

تحلیل پیوستگی رگرسیونی دوسطحی چندمتغیره و کاربرد آن در پیوستگی ژنتیکی صفات میزان HDL-C، تری گلیسیرید و چاقی تنه‌ای در بین ۹۱ خانواده ایرانی مبتلا به سندرم متابولیک

- مهدی اکبرزاده^۱ (M.Sc)، حمید علوی مجد^{۲*} (Ph.D)، مریم‌السادات دانشپور^۳ (Ph.D)، یدالله محرابی^۴ (Ph.D)، فریدون عزیزی^۵ (Ph.D)
- ۱- دانشگاه علوم پزشکی سبزوار، دانشکده بهداشت، گروه آمار زیستی
 - ۲- دانشگاه علوم پزشکی شهید بهشتی، دانشکده پیراپزشکی، گروه آمار زیستی
 - ۳- دانشگاه علوم پزشکی شهید بهشتی، مرکز تحقیقات پیشگیری و درمان چاقی، پژوهشکده علوم غدد درون‌ریز و متابولیسم
 - ۴- دانشگاه علوم پزشکی شهید بهشتی، دانشکده بهداشت، گروه اپیدمیولوژی

چکیده

سابقه و هدف: سندرم متابولیک یک فنوتیپ مرکب است که ۳۲ درصد از افراد ایرانی به آن مبتلا می‌باشند. این تحقیق به بررسی ژنتیکی نواحی کروموزومی اثرگذار بر میزان HDL-C، میزان تری گلیسیرید و دور کمر به کمک میکروستلایت‌ها و روش رگرسیونی چندمتغیره the جهت شناخت نواحی ژنی موثر در سندرم متابولیک در جمعیت ایرانی می‌پردازد.

مواد و روش‌ها: افراد مورد بررسی در این مطالعه از بین شرکت‌کنندگان در مطالعه‌ی قند و لیپید تهران انتخاب شده‌اند که شامل ۹۱ خانواده ایرانی بوده و انتخاب آن‌ها به این صورت بوده است که حداقل یک نفر از اعضای این خانوارها مبتلا به سندرم متابولیک (طبق معیار ATP III) و حداقل دو نفر از اعضای آن‌ها دچار کاهش HDL-C بودند. تکثیر ۱۲ قطعه مختلف در ۴ ناحیه کروموزومی افراد مربوط به ۹۱ خانواده ایرانی (۴۹۳ نفر)، با استفاده از تکنیک Fragment Analysis انجام گردید. تحلیل پیوستگی ژنتیکی صفات میزان HDL-C، میزان تری گلیسیرید و دور کمر با سه مدل تک متغیره، سه مدل دو متغیره و یک مدل سه متغیره انجام گردید.

یافته‌ها: ۹۱ خانوار کامل شامل ۴۹۳ نفر (۲۳۴ مرد و ۲۵۹ زن) بود. در مدل‌های تک متغیره، پیوستگی میزان HDL-C با مارکر D11S1998 و هم‌چنین پیوستگی تری گلیسیرید با مارکرهای D11S1934 و D12S1632 معنی‌دار شده است ($p < 0.05$). در مدل‌های دو متغیره، پیوستگی توام دو صفت میزان HDL-C و تری گلیسیرید، هم‌چنین میزان HDL-C و چاقی تنه‌ای با مارکر D11S1998 معنی‌دار شده است و پیوستگی توام چاقی تنه‌ای و تری گلیسیرید با مارکرهای D8S1743 و D11S934 معنی‌دار شده است. در مدل سه متغیره نیز پیوستگی توام میزان HDL-C، چاقی تنه‌ای و تری گلیسیرید با مارکر D8S1743 معنی‌دار شده است.

نتیجه‌گیری: نتایج نشان می‌دهند که، وقتی صفتی در بین مدل‌های متفاوت، مشترک است، مارکرهایی که با آن صفت پیوستگی نشان می‌دهند نیز اشتراک دارند. لذا با مشاهده این اشتراک‌ها در بین مارکرهای موجود، با روش‌های چند متغیره، با اطمینان بیش‌تری می‌توان فرض پیوستگی بین صفات مذکور با این مکان‌های ژنتیکی را پذیرفت. یافتن این نواحی ژنی توسط روش مذکور می‌تواند راه‌کار مناسبی برای طراحی مطالعات بعدی به منظور یافتن ژن‌های مستعدکننده سندرم متابولیک باشد.

