MHJ
مجلات لومورا

المجلة متعددة التخصصات للرعاية الصحية

المجلد 1، العدد 6 (2025) · e-ISSN 3152-8775

مقالة أصلية · DOI: غير مسجل

قابلية انتقال المواقع الجينية لدهون الدم المكتشفة لدى ذوي الأصول الأوروبية إلى مجموعة سكانية عربية في الشرق الأوسط: دراسة تكرار عابرة للأصول معايَرة بالقوة الإحصائية، مع تقييم للدرجة متعددة الجينات من الإحصاءات الموجزة وتحديد دقيق للمتغيرات السببية باستخدام الإحصاءات الموجزة لتسلسل الجينوم الكامل في البنك الحيوي القطري واتحاد علم وراثة الدهون العالمي لعام 2021

فاطمة العوض[1,*]، هديل رشاش المطيري[2]، هيلاء الشلوي[3]، حيدر ميرغني[4]، محمد الحسن[5]، نجلاء المانع[6]، زهوة دهمي الشمري[7]

تاريخ النشر	تاريخ القبول	تاريخ التقديم
30 يونيو 2025	25 يونيو 2025	20 يونيو 2025

[1] كلية الطب، جامعة الملك فيصل، الأحساء، المملكة العربية السعودية

[2] جامعة القصيم في كلية عنيزة، المملكة العربية السعودية

[3] مستشفى القوات المسلحة بالقصيم، القصيم، المملكة العربية السعودية

[4] جامعة تبوك، تبوك، المملكة العربية السعودية

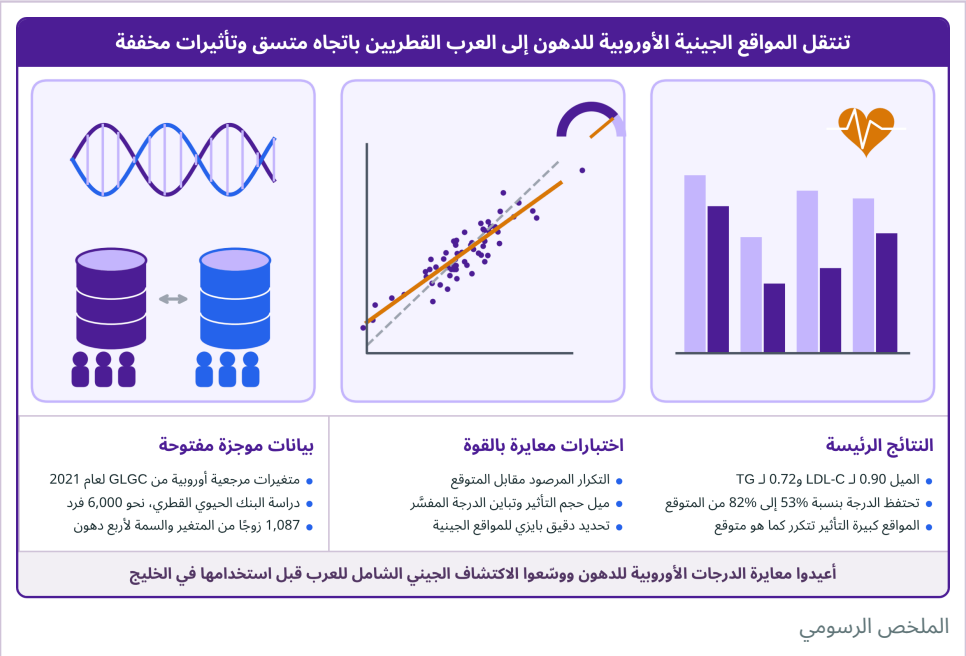
[5] جامعة الأمير سطام بن عبدالعزيز، الخرج، المملكة العربية السعودية

