پیش‌بینی بهتر است برای مقایسه انجام مدل‌ها بکار می‌رود، استفاده می‌شود. لازم به ذکر است که دقت بالای ۹۰ درصد عالی، بین ۹۰-۸۰ درصد خوب، بین ۸۰-۷۰ درصد قابل قبول، بین ۷۰-۶۰ درصد ضعیف و چنانچه دقت آزمونی زیر ۶۰ درصد گردد، غیر قابل قبول می‌باشد. در این مطالعه مقدار سطح زیر منحنی ROC برای مدل رگرسیون لجستیک شرطی (۰/۹۷-۰/۹۴ CI: ٪۹۵) ٪۹۵/۸۲ با حساست ۹۲/۹ درصد و ویژگی ۸۴/۲ درصد (شکل ۱) و برای مدل رگرسیون لجستیک (۰/۹۳-۰/۸۶ CI: ٪۹۵) ٪۸۹/۷۵ با حساست ۸۳/۶ درصد و ویژگی ۸۳/۱ درصد بدست آمد (شکل ۲).

جدول۱: عوامل موثر بر ابتلا مسمومیت با استفاده از دو مدل تک متغیره لجستیک شرطی و لجستیک معمولی

			مدل رگرسیون لجستیک معمولی		مدل رگرسیون لجستیک شرطی	
متغیر	**شاهدها (٪)**	**موردها (٪)**	**P-value**	**(فاصله اطمینان OR(٪۹۵**	**P-value**	**(فاصله اطمینان OR(٪۹۵**
سن (سال)						
≤۱	۴۰(۱۴/۳)	۱۶(۱۱/۴)	-			
۲-٤	۱۶۴(۵۸/۶)	۸۴(۶۰)		-	-	-
>٥	۷۶(۲۷۸/۱)	۴۰ (۲۸/۶)				
جنس						
پسر	۱۶۰ (۵۷/۱)	۸۰ (۵۷/۱)	-	-	-	-
دختر	۱۲۰ (۴۲/۹)	۶۰ (۴۲/۹)				
اشتغال مادر						
خانه دار	۲۶۴ (۹۴/۳)	۱۲۳(۸۹/۷)	۰/۰۲۴	۱	۰/۰۳	۱
سایر	۱۶ (۵/۷)	۱۷(۱۲/۹)		(۱/۱۱ - ۴/۶۶) ۲/۲۸		۲/۲ (۴/۵- ۱/۱)
تحصیلات مادر						
دانشگاهی	۴۹ (۱۷/۵)	۱۳ (۹/۳)	<۰/۰۰۱	۱	<۰/۰۰۱	۱
غیردانشگاهی	۲۳۱ (۸۲/۵)	۱۲۷ (۹۰/۷)		۰/۸۷ (۰/۸۲- ۰/۹۲)		(۰/۸۱ -۰/۹۲) ۰/۸۶
تحصیلات پدر						
دانشگاهی	۶۹ (۲۴/۷)	۱۶ (۱۱/۴)	<۰/۰۰۱	۱	<۰/۰۰۱	۱
غیردانشگاهی	۲۱۱ (۷۵/۳)	۱۲۴ (۸۸/۶)		۰/۸۷ (۰/۸۲- ۰/۹۱)		(۰/۸۲- ۰/۹۲) ۰/۸۷
سیگاری بودن پدر						
خیر	۲۱۵ (۷۶/۸)	۴۳ (۳۰/۷)	<۰/۰۰۱	۱	<۰/۰۰۱	۱
بله	۶۵ (۲۳/۲)	۹۷ (۶۹/۳)		۷/۴۷ (۴/۷- ۱۱/۷)		۷/۳ (۴/۴-۱۲/۲)

[Downloaded from zanko.muk.ac.ir on 2025-11-19]