واژه‌های کلیدی: سندرم متابولیک، میکروستلایت، روش رگرسیونی the چند متغیره

مطالعه ژنتیک بیماری‌ها فراهم کرده است. اما تکمیل این طرح انتهای این فرایند نخواهد بود، و در واقع نقطه شروع است.

مقدمه

تکمیل طرح پروژه ژنوم انسانی منابع بسیاری را برای

چاپ شد [۳]. در آوریل سال ۲۰۰۹ نیز برنامه مربوط به استفاده از این روش در نرم افزار S.A.G.E ارائه شد [۴]. سندرم متابولیک یک فنوتیپ مرکب است و با افزایش ریسک ابتلا به بیماری های قلبی - عروقی همراه می باشد که جهت شناسایی آن از بررسی تغییر عوامل مختلفی مانند افزایش تری گلیسرید، فشار خون، قند خون ناشتا، دور کمر و کاهش HDL-C استفاده نمود [۵]. ۳۲ درصد از افراد ایرانی مبتلا به این سندرم می باشند [۶]. فراوان ترین فنوتیپ این سندرم کاهش HDL-C است. برای شناسایی دلیل کاهش HDL-C نیاز به شناسایی ژن های اثرگذار بر متابولیسم آن می باشد.

این تحقیق به بررسی ژنتیکی نواحی کروموزومی اثرگذار بر میزان HDL-C، تری گلیسرید و دور کمر، به کمک میکروستلایت ها و روش رگرسیونی iHE چند متغیره جهت شناخت نواحی ژنی موثر در سندرم متابولیک در جمعیت ایرانی می پردازد و به علاوه نتایج حاصل از مدل های مختلف نیز با یکدیگر مقایسه می شود.

مواد و روش ها

جامعه مورد بررسی و حجم نمونه. افراد مورد مطالعه از میان شرکت کنندگان طرح مطالعه قند و لیپید تهران انتخاب شدند که در محدوده سنی ۶۹-۳ سالگی و در دو گروه مذکر و مونث قرار گرفته بودند. بدین ترتیب که خانواده هایی که حداقل یک نفر از افراد آن علائم سندرم متابولیک و حداقل دو نفر از اعضای آن دچار کاهش HDL-C بودند، انتخاب شدند [۷]. معیار انتخاب افراد بر اساس تعریف ATP III بوده است [۸]. در این مطالعه حجم نمونه با استفاده از فرمول مطرح شده توسط Oliver Hädike و همکاران به دست آمده است [۹]. برای محاسبه حجم نمونه با این روش، برنامه power.HE نرم افزار R اجرا شده است [۱۰]. با در نظر گرفتن سطح معنی داری ۰/۰۵، توان آزمون ۸۰ درصد، وراثت پذیری ۰/۴۰۳ و همبستگی مانده های مدل فاکتور برابر ۰/۳، حداقل حجم نمونه مورد نیاز با استفاده از این روش ۲۴۶ sibpair

آنچه هم اکنون باقی مانده است، تعیین ژن های تعیین کننده صفات مورد نظر، تغییرات آن ها در جامعه و این که این ژن ها چطور با سایر عوامل اثرات متقابل دارند، می باشد. لذا تعیین ژن های مربوط به بیماری ها همواره از اهداف مهم در مطالعات اپیدمیولوژی ژنتیک بوده است. یکی از مراحل مهم در این مسیر، تحلیل پیوستگی ژنتیکی می باشد. مسئله مهم در تحلیل پیوستگی، مرکب بودن برخی فنوتیپ ها است. ممکن است چند صفت هم بسته منجر به بروز یک بیماری پیچیده شوند، و هیچ یک از صفات به تنهایی نمی توانند تغییرات و خصوصیات بیماری مورد نظر را بیان کند. برای مثال در یک بیماری چند عاملی مانند بیماری های قلبی - عروقی یا سندرم متابولیک، عوامل مختلفی مانند عوامل محیطی، اجتماعی، تغییرات ژنتیکی و فاکتورهای بیوشیمیایی دخیل می باشند. لذا در این حالت بهتر است از تحلیل پیوستگی چند متغیره استفاده کرد.

