

بررسی تاثیر پلی مورفیسم‌ها و سایر عوامل خطر بر سطح کلسترول تام خون در طول زمان با استفاده از مدل اثرات تصادفی منطقی: مطالعه قند و لیپید تهران

پروین سربخش^۱ (Ph.D)، یدالله محرابی^{۲*} (Ph.D)، مریم‌السادات دانشپور^۳ (Ph.D)، فرید زابری^۴ (Ph.D)، مهشید نامداری^۵ (Ph.D)
۱- دانشگاه علوم پزشکی تبریز، دانشکده بهداشت، گروه آمار و اپیدمیولوژی
۲- دانشگاه علوم پزشکی شهید بهشتی، دانشکده بهداشت، گروه اپیدمیولوژی
۳- دانشگاه علوم پزشکی شهید بهشتی، پژوهشکده علوم غدد درون‌ریز، مرکز تحقیقات سلولی و مولکولی، گروه ژنتیک ملکولی
۴- دانشگاه علوم پزشکی شهید بهشتی، مرکز تحقیقات پروتئومیکس، گروه آمار زیستی
۵- دانشگاه علوم پزشکی شهید بهشتی، دانشکده دندان‌پزشکی، گروه آمار زیستی

چکیده

سابقه و هدف: تقابل‌های ژنتیکی از عوامل موثر بر سطح کلسترول هستند. این برهم‌کنش‌ها می‌توانند وابسته به زمان بوده و لذا بررسی آن‌ها در طول زمان نیز حائز اهمیت باشد. با توجه به این که تاکنون مطالعه‌ای به ارزیابی طولی این برهم‌کنش‌ها نپرداخته است، در مطالعه حاضر، تقابل‌های ژنتیکی موثر بر کلسترول در طول زمان با استفاده از مدل پیشنهادی اثرات تصادفی منطقی بررسی شدند.

مواد و روش‌ها: مدل اثرات تصادفی منطقی در قالب مدل‌های اثرات تصادفی و استفاده از مقدار منفی درست‌نمایی به عنوان تابع امتیاز، جهت تحلیل تقابل‌ها در داده‌های طولی با پاسخ کمی پیشنهاد شد و با استفاده از آن داده‌های طولی ۳۲۹ نفر از شرکت‌کنندگان مطالعه طولی آینده‌نگر قند و لیپید تهران جهت شناسایی اثرات متقابل موثر بر کلسترول در طول زمان تحلیل گردید. اثرات متقابل با استفاده از الگوریتم Annealing شناسایی و شدت اثر آن‌ها و نیز مقدار جزء تصادفی مدل با روش ماکزیمم درست‌نمایی برآورد شد.

یافته‌ها: ترکیب "مردان یا افرادی با دور کمر و فشارخون نرمال و ژنوتایپ GG برای ApoAIV" باعث کاهش معنی‌دار ۱۹/۸ واحدی (۲۵/۶۹ و ۱۳/۹۰) CI (%۹۵) کلسترول گردید. هم‌چنین تقابل "تری‌گلیسرید بالا یا ال ال e2 در ApoE" باعث افزایش معنی‌دار کلسترول به میزان ۱۶/۱ واحد ((۲۰/۵۵ و ۱۱/۶۴) CI (%۹۵) شد. کلسترول در فاز ۱ مطالعه نسبت به فازهای دیگر به طور معنی‌داری (۲۱/۴ واحد (۲۴/۴۴ و ۱۸/۳۵) CI (%۹۵) بالاتر بود. مقدار میانگین عرض از مبدا تصادفی مدل برابر ۱۶۵/۸۲ و انحراف معیار جزء اثر تصادفی برابر ۳۰/۸۹ برآورد شد ($p < 0.001$). نتیجه‌گیری: برخی از پلی مورفیسم‌ها و عوامل خطر دارای تاثیرات تقابلی بر کلسترول در طول زمان بودند. استفاده از روش آماری پیشنهادی توانست این برهم‌کنش‌ها را شناسایی کند.

واژه‌های کلیدی: کلسترول، رگرسیون منطقی، داده‌های طولی، مدل اثرات تصادفی

مقدمه

ابتلا به این بیماری‌ها شده است. در مطالعات مختلفی نشان داده شده است که مقادیر بالای کلسترول تام شانس ابتلا به بیماری‌های قلبی عروقی را بالا می‌برد [۲، ۱] لذا شناسایی

شیوع بالای بیماری‌های قلبی عروقی در دنیا زمینه‌ساز انجام مطالعاتی جهت بررسی و شناسایی عوامل تاثیرگذار بر

عوض می‌شود یا خیر. روش رگرسیون منطقی نیز به عنوان یکی از روش‌های تحلیل اثر عوامل ژنتیکی بر روی فنوتایپ طولی کم‌تر استفاده شده است و تنها مطالعه‌ای که اخیراً برای فنوتایپ دو حالتی طولی با رگرسیون منطقی انجام شده رگرسیون منطقی انتقال می‌باشد [۱۷] ولی برای تحلیل داده‌های ژنتیکی با فنوتایپ کمی طولی تا به حال از رگرسیون منطقی (Logic Regression) استفاده نشده است.

وجود داده‌های ژنتیکی با فنوتایپ کمی طولی و وجود اثرات متقابل تاثیرگذار بین این داده‌ها انگیزه‌ای برای انجام تحقیق حاضر بود. هدف این مطالعه ارائه مدل آماری جدید "مدل اثرات تصادفی منطقی" بر مبنای مدل رگرسیون منطقی و استفاده از آن جهت بررسی اثر پلی‌مورفیسم‌ها و سایر عوامل خطر بر روی سطح کلسترول تام خون در طول زمان در مطالعه کوهورت آینده‌نگر قند و لیپید تهران است به طوری که در تحلیل‌های انجام شده هم طولی بودن داده‌ها و هم اثرات متقابل بین متغیرها لحاظ می‌شود.