سیگاری بودن مادر						
خیر	۲۷۹ (۹۹/۶)	۱۳۶ (۹۷/۱)	۰/۰۶۱	۱	۰/۰۶	۱
بله	۱ (۰/۴)	۴ (۲/۹)		۸/۲ (۰/۹ – ۷۴/۱)		۸ (۰/۸۹– ۷۱/۱)
وضعیت سرپرستی						
کودک	۲۸۰ (۱۰۰)	۱۳۵ (۹۶/۴)	–	–	–	–
با هردو والد	۰	۵ (۳/۶)				
سایر						
مسمومیت قبلی						
خیر	۲۶۲ (۹۳/۶)	۹۹ (۷۰/۷)	<۰/۰۰۱	۱	<۰/۰۰۱	۱
بله	۱۸ (۶/۴)	۴۱ (۲۹/۳)		۶/۲ (۳/۳ – ۱۱)		۶/۵ (۳/۳– ۱۲/۸)
اعتیاد در خانواده						
خیر	۲۵۷ (۹۱/۸)	۵۳ (۳۷/۹)	<۰/۰۰۱	۱	<۰/۰۰۱	۱
بله	۲۳ (۸/۲)	۸۷ (۶۲/۱)		(۱۱/۴ – ۳۵/۲)		(۸/۷–۳۴/۶)
				۲۰/۱		۱۷/۴
بیماری روحی- روانی						
در خانواده	۲۶۵ (۹۴/۶)	۱۱۰ (۷۸/۶)	<۰/۰۰۱	۱	<۰/۰۰۱	۱
خیر	۱۵ (۵/۴)	۳۰ (۲۱/۴)		۴/۸ (۲/۵ – ۹/۳)		۴/۷ (۲/۴– ۹/۱)
بله						
دور از دسترس بودن						
محصولات مسمومیت زا	۲۷۰ (۹۶/۴)	۷۵ (۵۳/۶)	<۰/۰۰۱	۱	<۰/۰۰۱	۱
بله	۱۰ (۳/۶)	۶۵ (۴۶/۴)		۰/۴۳ (۰/۰۲– ۰/۸۸)		(۰/۰۱–۰ /۰۹)
خیر						۰/۰۳
بعد خانوار(تعداد)						
<٤	۱۲۸ (۴۵/۷)	۴۲ (۳۰)	<۰/۰۰۱	۱	<۰/۰۰۱	۱
≥٤	۱۵۲ (۵۴/۳)	۹۸ (۷۰)		۱/۶۲ (۱/۳– ۲/۰۴)		۱/۷۶ (۱/۴–۲/۳)
	شاهدها (٪)	موردها (٪)	P-value	(فاصله اطمینان ٪۹۵ OR	P-value	(فاصله اطمینان ٪۹۵ OR

جدول۲: عوامل موثر بر ابتلا به مسمومیت با استفاده از مدل چندگانه رگرسیون لجستیک شرطی