همه روش های تحلیل پیوستگی چند متغیره که تا سال ۲۰۰۶ ارائه شده است، بر اساس تعمیم روش مولفه های واریانس هستند، و تحلیل پیوستگی را با استفاده از روش (Likelihood ratio test LRT) برای یافتن مولفه واریانس مکان صفت کمی (Quantitative trait loci QTL) معنی دار انجام می دهد [۱]. واضح است که، در این روش ها، اعتبار روش مولفه های واریانس به برقراری فرض نرمال چند متغیره وابسته است. از آنجایی که دست یابی به پیش فرض نرمال چند متغیره در عمل کار ساده ای نیست، نتایج حاصل از روش مذکور چندان معتبر نبوده و شک برانگیز است. استفاده از یک روش استوار نسبت به انحراف از فرض نرمال چند متغیره، به خصوص زمانی که با صفاتی سروکار داریم که به وضوح دارای توزیع غیرنرمال هستند، می تواند توان آزمون را بهبود بخشد. چنین روشی اخیراً تحت عنوان روش (Two) the (level Haseman-Elston) چند متغیره معرفی شده که در واقع به عنوان تعمیمی از روش oHE (Original Haseman-Elston) است [۲]. در سال ۲۰۰۷ اولین مقاله کاربردی در مورد استفاده از این روش در بررسی پیوستگی صفات فشار خون سیستولیک، دیاستولیک و شاخص توده بدنی یا BMI

فرآیند ترتیب این کار انجام می‌شود. در این مقاله به بررسی سه مدل تک متغیره، سه مدل دو متغیره و یک مدل سه متغیره خواهیم پرداخت. شایان ذکر است که در تمام تحلیل‌ها، مدل‌های رگرسیونی مورد نظر با متغیرهای سن و جنس تعدیل شده‌اند. هم‌چنین نتایج متغیرهای کمی به صورت میانگین (انحراف معیار) ارائه، و سطح معنی‌داری، ۰/۰۵ در نظر گرفته شده است. جهت محاسبه P-value از روش ترتیب، تعداد تکرارها را حداقل ۱۰۰۰ و حداکثر ۱۰۰۰۰۰ در نظر گرفته‌ایم. نرم‌افزارهای مورد استفاده. در تهیه این مقاله از برنامه power.HE در نرم‌افزار R ویرایش ۲،۱۰،۰ جهت تعیین حجم نمونه [۱۰]، از نرم‌افزارهای Excel ویرایش ۲۰۰۷، SPSS ویرایش ۱۶ و نرم‌افزار PowerMarker ویرایش ۳،۲۵ [۱۱] در آماده‌سازی داده‌ها، و در نهایت از نرم‌افزار S.A.G.E [۴] ویرایش ۶،۰،۱ جهت اجرای آنالیزهای رگرسیونی HE استفاده شده است.

نتایج

۹۱ خانوار، شامل ۴۹۳ نفر (۲۳۴ مرد و ۲۵۹ زن)، دارای اطلاعات کاملی از مارکرها، متغیر پاسخ و متغیرهای مستقل جهت انجام آنالیز پیوستگی بوده‌اند. از این میان ۲۳۱ نفر Founder، و ۲۶۲ نفر Non-founder بوده‌اند. هم‌چنین اطلاعات مربوط به روابط خانوادگی در بین خانواده‌ها در جدول ۱ آمده است.

جدول (۱): روابط خانوادگی در جمعیت مورد مطالعه

تعداد	نوع رابطه
۵۲۴	والد- فرزند
۷۱	خواهر- خواهری
۷۰	برادر- برادری
۱۱۵	برادر- خواهری
۰	برادر(خواهری) ناتنی
۴۶	پدر(مادر)بزرگ- نوه
۲۸	دایی (عمو، خاله، عمه)- خواهرزاده (برادرزاده)
۳	پسر (دختر)-خاله (عمو، عمه، دایی)

برآورد شد. تکثیر ۱۲ قطعه مختلف در ۴ ناحیه کروموزومی مربوط به این افراد با استفاده از تکنیک Fragment Analysis انجام گردید.