مواد و روش‌ها

نوع مطالعه و افراد مورد بررسی. افراد مورد بررسی در این طرح از میان افراد شرکت‌کننده در مطالعه کوهورت آینده‌نگر قند و لیپید تهران که دارای معیارهای ورود زیر بودند، انتخاب شدند: ایرانی‌الاصل بودن، سن ۳ سال به بالا، سطح سرمی تری‌گلیسرید کم‌تر از ۴۰۰ mg/dl، بودن در رژیم غذایی معمول به مدت ۲ هفته قبل از خون‌گیری، باردار نبودن زنان، عدم مصرف داروهای موثر بر متابولیسم لیپیدها، نداشتن بیماری‌هایی که منجر به دیس لیپیدمی ثانویه شود، عدم وجود فاکتورهایی مانند انسداد وریدی، سکته مغزی، عفونت حاد، تروما.

مطالعه قند و لیپید تهران یک مطالعه آینده‌نگر است که روی یک نمونه از جمعیت منطقه ۱۳ تهران انجام شده و هدف آن تعیین شیوع بیماری‌های غیر واگیر و ترویج سبک زندگی سالم در این جمعیت می‌باشد. این مطالعه در ۴ فاز با فاصله متوسط ۳ ساله انجام شده است. در این تحقیق داده‌های ۳ فاز

عوامل موثر بر سطح این چربی می‌تواند در کنترل این بیماری‌ها موثر باشد. تغییرات ژنتیکی یا پلی‌مورفیسم‌های تک‌نوکلئیدی (Single nucleoid polymorphism) (SNP) در تنظیم میزان کلسترول تام نقش مهمی دارند [۳-۱۱]. بنابراین بررسی اثرات ژنتیکی موثر بر سطح این چربی نیز حائز اهمیت است.

مطالعات نشان می‌دهد که در بروز بیماری‌های پیچیده (Complex disease) ژن‌های متعددی با اثرات اصلی و انفرادی ضعیف ولی اثرات تقابلی شدید دخالت دارند بنابراین بررسی SNPها به صورت تکی صرفاً بخشی از تغییرات ژنتیکی مرتبط با فنوتایپ مورد بررسی را بیان می‌کند و باید اثرات تقابلی ژن‌ها با هم‌دیگر نیز بررسی شود [۱۲]. لذا زمانی که اثر تعدادی SNP روی فنوتایپ خاصی بررسی می‌شود به جای بررسی اثرات انفرادی آن‌ها باید از روش‌هایی استفاده شود که ترکیب و مجموعه‌هایی از SNPها را انتخاب و اثر آن بر فنوتایپ مدنظر ارزیابی می‌کند [۱۳]. روش‌های متعددی برای یافتن مجموعه‌های ژنی (Gene-Set) وجود دارد از جمله روش GRASS که در آن اطلاعات داخل یک ژن با استفاده از مولفه‌های اصلی خلاصه شده و مولفه‌های به دست آمده در مدل رگرسیونی استفاده می‌شوند [۱۴] و یا روش‌های مبتنی بر مدل‌سازی ساختار هم‌بستگی و خلاصه کردن SNPها در یک ژن با استفاده از توزیع چندمتغیره بیزی [۱۵].

روش دیگری که اخیراً برای یافتن مجموعه‌ای تاثیرگذار از SNPها بر روی فنوتایپ مدنظر استفاده می‌شود به طوری که برهم‌کنش بین SNPها هم در این مجموعه لحاظ شده باشد، رگرسیون منطقی است که نسبت به روش‌های دیگری برتری نسبی دارد [۱۶].

از طرفی در مطالعات ژنتیکی انجام گرفته به تحلیل ارتباط مجموعه‌ای از ژن‌ها با فنوتایپی که به صورت طولی در طول زمان ثبت شده، کم‌تر پرداخته شده است در حالی که بررسی اثرات اصلی و متقابل پلی‌مورفیسم‌ها بر روی فنوتایپ مورد بررسی در طول زمان نشان خواهد داد که آیا اثر تغییرات ژنی و برهم‌کنش بین آن‌ها بر فنوتایپ مورد بررسی در طول زمان

اول مطالعه مورد بررسی قرار می‌گیرد: فاز اول که مطالعه مقطعی بود از اسفند ۱۳۷۷ تا شهریور ۱۳۸۰ به طول انجامید؛ سپس افراد وارد فاز ۲ مطالعه شدند که از مهر ۱۳۸۰ شروع شد و در شهریور ۸۴ به پایان رسید و پس از آن فاز سوم اجرا شد که از سال ۸۴ تا اسفند ۸۶ ادامه پیدا کرد اطلاعات کامل مربوط به این مطالعه در منبع مربوطه آمده است [۱۹،۱۸].

از افراد واجد شرایط شرکت‌کننده در فاز ۳ مطالعه قند و لیپید تهران، تعداد ۳۲۹ نفر بالای ۲۰ سال بودند که داده‌های ژنتیکی برای آن‌ها ثبت شده بود و در فازهای ۱ و ۲ مطالعه TLGS هم شرکت داشتند و هیچ داده گمشده‌ای در متغیرهای مدنظر تحقیق حاضر نداشتند. لذا داده‌های طولی مربوط به ۳۲۹ نفر جهت بررسی ارتباط بین پلی مورفیسم‌ها و سایر عوامل خطر با سطح کلسترول تام خون در طول زمان با استفاده از مدل پیشنهادی مورد بررسی قرار گرفت.