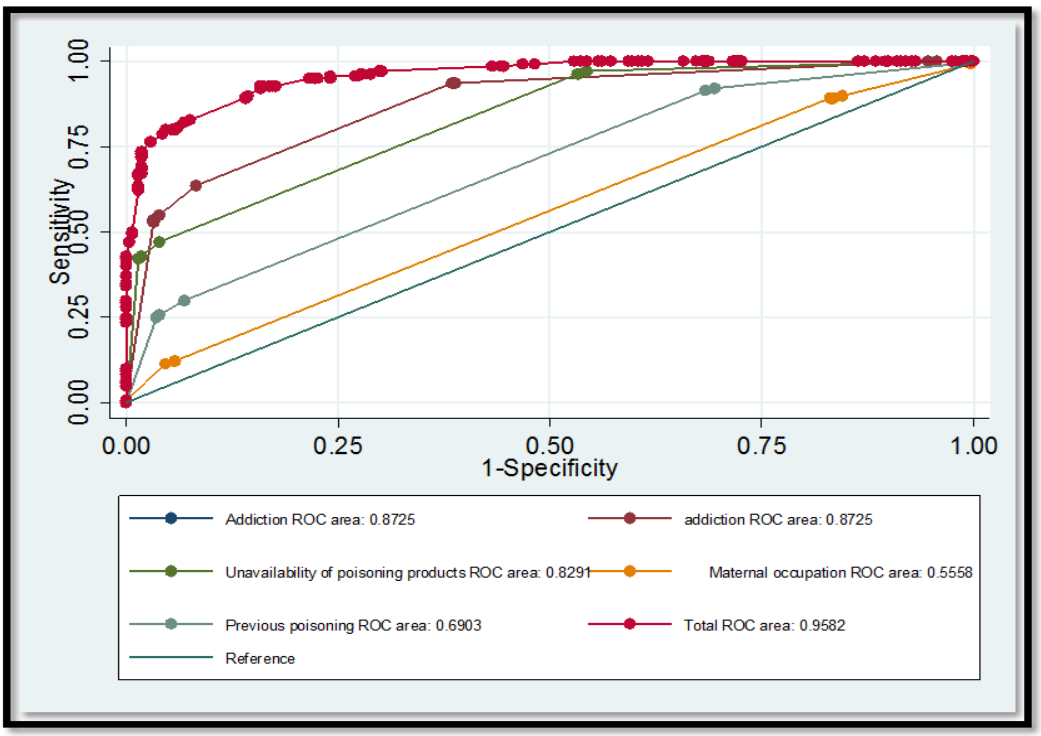
متغیرها	فاصله اطمینان ٪۹۵	OR	P-value
اشتغال مادر	۱/۳ – ۱۲/۳	۴	۰/۰۳
مسمومیت قبلی	۲/۴ – ۲۰/۲	۷	<۰/۰۰۱
اعتیاد در خانواده	۶/۲ –۳۴/۶	۱۴/۶	<۰/۰۰۱
دور از دسترس بودن محصولات مسمومیت زا	۰/۰۱ – ۰/۱۲	۰/۰۳	<۰/۰۰۱

[Downloaded from zanko.muk.ac.ir on 2025-11-19]

جدول ۲:عوامل موثر بر ابتلا به مسمومیت با استفاده از مدل چندگانه رگرسیون لجستیک معمولی

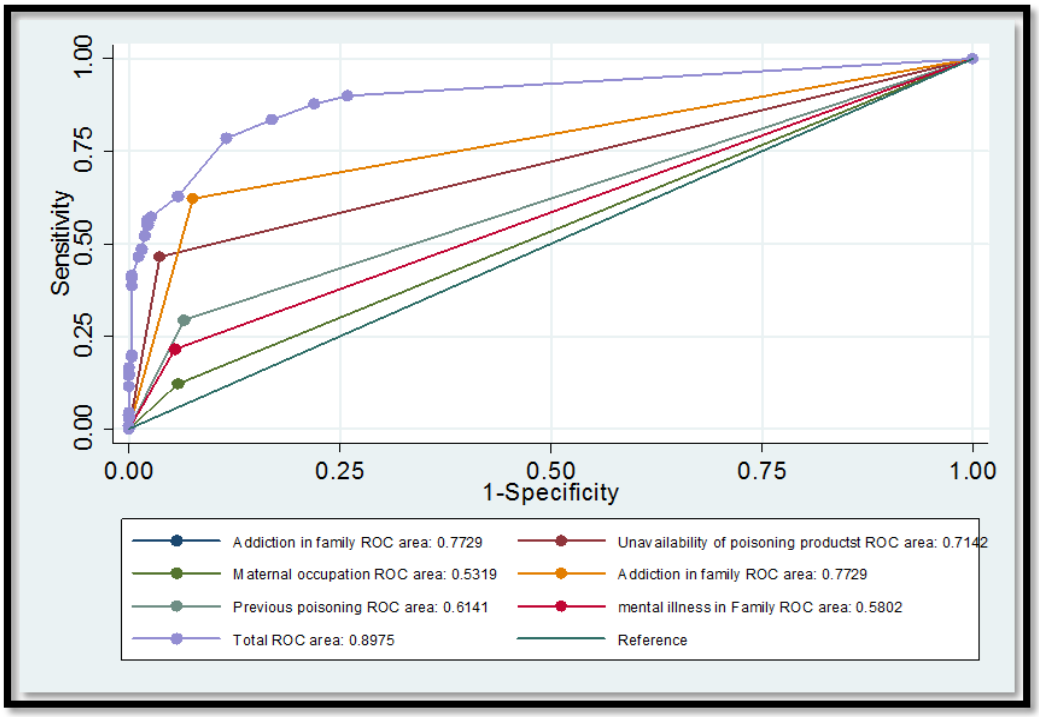
متغیرها	فاصله اطمینان ۹۵٪	OR	P-value
اشتغال مادر	۲/۳ – ۱۴/۸	۵/۸	۰/۰۳
مسمومیت قبلی	۲/۵ – ۱۳/۸	۵/۹	<۰/۰۰۱
اعتیاد در خانواده	۸/۹ -۳۴/۲	۱۷/۵	<۰/۰۰۱
دور از دسترس بودن محصولات مسمومیت زا	۰/۰۲ – ۰/۱۲	۰/۰۵	<۰/۰۰۱
بیماری روحی _روانی در خانواده	۱/۳ – ۹/۱	۳/۵	۰/۰۱