روش‌های آماری. استفاده از روش‌های رگرسیونی Haseman-Elston یا HE در تحلیل پیوستگی صفات کمی بسیار رایج است. در تمام تعمیم‌های روش HE به شجره نامه‌ها، این اشکال وجود دارد که نتوانسته‌اند تحلیل پیوستگی را به شجره نامه‌های کلی بسط دهند. روش tHE با در نظر گرفتن متغیرهای سطح فرد در مدل رگرسیونی، خطای مانده‌ای را کاهش داده و با در نظر گرفتن سطح شجره نامه نیز ناهمگنی داده‌ها را کنترل می‌کند، به عبارتی اثر متقابل محیط و ژنتیک را در مدل در نظر می‌گیرد. روش HE دو سطحی تحت چارچوب کلی رگرسیون چندسطحی و با استفاده از الگوریتم IGLS انجام می‌شود. در این مقاله از روش tHE چند متغیره جهت بررسی تحلیل پیوستگی ژنتیکی صفات میزان HDL-C، چاقی تنه ای و تری‌گلیسیرید استفاده شده است. مدل آماری این روش به صورت زیر است.

فرض کنید y_{ijk} مقدار صفت k ام ($k=1, 2, \dots, m$) برای عضو i ام ($i=1, 2, \dots, s_j$) در خانواده j ام ($j=1, 2, \dots, N$) باشد. فرض را بر این می‌گیریم که مقادیر صفت موردنظر با QTL موردنظر، اثرات مشترک پلی‌ژنیک، محیط و اثر تصادفی فرد، تعیین می‌شوند. برای سادگی کار فرض می‌کنیم که مدل به صورت افزایشی است، زیرا تعمیم این مدل به مدلی با اثر غالبیت ساده است. در مدل tHE میانگین نمونه‌ای در سطح اول رگرسیونی برآورد شده و پیوستگی را در سطح دوم رگرسیونی مورد بررسی قرار می‌دهد.

پارامترهای مدل رگرسیونی موردنظر عبارتند از:

$$\sigma_{pk}^2, \sigma_{pkt}^2, \sigma_{ckt}^2, \sigma_{ek}^2, \sigma_{ekt}^2, \sigma_{gk}^2, \sigma_{gkt}^2$$

برآوردهای این پارامترها و ماتریس واریانس-کوواریانس را می‌توان با استفاده از الگوریتم IGLS برآورد کرد، که تحت توزیع نرمال چند متغیره، این برآوردها همان برآوردهای حداکثر درست نمایی هستند. هم‌چنین محاسبه p-value در این روش ممکن است پیچیده باشد و با استفاده از

میانگین (انحراف معیار) HDL در افراد مورد بررسی برابر $111/01$ ($44/41$) mg/dl، تری‌گلیسیرید برابر $142/80$ ($86/53$) mg/dl و چاقی تنه ای برابر $15/61$ ($87/47$) cm بود.

در مدل‌های تک متغیره، پیوستگی میزان HDL-C با مارکر D11S1998 و همچنین پیوستگی تری‌گلیسیرید با مارکرهای D11S1934 و D12S1632 معنی‌دار شد و چاقی تنه ای با هیچ‌یک از مارکرهای موجود پیوستگی معنی‌داری نشان نداد.

در مدل‌های دو متغیره، پیوستگی توام دو صفت میزان HDL-C و تری‌گلیسیرید، همچنین میزان HDL-C و چاقی تنه‌ای با مارکر D11S1998، معنی‌دار شد و نیز پیوستگی توام چاقی تنه‌ای و تری‌گلیسیرید با مارکرهای D8S1743 و D11S934 نیز معنی‌دار شد. همچنین در مدل سه متغیره نیز پیوستگی توام میزان HDL-C، چاقی تنه‌ای و تری‌گلیسیرید با مارکر D8S1743 معنی‌دار شد. این نتایج در جدول ۲ آمده است.