متغیرهای مورد بررسی در تحقیق. اطلاعات مربوط به سن، جنس، مصرف سیگار و مصرف دارو با کمک پرسش‌نامه ثبت گردید. داده‌های مربوط به فشار خون و دور کمر اندازه‌گیری گردید. میزان کلسترول تام و تری‌گلیسرید سرم و قند خون ناشتا به وسیله کیست‌های تجاری (پارس‌آزمون-ایران) اندازه‌گیری شد. اثر متغیرهای چاقی شکمی (اندازه دور کمر ≤ 95 سانتی‌متر) [۲۰]، تری‌گلیسرید ≤ 150 میلی‌گرم بر دسی‌لیتر، فشار خون $\leq 130/85$ میلی‌متر جیوه یا درمان برای پرفشاری خون، گلوکز ناشتای پلاسما ≤ 110 میلی‌گرم بر دسی‌لیتر یا مصرف داروی پایین آورنده قند خون که مطابق تعریف مولفه‌های سندروم متابولیک بر اساس معیار ATP III تعریف شده بودند [۲۱] و سیگار نیز به عنوان سایر عوامل خطر موثر بر سطح کلسترول تام خون مورد بررسی قرار گرفت. افرادی که در زمان جمع‌آوری داده‌ها به صورت روزانه یا گه‌گاه سیگار مصرف می‌کردند، به عنوان سیگاری در نظر گرفته شدند.

پلی مورفیسم‌های XbaI در ژن ApoB، SstI در ژن ApoCIII، MspI در ژن ApoA1M1، XagI در ژن

ABCA1، AluI در ژن SRB1، پلی مورفیسم ژن ApoE، پلی مورفیسم ژن ApoA1M2 و پلی مورفیسم ژن ApoAIV که تصور می‌شود با سطح کلسترول تام خود در ارتباط باشند، [۲۲،۱۰-۳] بررسی شدند. هر پلی مورفیسم به صورت متغیر تصادفی X با مقادیر ۰، ۱ و ۲ در نظر گرفته شد (به عنوان مثال، نوکلئوتید AA، AG/GA و GG، به ترتیب با ۰ و ۱ و ۲ کدگذاری شدند). سپس این متغیر را به دو متغیر دوحالتی با عنوان‌های ژن غالب (dominant) (XD) و ژن مغلوب (recessive) (XR) تبدیل شد. متغیر ژن غالب به این صورت تعریف می‌شود: $XD=1$ اگر $X \geq 1$ باشد و $XD=0$ اگر $X=0$. متغیر ژن مغلوب نیز به این صورت تعریف می‌شود: $XR=1$ اگر $X=2$ باشد و $XR=0$ اگر $X \leq 1$ باشد. به این ترتیب، تعداد $2 \times p$ پیش‌بینی‌کننده دو حالتی از P تا SNP به دست می‌آید [۲۳]. نحوه دو حالتی کردن SNP‌ها به عنوان پیش‌پردازی برای داده‌های ژنتیکی در جدول ۱ ارائه شده است.

تجزیه و تحلیل آماری. رگرسیون منطقی یک روش رگرسیونی و کلاس‌بندی تعمیم یافته‌ای است که در آن متغیرهای جدید پیش‌بینی‌کننده به صورت ترکیبات منطقی (Logic combinations) ترکیباتی بر مبنای منطق بولی و عمل‌کردهای "و" و "یا" منطقی از متغیرهای دو حالتی اولیه ساخته و وارد مدل می‌شوند تا ارتباط و تقابل بین متغیرهای مستقل در قالب این ترکیب‌های منطقی ظاهر شود. جستجو برای یافتن بهترین ترکیبات منطقی از متغیرهای دو حالتی پیش‌بین با استفاده از الگوریتم جستجوی Simulated Annealing و بر مبنای کمینه کردن تابع امتیاز (Score function) مدل انجام می‌شود [۲۳-۲۵].

جدول ۱. نحوه دو حالتی کردن داده‌های ژنتیکی برای تحلیل آن‌ها با مدل

رگرسیون منطقی تصادفی

متغیر	حالت ۰	حالت ۱
ApoA1M1 غالب	+/+	-/- یا +/-
ApoA1M1 مغلوب	+/+ یا +/-	-/-

مدل، فرض بر این است که اثرات تصادفی b_i از فردی به فرد دیگر به طور مستقل و بر حسب یک توزیع مشترک تغییر می‌کنند. این توزیع معمولاً نرمال با میانگین صفر و واریانس σ_b^2 در نظر گرفته می‌شود. وابسته به مولفه‌های واریانس نامعلوم و پارامترهای هم‌بستگی بوده و نشانگر شدت پراکندگی و ناهمگنی بین افراد است. اثرات تصادفی b_i را می‌توان با عرض از مبدا مدل ادغام کرد و به صورت جز تصادفی یا عرض از مبدا تصادفی در مدل نشان داد $(\beta_{0i} = \beta_0 + b_i)$ که دارای تابع توزیع نرمال و مستقل از هم با میانگینی برابر عرض از مبدا جامعه و پارامتر واریانس نامعلوم \sum است.

$$\beta_{0i} \sim N(\beta_0 \text{ population}, \sigma_b^2)$$

β نیز بردار پارامترهای ضرایب رگرسیونی است. برنامه مربوط به مدل اثرات تصادفی منطقی و برآورد پارامترهای این مدل در محیط برنامه‌نویسی فورترن نوشته و تجزیه و تحلیل داده‌ها با استفاده از برنامه نوشته شده در محیط نرم‌افزار (2.14.1) R [۲۶] انجام شد.

جهت تجزیه و تحلیل داده‌ها با مدل پیشنهادی، مدل اثرات تصادفی منطقی که فضای جستجوی الگوریتم Annealing جهت یافتن بهترین ترکیبات منطقی شامل داده‌های پلی‌مورفیسم‌ها، متغیرهای سن، جنسیت، مصرف سیگار و فاز زمانی انجام مطالعه، تری‌گلیسرید بالا، فشار خون بالا، دور کمر بالا و قند خون ناشتا بالا بود، به داده‌های طولی سطح کلسترول تام برازش یافت. تعداد ترکیبات منطقی برای مدل برازش‌یافته ۳ و تعداد متغیرهای مشمول در این ترکیبات ۷ متغیر در نظر گرفته شد. برای برآورد شدت ناهمگنی بین افراد از برآورد انحراف معیار جمله اثر تصادفی استفاده شد. مقایسه مشخصات جمعیت‌شناختی شرکت‌کنندگان در سه فاز مطالعه نیز با استفاده از آنالیز اندازه‌های تکراری برای متغیرهای کمی و آزمون ککران برای متغیرهای کیفی انجام شد.