شکل ۱: سطح زیر منحنی راک، حساسیت و ویژگی حاصل از مدل چندگانه رگرسیون لجستیک شرطی



[Downloaded from zanko.muk.ac.ir on 2025-11-19]

شکل۲: سطح زیر منحنی راک، حساسیت و ویژگی حاصل از مدل چندگانه رگرسیون معمولی



بحث و نتیجه‌گیری

با توجه به اهمیت بیماری مسمومیت در کودکان و بالا بودن درصد ابتلا به آن در جوامع مختلف، امر پیش بینی این بیماری حائز اهمیت می‌باشد و در این راستا انتخاب مدل آماری مناسبی که بتواند مشاهدات را به درستی برای داشتن بیماری پیش‌بینی کند و یا اینکه آیا کودک در آینده شانس ابتلا به مسمومیت را دارد، دارای اهمیت زیادی است.

نتایج آنالیز تک متغیره هر دو مدل رگرسیون لجستیک شرطی و رگرسیون لجستیک نشان داد که اشتغال مادر، تحصیلات پدر، تحصیلات مادر، سیگاری بودن پدر، سیگاری بودن مادر، مسمومیت قبلی، اعتیاد در خانواده، بیماری روحی/روانی در خانواده، دور از دسترس بودن مواد مسمومیت‌زا و بعد خانوار بر ابتلا به مسمومیت تاثیر قابل توجهی دارد (P<۰/۱). نتایج حاصل از آنالیز مدل چندگانه رگرسیون لجستیک شرطی نیز نشان داد که ۴ مواجهه اشتغال مادر، مسمومیت قبلی، اعتیاد در خانواده و دور از دسترس بودن مواد مسمومیت‌زا بر ابتلا به مسمومیت تاثیر قابل توجهی دارد (P<۰/۰۵). اما نتایج حاصل از مدل چندگانه رگرسیون لجستیک نشان داد که علاوه بر ۴ مواجهه حاصل از مدل چندگانه رگرسیون لجستیک شرطی، بیماری روحی/روانی در خانواده کودک نیز رابطه معنی‌داری را با ابتلا به مسمومیت دارد (P<۰/۰۵). نکته قابل توجه این است که نسبت شانس‌های برآوردشده هم در حالت تک متغیره و هم چند متغیره برای هر کدام از مواجهه‌ها توسط مدل رگرسیون لجستیک شرطی در مقایسه با مدل رگرسیون لجستیک معمولی کوچکتر بوده و فاصله اطمینان های مربوط به آن دامنه کمتری را در بر می-گیرد و به عبارت دیگر دقت برآورد نسبت شانس در