[6] جامعة الملك سعود بن عبدالعزيز للعلوم الصحية، جدة، المملكة العربية السعودية

[7] جامعة حائل، المملكة العربية السعودية

*** المراسلة:** فاطمة العوض، fatemalumora@gmail.com

الملخص



الملخص الرسومي

الخلفية: اكتُشفت معظم المتغيرات الجينية المرتبطة بالدهون في مجموعات من ذوي الأصول الأوروبية، ولا تزال المجموعات السكانية العربية ممثلة تمثيلًا ناقصًا. وقد أفادت دراسات خليجية سابقة بانخفاض معدل تكرار هذه المتغيرات، لكنها لم تقارن التكرار المرصود بالتكرار المتوقع بناءً على القوة الإحصائية للدراسات العربية الأصغر حجمًا. وقد قمنا بقياس مدى انتقال المواقع الجينية للدهون المكتشفة لدى ذوي الأصول الأوروبية إلى العرب القطريين.

الطرق: استخدمت دراسة الارتباط الجيني العابرة للأصول هذه الإحصاءات الموجزة المتاحة للعموم فقط. أُخذت المتغيرات المرجعية لدى ذوي الأصول الأوروبية لكوليسترول البروتين الدهني منخفض الكثافة (LDL-C) وكوليسترول البروتين الدهني مرتفع الكثافة (HDL-C) والدهون الثلاثية (TG) والكوليسترول الكلي (TC) من اتحاد علم وراثة الدهون العالمي لعام 2021 (n = 1,320,016)، ثم طوبقت مع دراسة الارتباط على مستوى الجينوم المبنية على تسلسل الجينوم الكامل في البنك الحيوي القطري لعام 2021 (n = 5,972 إلى 6,017). وقدّرنا لكل سمة توافق اتجاه التأثير، والتكرار الاسمي والتكرار بتصحيح Bonferroni مقارنةً بالقيم المتوقعة من القوة الإحصائية، وميل حجم التأثير الموزون بمقلوب التباين، وعدم التجانس بين الأصول، والتباين المفسَّر بدرجة موزونة بالتأثيرات الأوروبية. وأجرينا تحديدًا دقيقًا للمواقع القطرية ذات الدلالة على مستوى الجينوم المتداخلة مع الإشارات الأوروبية باستخدام عوامل Bayes التقريبية لـ Wakefield بافتراض وجود متغير سببي واحد. واتُّبعت إرشادات STREGA في كتابة التقرير.

النتائج: من بين 1,600 متغير مرجعي مفهرس، كانت 1,087 زوجًا من أزواج المتغير والسمة قابلة للتحليل. وتراوح توافق الاتجاه بين 69.5% (LDL-C) و75.1% (HDL-C)، وهو أدنى من القيم المتوقعة عند الانتقال الكامل (77.3% إلى 80.2%). وكان التكرار الاسمي أقل من المتوقع لـ LDL-C (29 مقابل 42.9؛ النسبة 0.68) وTG (23 مقابل 44.7؛ النسبة 0.51) وTC (36 مقابل 47.6؛ النسبة 0.76)، في حين طابق التكرار ذو الدلالة بتصحيح Bonferroni القيم المتوقعة (5 إلى 7 متغيرات لكل سمة). وكانت التأثيرات القطرية مخففة مقارنةً بالتأثيرات الأوروبية لـ TC (الميل 0.88، 95% CI 0.78–0.97) وHDL-C (0.78، 0.68–0.88) وTG (0.72، 0.64–0.81)؛ أما لـ LDL-C فبلغ الميل 0.90 (0.80–1.00) وامتد فاصل الثقة حتى 1. وفسّرت الدرجة الموزونة بالتأثيرات الأوروبية 6.46% من تباين LDL-C (النسبة إلى المتوقع 0.82، 95% CI 0.66–0.99)، و5.63% من تباين TC (0.77)، و4.13% من تباين TG (0.53)، و3.57% من تباين HDL-C (0.60). ومن بين 16 موقعًا قطريًا أحادي الإشارة، احتوت مجموعة الموثوقية 95% في الوسيط على 6 متغيرات، وانحصرت الإشارة في متغير واحد عند APOE rs7412 وGCKR rs1260326. واستُبعدت من التحديد الدقيق خمسة مواقع منخفضة التكرار متعددة الإشارات على الكروموسوم 11q23، ويتسق نمطها مع نمط فرداني قطري ممتد دون أن يثبته. وتعتمد تقديرات الدرجة على استقلال المتغيرات وليست تحققًا ببيانات فردية؛ ولم يستمر دليل تخفيف TC عند تقليم المسافة إلى 5 Mb.