نتایج

-/یا+/-	+/+	ApoA1M2 غالب
-/-	+/-یا+/+	ApoA1M2 مغلوب
+/-یا+/+	-/-	APOB غالب
+/+	-/یا+/-	APOB مغلوب
GT یا TT	GG	ApoAIV غالب
TT	GG یا GT	ApoAIV مغلوب
GG یا CG	CC	ApoCIII غالب
GG	CG یا CC	ApoCIII مغلوب
AA یا GA	GG	ABCA1 غالب
AA	GA یا GG	ABCA1 مغلوب
AA یا GA	GG	SRB1 غالب
AA	GA یا GG	SRB1 مغلوب
e 4 یا 2	e 3	ApoE

در این مطالعه، "مدل اثرات تصادفی منطقی" در قالب مدل‌های اثرات تصادفی (Random Effects Model) و استفاده از مقدار منفی درست‌نمایی مدل (Negative likelihood) به عنوان تابع امتیاز، جهت توسعه رگرسیون منطقی برای تحلیل داده‌های طولی کمی پیشنهاد شد که در آن برای لحاظ کردن هم‌بستگی بین مشاهدات طولی از ساختار اثر تصادفی استفاده می‌شود. در مدل اثرات تصادفی منطقی که برای تحلیل داده‌های طولی کمی ارائه می‌شود، فرض بر این است که زیر مجموعه‌ای از ضرایب رگرسیون با توزیعی از یک فرد به فرد دیگر تغییر می‌کند. اثر تصادفی قادر است بسیاری از عوامل ایجاد ناهمگنی که قابل اندازه‌گیری نیستند را بیان کند. در حالت کلی، مدل اثرات تصادفی منطقی برای زامین مشاهده از فرد i ام به شکل زیر تعریف شد:

$$Y_{ij} = \beta_0 + L_{ij}\beta + b_i + e_{ij}, b_i \sim N(0, \sigma_b^2)$$

که در آن بردار L_{ij} ها بهترین ترکیبات منطقی از متغیرهای پیش‌بینی‌کننده دوحالتی هستند که توسط الگوریتم جستجوی Annealing یافت می‌شود طوری که تابع امتیاز مدل کم‌ترین و برازش مدل بهترین شود. بنابراین الگوریتم جستجوی Annealing ترکیبات منطقی را جستجو کرد که طبق آماره منفی درست‌نمایی، کم‌ترین امتیاز و در نتیجه بهترین برازش را در مدل اثرات تصادفی منطقی داشتند. در این

آمار توصیفی مربوط به ۳۲۹ نفر از شرکت‌کنندگان ۳ فاز متوالی مطالعه TLGS شامل متغیرهای جمعیت‌شناختی و هم‌چنین شاخص‌های خونی و فراوانی ال‌ها و ژنوتایپ‌های پلی مورفیسم ژن‌های مورد بررسی در جدول ۲ آمده است. مطابق این جدول، مشاهده شد که سطح کلسترول خون در فاز ۱ بیش‌ترین مقدار را داشته و در فاز ۲ با کاهش حدود ۱۷ واحدی به ۱۹۳ میلی‌لیتر رسیده است که این مقدار در فاز ۳ تغییر اندکی داشته و حدود ۱۹۴ میلی‌گرم در دسی‌لیتر بود. به نظر عوامل خطر تری‌گلیسرید بالا، قند خون بالا و مصرف سیگار روند نسبتاً ثابتی را در طی دوره پیگیری افراد مورد بررسی داشتند در حالی که عوامل خطر سن و دور کم‌ر بالا روند افزایشی و عامل خطر فشار خون روند کاهشی را در این دوره نشان داد.

فراوانی ال‌لی و ژنوتایپ پلی مورفیسم‌های مورد بررسی نیز در جدول ۳ آمده است. طبق این جدول، ژنوتایپ ++ در پلی مورفیسم Apo AIMI بیش‌ترین فراوانی (۹۱/۲٪) و ژنوتایپ TT در پلی مورفیسم Apo AIV کم‌ترین فراوانی (۰/۳٪) را داشتند. داده‌های این جدول تاییدکننده تبعیت توزیع ژنوتیپ‌ها از تعادل هاردی و واینبرگ است.

یافته‌های حاصل از برازش مدل اثرات تصادفی منطقی به داده‌ها که در جدول ۴ گزارش شده است، نشان می‌دهد که اولین ترکیب منطقی (اثر متقابل) بر سطح کلسترول خون در طول زمان تاثیر معنی‌دار داشت به این مفهوم که "مردان یا افرادی که دارای دور کم‌ر و فشار خون نرمال بوده و ژنوتایپ مربوط به ژن ApoAIV آن‌ها به صورت GG است، ۱۹/۸ واحد سطح کلسترول خونشان پایین‌تر از بقیه افراد فاقد این ترکیب منطقی بود. داشتن این ترکیب منطقی با فاصله اطمینان ۹۵٪ (۲۵/۶۹ و ۱۳/۹۰) در کاهش سطح کلسترول خون معنی‌دار بود ($p < 0.001$).

طبق دومین ترکیب منطقی (اثر متقابل) یافت شده در مدل اثرات تصادفی منطقی توسط الگوریتم Annealing، "داشتن تری‌گلیسرید بالا یا ال‌e2 در ژن ApoE" باعث افزایش سطح کلسترول خون به میزان ۱۶/۱ میلی‌گرم در دسی‌لیتر شد. این ترکیب نیز با فاصله اطمینان ۹۵٪ (۲۰/۵۵ و ۱۱/۶۴) بر سطح کلسترول تاثیر معنی‌دار داشت ($p < 0.001$).