[Downloaded from zanko.muk.ac.ir on 2025-11-19]

این مدل بیشتر از مدل لجستیک معمولی است. علت آن را می‌توان اینگونه توجیه کرد که در یک مطالعه مورد-شاهدی همسان‌سازی شده توزیع متغیرهای مخدوش‌کننده‌ای که همسان‌سازی بر مبنای آنها صورت گرفته در گروه‌های مورد و شاهد مشابه است و در نتیجه در همه طبقات ایجاد شده تعداد موردها و شاهدها تقریباً مشابه است، و به ویژه طبقه‌ای نخواهد بود که تنها مورد یا شاهد در آن باشد و در گروه مقابل فردی برای مقایسه وجود نداشته باشد که این تعادل می‌تواند باعث کاهش واریانس در پارامترهای مورد بررسی و به دنبال آن نیز افزایش efficiency آماری (efficiency statistical) و در نهایت افزایش دقت مطالعه نسبت به حالتی که مطالعه همسان‌سازی نشده خواهد شد (۱۶و۱۷).

در مطالعه حاضر حساسیت و ویژگی حاصل از مدل لجستیک شرطی (حساست ۹۲/۹٪، ویژگی ۸۴/۲٪) بالاتر از مدل رگرسیون لجستیک (حساسیت ۸۳/۶٪، ویژگی ۸۳/۱٪) بود که می‌توان گفت این مدل پیش-بینی مناسب‌تری از مشاهداتی که به گروه‌های مربوطه تخصیص داده شده‌اند را اعمال کرده است. در واقع درصد کودکان مبتلا به مسمومیت که توسط این مدل به درستی تشخیص داده شده‌اند ۹۲/۹٪ و نیز درصد کودکان غیر مسموم که توسط این مدل به درستی پیش‌بینی شده‌اند ۸۴/۲٪ بوده است. همچنین مقدار سطح زیر منحنی ROC برای مدل رگرسیون لجستیک شرطی با ۹۵/۸۲٪ بیشتر از سطح زیر این منحنی برای مدل رگرسیون لجستیک (۸۹/۷۵٪) بود، که با نتایج مطالعه Rahman و همکاران که به مقایسه سطح زیر منحنی ROC در دو مدل رگرسیون لجستیک و رگرسیون لجستیک شرطی پرداخته بودند همخوانی دارد (۱۶). در واقع بالاتر بودن سطح زیر منحنی ROC نشانگر مناسب‌تر بودن مدل رگرسیون لجستیک شرطی است چرا که توانایی پیش‌بینی بهتری را دارد.