الاستنتاجات: أظهرت مواقع الدهون المكتشفة لدى ذوي الأصول الأوروبية توافقًا في اتجاه الارتباط لدى العرب القطريين، مع انخفاض ميول أحجام التأثير لـ HDL-C وTG. وتعتمد تقديرات تباين الدرجة المستنتجة من الإحصاءات الموجزة على استقلال المتغيرات وتوافق مقاييس السمات؛ وهي ليست قياسًا متحققًا منه للأداء التنبئي الفردي. وتدعم النتائج الاكتشاف الجيني الشامل للعرب والتحقق المحلي قبل الاستخدام السريري.

الكلمات المفتاحية: دراسة الارتباط على مستوى الجينوم؛ الدهون؛ الدرجة متعددة الجينات؛ قابلية الانتقال؛ التحديد الدقيق للمتغيرات السببية؛ قطر؛ العرب

للاستشهاد: فاطمة العوض، هديل رشاش المطيري، هيلاء الشلوي، حيدر ميرغني، محمد الحسن، نجلاء المانع، وآخرون. قابلية انتقال المواقع الجينية لدهون الدم المكتشفة لدى ذوي الأصول الأوروبية إلى مجموعة سكانية عربية في الشرق الأوسط: دراسة تكرار عابرة للأصول معايَرة بالقوة الإحصائية، مع تقييم للدرجة متعددة الجينات من الإحصاءات الموجزة وتحديد دقيق للمتغيرات السببية باستخدام الإحصاءات الموجزة لتسلسل الجينوم الكامل في البنك الحيوي القطري واتحاد علم وراثة الدهون العالمي لعام 2021. المجلة متعددة التخصصات للرعاية الصحية. MHJ-2026-000374:(6)1;2025. © المؤلفون. منشور بنظام الوصول المفتوح بموجب رخصة المشاع الإبداعي نَسب المُصنَّف 4.0 الدولية (CC BY 4.0)، https://creativecommons.org/licenses/by/4.0/.

1. المقدمة

يُعد مرض القلب والأوعية الدموية التصلبي العصيدي السبب الرئيس للوفاة على مستوى العالم، ويُعد ارتفاع كوليسترول البروتين الدهني منخفض الكثافة (LDL-C) عامل خطر سببيًا [1,2]. وفي المملكة العربية السعودية ومنطقة الخليج عمومًا، يشيع اضطراب شحوم الدم منذ المراهقة، وتضعه الإرشادات الوطنية الآن في صميم الوقاية من أمراض القلب والأوعية الدموية [3,4]. ويمكن للمعلومات الجينية أن تدعم هذا التوجه، من خلال الدرجات متعددة الجينات التي تحدد الأشخاص المعرضين لخطر مرتفع مدى الحياة، ومن خلال المتغيرات السببية التي تشير إلى أهداف دوائية. غير أن معظم هذه المعرفة الجينية مستمد من دراسات على ذوي الأصول الأوروبية [5,6]. وتفقد الدرجات متعددة الجينات المبنية على هذه الدراسات دقتها في الأصول الأخرى، مما ينذر بتوسيع الفوارق الصحية [7,35].