جدول ۲: مشخصات جمعیت‌شناختی و شاخص‌های خونی افراد شرکت‌کننده در ۳ فاز متوالی از مطالعه TLGS

متغیر	فاز مطالعه		
	فاز ۱ (تعداد=۳۲۹ نفر)	فاز ۲ (تعداد=۳۲۹ نفر)	فاز ۳ (تعداد=۳۲۹ نفر)
*سن (سال)	۴۷/۴۶ ± ۱۵/۶۲	۴۴/۸۴ ± ۱۵/۷۱	۴۱/۰۹ ± ۱۵/۸۲
*سطح کلسترول تام (mg/dl)	۱۹۴/۹۴ ± ۴۰/۹۲	۱۹۳/۶۲ ± ۴۲/۴۵	۲۱۱/۳۱ ± ۴۸/۸۰
مصرف سیگار*	۲۷ (۸/۲)	۳۰ (۹/۱)	۲۷ (۸/۲)
دور کم‌ر بالا*	۱۵۱ (۴۵/۹)	۱۳۴ (۴۰/۷)	۹۰ (۲۷/۴)
قند خون بالا*	۴۱ (۱۲/۵)	۳۹ (۱۱/۹)	۳۹ (۱۱/۹)
فشار خون بالا*	۸۲ (۲۴/۹)	۹۷ (۲۹/۵)	۱۱۱ (۳۳/۷)
تری‌گلیسرید بالا*	۱۴۱ (۴۲/۹)	۱۴۰ (۴۲/۶)	۱۳۹ (۴۲/۲)
جنسیت (زن)†	۲۰۲ (۶۱/۴)	۲۰۲ (۶۱/۴)	۲۰۲ (۶۱/۴)

مقادیر به صورت (درصد) تعداد برای متغیرهای کیفی و انحراف معیار ± میانگین برای متغیرهای کمی گزارش شده‌اند. متغیرهایی که در طول زمان قابلیت تغییر دارند با علامت (*) و متغیرهایی که در طول زمان ثابت هستند با علامت (†) مشخص شده‌اند.

جدول ۳: فراوانی اللی و ژنوتایپ پلی مورفیسم‌های مورد بررسی مرتبط با سطح HDL

الل ApoE	e2	e3	e4
ژنوتایپ ApoA1M1	+/+	-/-	+/-
ژنوتایپ ApoA1M2	+/+	-/-	+/-
ژنوتایپ ApoB	+/+	-/-	+/-
ژنوتایپ ApoAIV	TT	GG	GT
ژنوتایپ ApoCIII	CC	GG	CG
ژنوتایپ ABCA1	GG	AA	GA
ژنوتایپ SRB1	GG	AA	GA

داده‌ها به صورت فراوانی (درصد) گزارش شده‌اند.

جدول ۴: یافته‌های حاصل از برازش مدل اثرات تصادفی منطقی با ۳ ترکیب منطقی و ۷ متغیر برای بررسی تاثیر پلی مورفیسم‌ها و سایر عوامل خطر بر سطح کلسترول تام در طول زمان

ترکیب منطقی یافت شده	برآورد ضریب رگرسیون	خطای استاندارد برآورد ضریب	فاصله اطمینان ۹۵٪ برای ضریب برآورد شده	p-value
(داشتن دور کمر نرمال و فشار خون نرمال و ApoAIV=GG یا جنسیت مرد)	-۱۹/۸	۳/۰۱	(-۱۳/۹۰ و -۲۵/۶۹)	<۰/۰۰۱
(داشتن تری گلیسرید بالا یا الل e2 در ژن ApoE)	۱۶/۱	۲/۲۷	(۱۱/۶۴ و ۲۰/۵۵)	<۰/۰۰۱
فاز ۱ نسبت به فازهای دیگر	۲۱/۴	۱/۵۵	(۱۸/۳۵ و ۲۴/۴۴)	<۰/۰۰۱
اثر تصادفی	۱۶۵/۸۲	۳۰/۸۹	(۱۰۵/۵۶ و ۲۲۶/۳۸)	<۰/۰۰۱

* فضای جستجوی الگوریتم Annealing برای یافتن بهترین ترکیبات منطقی جهت برازش مدل اثرات تصادفی منطقی شامل پلی مورفیسم‌ها و متغیرهای سن، جنسیت، مصرف سیگار و فاز اجرای مطالعه، فشار خون بالا، تری گلیسرید بالا، قند خون ناشتا بالا و دور کمر بالا بود.

اهمیت فاکتور زمان و لزوم در نظر گرفتن این فاکتور در تحلیل‌ها دلالت دارد. اثر این ترکیب بر سطح کلسترول با فاصله اطمینان ۹۵٪ (۲۴/۴۴ و ۱۸/۳۵) معنی‌دار بود ($p < 0/001$). در مدل برازش یافته متغیر سن هم که به صورت کمی در مدل تعدیل شده بود با ضریب ۰/۷۳ و فاصله اطمینان ۹۵٪ برابر (۰/۹۶ و ۰/۴۹) دارای اثر افزایشی معنی‌دار بر سطح کلسترول بود ($p < 0/001$) به این معنی که افزایش یک

طبق سومین ترکیب منطقی (اثر متقابل) یافت شده در مدل اثرات تصادفی منطقی، بودن در فاز ۱ مطالعه تاثیر معنی‌داری بر سطح کلسترول داشت. یعنی فاز ۱ مطالعه نسبت به فازهای دیگر دارای ۲۱/۴ واحد کلسترول بالاتری بود. این یافته‌ها با یافته‌های حاصل از جدول توصیفی نیز هم‌خوانی دارد چرا که طبق آمارهای توصیفی مشاهده شد که فاز ۱ بیش‌ترین میانگین کلسترول را داشت. یافت شدن این ترکیب منطقی به

سال سن باعث افزایش ۰/۷۳ میلی‌گرم در دسی‌لیتر کلسترول خون می‌شود.

انحراف معیار جزء اثر تصادفی مدل برابر ۳۰/۸۹ برآورد شد که اختلافش با صفر معنی‌دار بود ($p < 0.001$). میزان بالای انحراف معیار جمله اثر تصادفی (نسبت به صفر) و معنی‌دار بودن آن نشانگر ناهمگنی افراد با یک‌دیگر و تاثیرگذار بودن ویژگی‌های فردی بر سطح کلسترول خون است. مقدار جز تصادفی مدل یعنی پارامتر میانگین عرض از مبدا تصادفی مدل نیز برابر ۱۶۵/۸۲ برآورد شد.