جدول ۲. p-value های حاصل از روش رگرسیونی the

مارکر	نوع مدل						
	مدل سه متغیره	مدل دو متغیره			مدل تک متغیره		
		HDL/TG/waist	HDL/TG	HDL/waist	TG/waist	HDL	TG
D8S1132	۰/۶۹۷۹	۰/۴۳۲	۰/۶۶۵	۰/۵۳۳	۱/۰۰۰	۰/۱۶۶	۰/۱۴۶
D8S1779	۰/۶۰۷	۰/۵۰۵	۰/۶۹۱	۰/۵۴۱	۱/۰۰۰	۰/۰۵۷	۰/۱۳۱
D8S514	۰/۸۶۸	۰/۶۵۵	۰/۶۶۰	۱/۰۰۰	۰/۳۹۰	۱/۰۰۰	۰/۴۷۸
D8S1743	۰/۰۳۰*	۰/۶۲۴۶	۰/۲۸۰۸	۰/۰۳۵*	۰/۲۵۰	۰/۰۹۸	۱/۰۰۰
D11S1998	۰/۰۹۹	۰/۰۴۴*	۰/۰۲۹*	۰/۵۲۳	۰/۰۱۲*	۰/۱۳۲	۰/۳۶۰
D11S934	۰/۱۸۳	۰/۱۱۹	۰/۲۹۳	۰/۰۲۵*	۱/۰۰۰	۰/۰۱۵*	۰/۳۲۲
D11S1304	۰/۴۵۶	۰/۶۹۷	۰/۲۷۹	۰/۶۰۱	۱/۰۰۰	۰/۴۱۵	۰/۱۴۸
D12S1632	۰/۴۱۵	۰/۲۶۱	۰/۳۶۵	۱/۸۰	۰/۳۳۷	۰/۰۱۶*	۰/۳۴۲
D12S96	۰/۴۴۵	۰/۲۲۹	۰/۲۰۲	۰/۶۱۶	۰/۱۱۶	۰/۲۴۰	۰/۲۳۱
D12S329	۰/۳۷۴	۰/۲۶۳	۰/۲۳۸	۰/۳۱۷	۰/۱۵۱	۰/۱۳۴	۱/۰۰۰
D16S2624	۰/۳۷۳	۰/۲۳۱	۰/۲۵۱	۰/۶۱۶	۰/۲۱۰	۰/۱۰۸	۱/۰۰۰
D16S3096	۰/۴۶۹	۰/۴۴۵	۰/۲۵۱	۰/۴۱۴	۰/۱۹۵	۰/۱۹۳	۰/۰۶۵

* : p-value < ۰/۰۵

بحث و نتیجه‌گیری

این مقاله به بررسی ژنتیکی نواحی کروموزومی اثرگذار بر میزان HDL-C، تری‌گلیسیرید و دور کمر با استفاده از روش رگرسیونی the چندمتغیره برای شناخت نواحی ژنی موثر در سندرم متابولیک در جمعیت ایرانی پرداخت. با توجه به نتایج ارائه شده در جدول ۲ وقتی صفتی در بین مدل‌های متفاوت، مشترک است، مارکرهایی که با آن صفت پیوستگی نشان می‌دهند نیز تقریباً اشتراک دارند و پیام مشابه‌ای را

می‌رسانند. در مورد HDL-C این مارکر، مارکر D11S1998 می‌باشد. در مورد تری‌گلیسیرید نیز مارکرهای D8S1743 و D11S934 این‌گونه هستند. لذا با مشاهده این اشتراک‌ها در بین مارکرهای موجود، با روش‌های چند متغیره، با اطمینان بیش‌تری می‌توان فرض پیوستگی بین صفات مذکور با این مکان‌های ژنتیکی را پذیرفت. همچنین نتایج حاصل از این مقاله نشان داد که خروجی حاصل از مدل‌های تک متغیره و چند متغیره با یکدیگر متفاوت است و این خود نشان از اثر pleiotropic در بین صفات مذکور (در این مکان‌ها) می‌باشد.