با تمام این اوصاف با قاطعیت نمی‌توان مدل مشخصی را به عنوان مناسب‌ترین مدل معرفی کرد، اما توصیه شده است اولاً در تصمیم‌گیری در مورد استفاده از مدل رگرسیون لجستیک شرطی و مدل رگرسیون لجستیک باید به تعداد پارامترها (parameters) نسبت به تعداد افراد (subjects) در مدل توجه کرد. مدل رگرسیون لجستیک زمانی که تعداد پارامترها نسبت به تعداد افراد کوچک باشد ترجیح داده می‌شود. ثانیاً زمانی که همسان‌سازی فردی صورت گرفته باشد باید مدل رگرسیون لجستیک شرطی استفاده شود و اگر بصورت گروهی همسان سازی صورت گرفته یا اصلا همسان‌سازی صورت نگرفته از مدل رگرسیون لجستیک استفاده شود. (۱۸) چرا که در مواردی که همسان‌سازی فردی وجود دارد با تعداد زیادی از طبقات و به تبع آن با حجم نمونه کم در طبقات روبه رو هستیم و با افزایش حجم نمونه تعداد پارامترها زیاد می‌شود برآورد حاصل از برازش مدل رگرسیون لجستیک استاندارد دارای اریبی خواهد بود. از طرفی برآورد پارامتر مربوط به طبقات در مطالعات همسان-سازی شده اهمیت چندانی ندارد لذا با بکارگیری مدل رگرسیون لجستیک شرطی بجای مدل رگرسیون لجستیک استاندارد با حذف پارامتر طبقات (پارامترهای مزاحم) برآورد نااریبی از پارامترهای مهم که ارتباط عوامل خطر و پاسخ را نشان می‌دهد خواهیم داشت (۱۴, ۱۵). نتیجه نهایی اینکه اگرچه در مطالعه حاضر مدل رگرسیون لجستیک شرطی برازش شده با توجه به ملاک‌های دقت و قدرت آن از جمله حساسیت و ویژگی، نسبت به لجستیک معمولی مدل مناسب‌تری برای پیش‌بینی پیامد است اما باید توجه داشت استفاده از هرکدام از رویکردهای رگرسیون لجستیک شرطی

[Downloaded from zanko.muk.ac.ir on 2025-11-19]

یا لجستیک معمولی به خصوصیات داده‌های مورد بررسی وابسته است. اگر همسان‌سازی فردی صورت گرفته باشد باید از لجستیک شرطی استفاده نمود چرا که اگر از رویکر رگرسیون لجستیک معمولی بجای آن استفاده شود نسبت شانس (OR) برآورد شده بزرگنمایی (overestimate) می‌شود (۱۶, ۱۷).

References

1- Antonogeorgos G, Panagiotakos DB, Priftis KN, Tzonou A. Logistic regression and linear discriminant analyses in evaluating factors associated with asthma prevalence among 10-to 12-years-old children: Divergence and similarity of the two statistical methods. International journal of pediatrics. 2009;2009.

2- Fernando R, Fernando DN. Childhood poisoning in Sri Lanka. The Indian Journal of Pediatrics. 1997;64(4):457-60.

3- Rice DP, MacKenzie EJ, Jones A, Kaufman S, DeLissovoy G, Max W, et al. Cost of injury in the United States: a report to Congress. 1989.

4- Mansori M, Soori H, Farnaghi f, Sohaila K. Assessment Risk factors for unintentional childhood poisoning :a case-control study in Tehran. Journal of Safety Promotion and Injury Prevention. 2013;1(4):183-9

5- Aviva P, Sabin C. Medical statistics at a glance. Blackwell Science Ltd., Padstow, UK; 2000.

6- Langholz B, Goldstein L. Conditional logistic analysis of case-control studies with complex sampling. Biostatistics. 2001;2(1):63-84.

7- Breslow NE ,Day NE. Statistical methods in cancer research: International Agency for Research on Cancer Lyon; 1987.

8- Hanley JA, Miettinen OS. An ‘Unconditional-like’Structure for the Conditional Estimator of Odds Ratio from 2× 2 Tables. Biometrical journal. 2006;48(1):23-34

9- Sedehi M, Mehrabi YE, Kazemnejad A, Hadaegh F. Comparison of artificial neural network, logistic regression and discriminant analysis methods in prediction of metabolic syndrome. 2009;11(6):638-46.

10- Long JD, Loeber R, Farrington DP. Marginal and random intercepts models for longitudinal binary data with examples from criminology. Multivariate behavioral research. 2009;44(1):28-58.

11- Peng C-YJ, Lee KL, Ingersoll GM. An introduction to logistic regression analysis and reporting. The Journal of Educational Research. 2002;96(1):3-14.

12- Worth AP, Cronin MT. The use of discriminant analysis, logistic regression and classification tree analysis in the development of classification models for human health effects. Journal of Molecular Structure: Theochem . 2003;622(1):97-111.