بحث و نتیجه‌گیری

نتایج مطالعه نشان داد که افراد با ترکیب خاصی از جنسیت، دور کمر، فشار خون و پلی مورفیسم ApoAIV دارای سطح کلسترول پایین‌تری نسبت به بقیه افراد فاقد این ترکیب بودند. هم‌چنین افراد با تری‌گلیسرید بالا یا ال‌ال e2 در ژن ApoE سطح کلسترول بالاتری نسبت به افراد با تری‌گلیسرید نرمال و ال‌الی غیر از e2 در ژن ApoE داشتند. به طور متوسط نیز سطح کلسترول در فاز ۱ مطالعه نسبت به فازهای دیگر بالاتر بود.

داده‌های مورد بررسی در این مطالعه از نوع داده‌های ژنتیکی طولی هم‌بسته بودند که متغیر پاسخ یعنی سطح کلسترول تام خون در دوره پیگیری ۳ بار اندازه‌گیری شده و شرط استقلال در داده‌ها برقرار نبود لذا برای لحاظ کردن این هم‌بستگی و عدم استقلال در داده‌ها نیاز به استفاده از مدل‌های مناسب جهت تحلیل داده‌های طولی مانند مدل اثرات تصادفی بود. از طرفی با توجه به این‌که در مطالعات ژنتیکی برای بررسی ارتباط بین عوامل محیطی و ژنوتایپ‌ها با فنوتایپ مدنظر به دلایل متعدد از جمله شناسایی تقابل و تعامل بین متغیرهای ژنتیکی و محیطی تمایل داریم تا از بعد داده‌ها کاسته و اطلاعات آن‌ها در قالب چند مجموعه و ترکیب محدود خلاصه کنیم لذا با ترکیب دو روش مدل اثرات تصادفی (جهت در نظر گرفتن هم‌بستگی مشاهدات پاسخ) و مدل رگرسیون منطقی (روشی برای تلخیص داده‌ها و شناسایی کردن اثرات

متقابل) مدل جدیدی با عنوان "مدل اثرات تصادفی منطقی" پیشنهاد شد و از آن جهت بررسی اثر پلی مورفیسم‌ها و سایر عوامل خطر ساز بر سطح کلسترول استفاده شد.

این ترکیب‌های خاص یا اثرات متقابل با روش مدل اثرات تصادفی منطقی شناسایی شد و روش‌های موجود توانایی شناسایی و تعیین شدت اثر این تقابل‌ها را نداشتند.

مقایسه نتایج مدل اثرات تصادفی منطقی با مدل مرسوم (مدل اثرات تصادفی با اثرات اصلی همه متغیرها) نشان داد که مدل اثرات تصادفی منطقی برآزش بهتری در مقایسه با مدل با اثرات تصادفی صرفاً با اثرات اصلی داشت به طوری که معیار اطلاع آکائیکه (AIC) که برای مقایسه برآزش مدل‌ها استفاده می‌شود (مدل با AIC کم‌تر برآزش بهتری دارد) برای مدل اثرات تصادفی منطقی برآزش یافته برابر ۷۷۱۰/۸۴ و برای مدل مرسوم برابر ۷۷۲۳/۶۳ به دست آمد و این برآزش بهتر موید این مطلب بود که خلاصه کردن اطلاعات تعداد زیادی متغیر در تعداد محدودی مجموعه و ترکیب مناسب نه تنها باعث از دست رفتن اطلاعات داده‌ها نشد بلکه انتخاب مجموعه و ترکیب مناسبی که هم‌بستگی و تقابل بین متغیرها را نیز لحاظ کرده است باعث بهبود برآزش مدل شده و بیش و تفسیر بیولوژیکی جدیدی نیز ارائه داد که می‌تواند باعث شناسایی تقابل‌ها و شرایطی که احتمالاً دلیل ایجاد بیماری است، گردد. علاوه بر موارد ذکر شده، به علت حضور بخش تصادفی در مدل رگرسیون طولی، تاثیر عوامل فردی ناشناخته غیر قابل اندازه‌گیری بر کلسترول خون نیز مورد بررسی قرار گرفت که با توجه به معناداری جمله تصادفی می‌توان نتیجه گرفت که علاوه بر تمامی عوامل بررسی شده، مشخصات فردی شامل عوامل ناشناخته نیز موثر بر سطح کلسترول خون باشد که در نظر گرفتن این جزء تصادفی باعث بهبود برآزش مدل شد.

ترکیبات منطقی یافت شده به عنوان مجموعه‌های تلخیص شده از تمام متغیرهای مورد بررسی، تقابل و برهمکنش بین پلی مورفیسم‌های ژنی و عوامل محیطی را نشان داد. این تقابل‌های معنی‌دار به این معنی هستند که ممکن است ترکیب و

در پژوهش حاضر، از داده‌های طرح قند و لیپید تهران که توسط پژوهشکده‌ی علوم غدد درون‌ریز و متابولیسم دانشگاه علوم پزشکی شهید بهشتی اجرا شده، استفاده گردید. پژوهشگران از تمام کسانی که در طراحی و جمع‌آوری داده‌های TLGS مشارکت داشتند نهایت قدردانی را به عمل می‌آورند. هم‌چنین برای انجام محاسبات این پژوهش از امکانات گروه علوم کامپیوتر پژوهشگاه دانش‌های بنیادی (IPM) انجام شده است لذا نویسندگان مراتب قدرشناسی خود را نسبت به مسئولین و کارکنان مرکز IPM ابراز می‌دارند.