HDL-C به ناحیه کروموزومی 12q14.1 و مارکر D12S334 رسیدند [۱۷]. در مطالعه دیگری که هورن بنجامین و همکاران در سال ۲۰۰۳ بر داده‌های حاصل از Genome wide scan از ۹۳۰ نفر داده‌های مطالعه قلب فارمینگهام، و با روش‌های مدل مینا انجام شد، در مورد نسبت TG/HDL-C به ناحیه کروموزومی ۲ رسیدند [۱۸]. هم‌چنین در مطالعه دیگری که در دسامبر سال ۲۰۰۹ توسط بیگر و همکاران انجام شد و شامل ۷۰۰ نفر با sib-pair ۱۰۵۲ از ۳۳۴ خانواده مربوط به مطالعه قلب فارمینگهام بود. این مطالعه با روش the تک متغیره و چند متغیره بر مولفه‌های سندرم متابولیک و با استفاده از دستور RELPAL در نرم‌افزار S.A.G.E انجام شده است که در تحلیل تک متغیره مربوط به HDL-C به ناحیه کروموزومی ۱۱ با $p\text{-value}=1.80 \times 10^{-4}$ و در آنالیز چند متغیره نیز به نواحی کروموزومی ۱۶ و ۱۷ رسیده‌اند [۱۹].

همان‌طور که مشاهده می‌شود، مطالعات اندکی تا سال ۲۰۱۰ میلادی در ارتباط با تحلیل پیوستگی سندرم متابولیک و فاکتورهای آن انجام شده است، که نتایج مطالعه حاضر در مورد نواحی کروموزومی HDL-C در حد زیادی با مطالعات مذکور هم‌خوانی دارد، و در مواردی که عدم مطابقت بین نتایج مطالعه حاضر و مطالعات مذکور مشاهده می‌شود را می‌توان به دلیل یکسان نبودن جوامع و نژادهای مورد بررسی دانست. هم‌چنین از آنجایی که روش the چند متغیره اخیراً ارائه شده، تنها در یکی از این مطالعات از روش the چند متغیره استفاده شده است و نتایج مدل تک متغیره مطالعه حاضر در مورد HDL-C با مطالعه مذکور مطابقت دارد [۱۹].

تشکر و قدردانی

بدین وسیله از پژوهشکده غدد درون‌ریز و متابولیسم دانشگاه علوم پزشکی شهید بهشتی به خاطر در اختیار قرار دادن داده‌های مربوط به مطالعه TLGS، از آقای دکتر Robert Elston به خاطر در اختیار قرار دادن نرم‌افزار ارزش‌مند S.A.G.E، و از داوران مجله کومش برای نظرات سازنده ایشان کمال تشکر و قدردانی را به عمل می‌آوریم.