13- Bagley SC, White H, Golomb BA. Logistic regression in the medical literature:: Standards for use and reporting, with particular attention to one medical domain. Journal of clinical epidemiology. 2001;54(10):979-85.

14- Kleinbaum DG, Klein M. Analysis of Matched Data Using Logistic Regression: Springer; 2010.

15- McFadden D. Conditional logit analysis of qualitative choice behavior. 1973.

16- Rahman M, Fukui T. Conditional versus unconditional logistic regression in the medical literature. Journal of clinical epidemiology. 2003;56(1):101-2.

17- Campos-Filho N, Franco EL. A microcomputer program for multiple logistic regression by unconditional and conditional maximum likelihood methods. American Journal of Epidemiology. 1989;129(2):439-44

18- Fritz W. Statistical Methods in Cancer Research. Vol. 1 The Analysis of Case-Control Studies. Herausgegeben von NE BRESLOW, NE DAY und W. DAVIS. WHO-IARC Scientific Publications No. 32. 338 Seiten. International Agency for Research on Cancer ,Lyon 1980. Preis: 50,-sfrs; 30.00 US$. Food/Nahrung. 1982;26(3):329.

[Downloaded from zanko.muk.ac.ir on 2025-11-19]

Original paper

Compare the estimated odds ratios from logistic regression and conditional logistic regression in the case-control study determination risk factors for unintentional childhood poisoning of children in Tehran

Mansori K [1], Mansori Hanis SH[2*], Soori H[1], Khosravi Shadmani F[3], Moradpure F[4], Rajabi A [4], Gholami A[4]

1- Department of Epidemiology, Faculty of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

2- Student Research Committee, Kurdistan University of Medical Sciences, sanandaj, Iran; (Corresponding author). Email: mansorishiva@gmail.com

3- Department of Epidemiology, Faculty of Public Health, Kermanshah University of Medical Sciences, Kermanshah, Iran.

4- Department of Epidemiology, Faculty of Public Health ,Iran University of Medical Sciences, Tehran, Iran.

Abstract

Background and Aim: Identifying risk factors affecting on incidence poisoning is a fundamental measure to prevent poisoning in the community, In the meantime, find an equation for determination effects of factors and severity association of them has great importance. Purpose of this study compare the performance of two models ordinary logistic regression and conditional logistic regression in determination risk factors for unintentional childhood poisoning.

Materials and Methods: In the present study was used of data a case-control study that to determination risk factors for unintentional childhood poisoning in Tehran was carried out. After collecting the relevant data into the software stata and to determine the risk factors on incidence poisoning of two conditional logistic regression model and ordinary logistic regression was used. Odds ratios (OR) with confidence interval (CI) 95% for crude and adjusted with sensitivity, specificity, and area under the ROC curve were estimated for each of the proposed models, then they were compared.

Results: In multiple conditional logistic regression model addiction in the family, previous poisoning, maternal occupation and inaccessibility of poisoning products have shown a significant impact on incidence of poisoning. But In multiple logistic regression model, addition to the items listed, mental -mental illness in the family showed a significant relationship with outcome (poisoning). The AUC of ROC for conditional logistic regression model %95.82 (CI: 95% 0.94-0.97) with sensitivity and specificity 92.9% and 84.2% respectively and for logistic regression model %89.75 (CI: 95% 0.86-0.93) with sensitivity and specificity 83.6% and 83.1% respectively was obtained.

Conclusion: Although confidence interval associated with conditional logistic regression model compared with logistic regression is narrow and on the other hand, higher under curve area of ROC is a good indication of the overall precision of this model in the diagnosis of poisoning. But should be noted that use of each of these models depends on the feature of data (matched or unmatched).

Keywords: Childhood Poisoning, Case-Control, Conditional Logistic Regression, Logistic Regression

[Downloaded from zanko.muk.ac.ir on 2025-11-19]