منابع

- [1] Tanabe N, Iso H, Okada K, Nakamura Y, Harada A, Ohashi Y, et al. Serum total and non-high-density lipoprotein cholesterol and the risk prediction of cardiovascular events - the JALS-ECC. *Circ J* 2010; 74: 1346-1356.
- [2] NIPPON DATA80 Research Group. Risk assessment chart for death from cardiovascular disease based on a 19-year follow-up study of a Japanese representative population. *Circ J* 2006; 70: 1249-1255.
- [3] Daneshpour MS, Faam B, Hedayati M, Eshraghi P, Azizi F. ApoB (XbaI) polymorphism and lipid variation in Teharnian population. *Eur J Lipid Sci Technol* 2011; 113: 436-440.
- [4] Brown CM, Rea TJ, Hamon SC, Hixson JE, Boerwinkle E, Clark AG, Sing CF. The contribution of individual and pairwise combinations of SNPs in the APOA1 and APOC3 genes to interindividual HDL-C variability. *J Mol Med (Berl)* 2006; 84: 561-572.
- [5] McCarthy JJ, Lehner T, Reeves C, Moliterno DJ, Newby LK, Rogers WJ, Topol EJ. Association of genetic variants in the HDL receptor, SR-B1, with abnormal lipids in women with coronary artery disease. *J Med Genet* 2003; 40: 453-458.
- [6] Frikke-Schmidt R. Context-dependent and invariant associations between APOE genotype and levels of lipoproteins and risk of ischemic heart disease: a review. *Scand J Clin Lab Invest Suppl* 2000; 233: 3-25.
- [7] Daneshpour MS, Faam B, Mansournia MA, Hedayati M, Halalkhor S, Mesbah-Namin SA, et al. Haplotype analysis of Apo A1-CIII-AIV gene cluster and lipids level: Tehran lipid and glucose study. *Endocrine* 2012; 41: 103-110.
- [8] Daneshpour M, Zarkesh M, Hedayati M, Mesbah SN, Halalkhor S, Faam B, Azizi F. The G360t polymorphism in the APO AIV gene and its association with combined HDL/LDL cholesterol phenotype: Tehran lipid and glucose study. *Int J Endoc Metab* 2010; 8: 32-38.
- [9] Daneshpour MS, Hedayati M, Eshraghi P, Azizi F. Association of Apo E gene polymorphism with HDL level in Teheranian population. *Eur J Lipid Sci Technol* 2010; 112: 810-816.
- [10] Halalkhor S, Mesbah-Namin SA, Daneshpour MS, Hedayati M, Azizi F. Association of ATP-binding cassette transporter-A1 polymorphism with apolipoprotein AI level in Teheranian population. *J Genet* 2011; 90: 129-132.
- [11] Zarkesh M, Daneshpour MS, Faam B, Hedayati M, Azizi F. Is there any association of apolipoprotein E gene polymorphism with obesity status and lipid profiles? *Tehran Lipid and Glucose Study (TLGS)*. *Gene* 2012; 509: 282-285.
- [12] Deelen J, Uh HW, Monajemi R, van Heemst D, Thijssen PE, Böhringer S, van den Akker EB, et al. Gene set analysis of GWAS data for human longevity highlights the relevance of the insulin/IGF-1 signaling and telomere maintenance pathways. *Age (Dordr)* 2013; 35: 235-249.

چینش خاصی از عوامل خطر بر سطح کلسترول تاثیرگذار باشد نه خود عوامل خطر به تنهایی و بعضی از عوامل خطر یا پلی‌مورفیسم‌ها در حضور هم می‌توانند خطر افزایش سطح کلسترول یا هر فنوتایپ دیگری را افزایش دهند لذا در مطالعات ژنتیکی توجه به اثرات متقابل و شناسایی آن‌ها حایز اهمیت است.

متغیرهای ظاهر شده در ترکیبات منطقی از جمله جنسیت، سن، تری‌گلیسرید، فشار خون، پلی‌مورفیسم ApoAIV، ApoE و دور کمر همگی طبق مطالعات قبلی بر سطح کلسترول خون معنی‌دار بودند [۲۷،۹،۸] ولی در اغلب این مطالعات تاثیر این متغیرها یا به صورت انفرادی و نه تقابلی بررسی شده بود [۲۷،۹،۸] و یا به صورت تقابلی ولی برای داده‌های مقطعی ارزیابی شده بود [۲۸]. در این مطالعه به بررسی تاثیر پلی‌مورفیسم‌ها و سایر عوامل خطر بر سطح کلسترول در طول زمان با مدل پیشنهادی اثرات تصادفی منطقی و رویکرد تلخیص داده‌ها و ساختن مجموعه و ترکیب ژنی پرداخته شد.

پیشینه بودن ساختار داده‌های ژنی و وجود انواع هم‌بستگی و تقابل و ارتباطات ژن‌ها با هم باعث شده است که مدل‌های آماری معمول در بررسی اثر ژنوتایپ‌ها بر فنوتایپ توان کم‌تری داشته باشند و لذا هر روزه مدل‌های آماری پیشرفته‌ای با توان و کارایی بیشتر برای تحلیل این داده‌ها معرفی می‌شود. در این مقاله نیز مدل اثرات تصادفی منطقی به عنوان روشی برای تحلیل داده‌های ژنی طولی توسط محققین پیشنهاد و از آن جهت تحلیل داده‌های TLGS استفاده شد.

محدودیت‌های پژوهش

به دلیل مشکل داده‌های گمشده و عدم امکان جانی آن‌ها، تنها داده‌های ۳ فاز از مطالعه TLGS قابل تحلیل و استفاده بودند. هم‌چنین تعداد پلی‌مورفیسم‌های ارزیابی شده برای افراد در مقایسه با پلی‌مورفیسم‌های مرتبط با کلسترول کم بود.