در مطالعه‌ای که توسط المسی و همکاران در سال ۱۹۹۹ بر داده‌های حاصل از Genome wide scan از ۴۷۷ نفر (۱۰ خانواده بزرگ)، زیر نمونه‌ای از مطالعه قلب خانواده‌های مکزیکی سان آنتیو، و با روش‌های مدل مینا انجام شد، به این نتیجه رسیدند که دو مکان بر کروموزوم‌های ۸ و ۱۵ میزان HDL-C را کنترل می‌کنند. در این مطالعه بر روی کروموزوم‌های ۲، ۴، ۵، ۸، ۱۲ و ۱۵ کار شده بود [۱۲]. کورت و همکاران در سال ۲۰۰۰ ارتباط ناحیه D11S1998 با تغییرات میزان HDL-C را نشان داده‌اند. ناحیه مورد نظر در این مطالعه، ناحیه 11q23.3 بوده است که 1.7Mbp با ژن کدکننده کلاستر ژنی آپولیپوپروتئین‌های APOA1/C3/A4/A5 فاصله دارد [۱۳]. در مطالعه‌ای که توسط هیلاری و همکاران در سال ۲۰۰۱، بر روی ۱۸۲ سفیدپوست آمریکایی و ۱۹۸ سیاه‌پوست آمریکایی، انجام شد به این نتیجه رسیدند که در خانواده‌های آمریکایی میزان HDL-C (تعدیل شده با TG و تعدیل نشده) با مکان خاصی پیوستگی ژنتیکی ندارد و در خانواده‌های سفیدپوست میزان HDL-C در حالت تعدیل شده با TG دارای پیوستگی معنی‌داری با ناحیه کروموزوم ۵ و در حالتی که با متغیری تعدیل نشده با کروموزوم ۱ پیوستگی معنی‌داری نشان می‌دهد. تحلیل‌های آماری این مقاله با روش‌های مدل مینا انجام شده است [۱۴]. در مطالعه‌ای که توسط مهانی و همکاران در سال ۲۰۰۲ بر داده‌های حاصل از Genome wide scan از ۴۷۲ نفر داده‌های مطالعه قلب خانواده‌های مکزیکی سان آنتیو و با روش‌های مدل مینا انجام شد، در مورد میزان HDL-C به مارکر D16S518 رسیدند [۱۵]. در مطالعه دیگری که آریا و همکاران در سال ۲۰۰۳ بر داده‌های حاصل از Genome wide scan از ۱۷۰۲ نفر داده‌های مطالعه قلب فارمینگهام، و با روش‌های مدل مینا دو متغیره انجام شد، در مورد میزان HDL-C و BMI به مارکرهای D261334 و D6S1009 رسیدند [۱۶]. در مطالعه دیگری که بوس و همکاران در سال ۲۰۰۳ بر داده‌های حاصل از Genome wide scan از ۹۳۰ نفر داده‌های مطالعه قلب فارمینگهام، و با روش‌های مدل مینا انجام شد، در مورد میزان

منابع

- [11] K. Liu, S.V. Muse. PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics* 2004, 21: 2128-2129.
- [12] Almasy L, Hixson JE, Rainwater DL, Cole S, Williams JT, Mahaney MC, et al. Human pedigree-based quantitative-trait-locus mapping: Localization of two genes influencing HDL-cholesterol metabolism. *Am J Hum Genet* 1999; 64: 1686-1693.
- [13] Kort EN, Ballinger DG, Ding W, Hunt SC, Bowen BR, Abkevich V, et al. Evidence of linkage of familial hypoalphalipoproteinemia to a novel locus on chromosome 11q23. *Am J Hum Genet* 2000; 66: 1845-1856.
- [14] Coon H, Leppert MF, Eckfeldt JH, Oberman A, Myers RH, Peacock JM, et al. Genome-wide linkage analysis of lipids in the hypertension genetic epidemiology network blood pressure study. *Arterioscler thromb vasc biol* 2001; 21: 1969-1976.
- [15] Mahaney MC, Almasy L, Rainwater DL, VandeBerg JL, Cole SA, Hixson JE, et al. A quantitative trait locus on chromosome 16q influences variation in plasma HDL-C levels in mexican americans. *Arterioscler Thromb Vasc Biol* 2003; 23: 339-345.
- [16] Arya R, Lehman D, Hunt KJ, Schneider J, Almasy L, Blangero J, et al. Evidence for bivariate linkage of obesity and HDL-C levels in the framingham heart study. *BMC Genet* 2003; 4: S52.
- [17] Bosse Y, Chagnon YC, Despres JP, Rice T, Rao DC, Bouchard C, et al. Genome-wide linkage scan reveals multiple susceptibility loci influencing lipid and lipoprotein levels in the quebec family study. *J Lipid Res* 2004; 45: 419-446.
- [18] Horne BD, Malhotra A, Camp NJ. Comparison of linkage analysis methods for genome-wide scanning of extended pedigrees, with application to the TG/HDL-C ratio in the framingham heart study. *BMC Genet* 2003; 4: S93.
- [19] Baker AR, Goodloe RJ, Larkin EK, Baechle DJ, Song YE, Phillips LS, Gray-McGuire CL. Multivariate association analysis of the components of metabolic syndrome from the framingham heart study. *BMC Proc* 2009; 3: S42.
- [1] Allison DB, Thiel B, St. Jean P, Elston RC, Infante MC, Schork NJ. Multiple phenotype modeling in gene-mapping studies of quantitative traits: Power advantages. *Am J Hum Genet* 1998; 63: 1190-1201.
- [2] Wang T. Extensions of Haseman-Elston regression for linkage analysis. Cleveland, Ohio: CASE WESTERN UNIVERSITY; 2006.
- [3] Wang T, Elston RC. Regression-based multivariate linkage analysis with an application to blood pressure and body mass index. *Ann Hum Genet* 2007; 71: 96-106.
- [4] Elston RC. Statistical analysis for genetic epidemiology (S.A.G.E). 6.0.1 ed. Cleveland, Ohio: Department of Epidemiology and Biostatistics, Case Western Reserve University; April, 2009.
- [5] Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. *Lancet* 2005; 365: 1415-1428.
- [6] Azizi F, Salehi P, Etemadi A. and Zahedi-Asl S. Prevalence of metabolic syndrome in an urban population: Tehran lipid and glucose study. *Diabetes Res Clin Pract* 2003; 61: 29-37.
- [7] Azizi F, Rahmani M, Emami H, Madjid M. Tehran Lipid and glucose study: Rationale and design. *CVD prevention* 2000; 3: 242-247.
- [8] Expert Panel on Detection, Evaluation, and treatment of high blood cholesterol in adults. executive summary of the third report of the national cholesterol education program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults. *JAMA* 2001; 285: 2486-2497.
- [9] Hädike O, Pahlke F, Ziegler A. A general approach for sample size and power calculations based on the Haseman-Elston method. *Biom J* 2008; 50: 257-269.
- [10] Hädike O. Power.HE. 1.2 ed: Universität Lübeck - Institut für Medizinische Biometrie und Statistik; 2007.