تشکر و قدردانی

- [21] Third report of the national cholesterol education program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III) final report. *Circulation* 2002; 106: 3143-3421.
- [22] Zarkesh M, Daneshpour MS, Faam B, Hedayati M, Azizi F. Is there any association of apolipoprotein E gene polymorphism with obesity status and lipid profiles? *Tehran Lipid and Glucose Study (TLGS)*. *Gene* 2012; 509: 282-285.
- [23] Ruczinski I, Kooperberg C, LeBlanc M. Logic Regression. *Journal of Computational and Graphical Statistics* 2003; 12: 475-511.
- [24] Mehrabi Y, Khadem-Maboudi A, Hadaegh F, Sarbakhsh P. Prediction of diabetes using logic regression. *Iran J Endoc Metabol* 12: 16-24. (Persian).
- [25] Sarbakhsh P, Mehrabi y, Khadem Maboudi AA, Hadaegh F. Logic regression and its application in predicting diseases. *Andish-ye-amari* 2011; 16: 34-46.
- [26] R Development Core Team (2011), R: ISBN: 3-900051-07-0. A language and environment for statistical computing. vienna, austria : the R foundation for statistical computing. Available online at <http://www.R-project.org/>
- [27] Lewington S, Whitlock G, Clarke R, Sherliker P, Emberson J, Halsey J, et al. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *Lancet* 2007; 370: 1829-1839.
- [28] Sarbakhsh P, Mehrabi Y, Daneshpour MS, Zayeri F, Zarkesh M. Logic regression analysis of association of gene polymorphisms with low HDL: Tehran lipid and glucose study. *Gene* 2013; 513: 278-281.
- [13] Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown K, et al. A versatile gene-based test for genome-wide association studies. *Am J Hum Genet* 2010; 87: 139-145.
- [14] Chen LS, Hutter CM, Potter JD, Liu Y, Prentice RL, Peters U, Hsu L. Insights into colon cancer etiology via a regularized approach to gene set analysis of GWAS data. *Am J Hum Genet* 2010; 86: 860-871.
- [15] Tsonaka R, van der Helm-van Mil AH, Houwing-Duistermaat JJ. A two-stage mixed-effects model approach for gene-set analyses in candidate gene studies. *Stat Med* 2012; 31: 1190-1202.
- [16] Schwender H, Ruczinski I. Logic regression and its extensions. *Adv Genet* 2010; 72: 25-45.
- [17] Sarbakhsh P, Mehrabi Y, Daneshpour M, Zayeri F, Zarkesh M, Azizi F. Studying the interaction effects of gene polymorphisms on low level of HDL over time using transition logic regression: Tehran lipid and glucose study. *Iran J Endoc Metabol* 2013; 15: 352-359.
- [18] Azizi F, Rahmani M, Emami H, Mirmiran P, Hajipour R, Madjid M, et al. Cardiovascular risk factors in an Iranian urban population: Tehran lipid and glucose study (Phase 1). *Soz Praventivmed* 2002; 47: 408-426.
- [19] Azizi F, Mirmiran P, Azadbakht L. Predictors of cardiovascular risk factors in Tehranian adolescents: Tehran Lipid and Glucose Study. *Int J Vitam Nutr Res* 2004; 74: 307-312.
- [20] Azizi F, Khalili D, Aghajani H, Esteghamati A, Hosseinpanah F, Delavari A, et al. Appropriate waist circumference cut-off points among Iranian adults: the first report of the Iranian national committee of obesity. *Arch Iran Med* 2010; 13: 243-244.

Association of polymorphisms and other risk factors with cholesterol level over time using logic random effect model: Tehran Lipid and Glucose Study

Parvin Sarbakhsh (Ph.D)¹, Yadollah Mehrabi (Ph.D)^{*2}, Farid zayeri (Ph.D)³, Maryam Daneshpour (Ph.D)⁴, Mahshid Namdari (Ph.D)⁵

1 – Dept. of Statistics and Epidemiology, School of Health, Tabriz University of Medical Sciences, Tabriz, Iran

2 – Dept. of Epidemiology, School of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran

3 - Proteomics Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

4 - Cellular and Molecular Endocrine Research Center, Research Institute for Endocrine Sciences

5 - Dental School, Shahid Beheshti University of Medical Sciences, Tehran, Iran

(Received 25 Sep 2013; Accepted: 6 Jul 2014)

Introduction: changes in cholesterol level can be a product of Genetic interactions of some factors. These interactions are time-dependent. Since a time-dependent genetic interaction study related to blood cholesterol levels has not been performed until now, this study is proposed to evaluate the time-dependent association of SNPs and cholesterol by using a logic random effect model of assessment.

Materials and Methods: logic random effect model was proposed in the frame of random effects models by using negative likelihood as the score function to analyze the interactions related to longitudinal quantitative responses. Data of 329 participants in TLGS prospective cohort study was analyzed using proposed model to identify effective interactions related to cholesterol. Annealing algorithm was used to search interaction effects. Coefficient of these interactions and random effect were estimated using maximum likelihood method.

Results: The results for male or other subjects with GG genotype for ApoAIV who had normal blood pressure and waist circumference had shown significant lower cholesterol levels (19.8 mg/dl, CI 95%: 13.9, 25.69) ($p < 0.001$) than other subjects without this combination. Also, having (high triglyceride or allele e2 for ApoE) was significantly associated with an increasing effect on cholesterol ($\beta = 16.1$, CI 95%: 11.64, 20.55). The level of the cholesterol in phase 1 of study was 21.4 mg/dl (CI 95%: 18.35, 24.44) more than other phases ($p < 0.001$). The model's random effect was estimated 165.82 with standard deviation 30.89 ($p < 0.001$).

Conclusion: We identified some interactions among SNPs and other covariates related to cholesterol level using logic random effect model.

Keywords: Cholesterol, Logic regression, Longitudinal data, Random effect model

* Corresponding author. Fax: +98 21 22416264; Tel +98 21 22432500
mehrabi@sbmu.ac.ir

How to cite this article:

Sarbakhsh P, Mehrabi Y, Daneshpour M, Zayeri F, Namdari M. Association of polymorphisms and other risk factors with cholesterol level over time using logic random effect model: Tehran Lipid and Glucose Study. *koomesh*. 2015; 16 (2) :193-201

URL http://koomeshjournal.semums.ac.ir/browse.php?a_code=A-10-2151-1&slc_lang=fa&sid=1