Analyzing of multivariate two levels Haseman-Elston regression and its application in genetic linkage of HDL-C, triglycerides and waist in 91 Iranian families with metabolic syndrome

Mahdi Akbarzadeh (M.Sc)¹, Hamid Alavi Majd (Ph.D)^{*2}, Maryam Sadat Daneshpour (Ph.D)³, Yadolah Mehrabi (Ph.D)⁴, Freydoon Azizi (Ph.D)³

1 - Dept. of Biostatistics, Faculty of Public Health, Sabzevar University of Medical Sciences, Sabzevar, Iran

2 - Dept. of Biostatistics, Faculty of Paraclinical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

3 - Obesity Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

4 - Dept. of Epidemiology, Faculty of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran

(Received: 06 Jun 2010 Accepted: 05 Oct 2010)

Introduction: Metabolic syndrome is a complex trait and its prevalence is 32% in Iranian population. The present study was conducted to find chromosomal area locus of HDL-C, triglycerides and waist with microsatellites and multivariate two level Haseman-Elston regressions in Iranian families with metabolic syndrome.

Materials and Methods: 91 Iranian families (493 people) with at least one member with metabolic syndrome were selected from database of TLGS. We performed the Fragment Analysis technique to reproduce 12 different pieces from 4 chromosomal areas and to identify loci related to metabolic syndrome; both single and multi variable two level Haseman-Elston regression methods were used for traits of triglycerides, high-density lipoprotein and waist. We performed three single variable models, three double variable models and one triple variable model of these traits.

Results: 91 Iranian families included 493 people, 234 males and 259 females. In single variable models: genetic linkage of HDL-C was significant with D11S1998 marker; genetic linkage of triglycerides was significant with D11S1934 and D12S1632 markers. In double variable models genetic linkage of HDL-C and triglyceride, HDL-C and waist was significant with D11S1998 marker and the genetic linkage of HDL-C and triglyceride, triglyceride and waist was significant with D8S1743 and D11S934 marker. In triple variable model genetic linkage of HDL-C, triglyceride and waist was significant with D8S1743 marker.

Conclusion: These results showed when a trait is common in different models; the linked markers of them are also common. We concluded that the multivariate methods can detect linked loci of mixed disease better than single variable models and these results are useful for more future studies in Iranian population.

Key words: Metabolic Syndrome, HDL-C, Microsatellite, Two levels Haseman-Elston Regression

* Corresponding author: Fax: +98 21 22721150; Tel: +98 9121483687
alavimajd@gmail.com