أجرى اتحاد علم وراثة الدهون العالمي (GLGC) في عام 2021 تحليلًا تلويًا شمل 1.65 مليون فرد. وقد هيمن المشاركون من ذوي الأصول الأوروبية على هذا التحليل التلوي، وبيّن الاتحاد أن إضافة أصول متنوعة حسّنت التحديد الدقيق للمتغيرات السببية وقابلية نقل الدرجات [8]. ولفقدان قابلية النقل بين الأصول مصادر عدة: الاختلافات في تكرار الأليلات، والاختلافات في اختلال التوازن الارتباطي (LD) بين المتغيرات الواسمة والمتغيرات السببية، والاختلافات الحقيقية في حجم التأثير [9–12]. ولهذه المصادر عواقب عملية؛ فإذا كان ضعف التكرار في مجموعة سكانية جديدة يعكس انخفاض القوة الإحصائية فحسب، أمكن استخدام المواقع الأوروبية كما هي، أما إذا كان يعكس تأثيرات مخففة، فإن الدرجات تحتاج إلى إعادة معايرة محلية.

تتميز المجموعات السكانية العربية في شبه الجزيرة العربية بتواريخ سكانية مميزة، ومعدلات مرتفعة من زواج الأقارب، وامتدادات طويلة من التماثل الزيجوتي [13–15]. وقد أتاح برنامج الجينوم القطري نتائج الارتباط المبنية على تسلسل الجينوم الكامل لـ 45 سمة سريرية، منها أربعة من دهون الدم [16,17]. وخلصت تقييمات سابقة لقابلية الانتقال إلى شبه الجزيرة العربية إلى أن نسبة صغيرة فقط من المتغيرات الأيضية المبلغ عنها قد تكررت [18]. وكررت دراسة قطرية حديثة عن LDL-C عدد 168 موقعًا من مواقع GLGC وأفادت بترابط متوسط في أحجام التأثير [19]، كما قُيّمت الدرجات متعددة الجينات للأمراض القلبية الأيضية لدى العرب السعوديين [20]. وحددت دراسات سعودية على مستوى الجينوم مواقع لاحتشاء عضلة القلب وفرط ثلاثيات الغليسريد في الدم [21,22]. ولم تتساءل أي من هذه الدراسات عن مقدار التكرار المتوقع في ضوء القوة الإحصائية لمجموعة عربية تضم نحو 6,000 شخص، ولم تجمع أي منها، عبر سمات الدهون القياسية الأربع، بين تحليل تكرار معايَر بالقوة الإحصائية، وميل لمعايرة حجم التأثير، وتقدير على مستوى الإحصاءات الموجزة لانتقال الدرجة متعددة الجينات، وتحديد إحصائي دقيق للمتغيرات السببية باستخدام بيانات مفتوحة بالكامل.

لذلك قيّمنا قابلية انتقال المتغيرات المرجعية لدى ذوي الأصول الأوروبية من GLGC لعام 2021 الخاصة بـ LDL-C وكوليسترول البروتين الدهني مرتفع الكثافة (HDL-C) والدهون الثلاثية (TG) والكوليسترول الكلي (TC) إلى المجموعة السكانية القطرية، باستخدام الإحصاءات الموجزة المتاحة للعموم فقط. وتمثلت أهدافنا في: (i) مقارنة توافق الاتجاه والتكرار المرصودين بالقيم المتوقعة من القوة الإحصائية؛ (ii) تقدير تخفيف أحجام التأثير وأثره في التباين المفسَّر بدرجة متعددة الجينات موزونة بالتأثيرات الأوروبية؛ (iii) اختبار عدم التجانس بين الأصول عند المواقع المفردة؛ (iv) إجراء تحديد دقيق للمواقع القطرية ذات الدلالة على مستوى الجينوم المتداخلة مع الإشارات الأوروبية.

2. الطرق

2.1 تصميم الدراسة وإعداد التقرير

كانت هذه دراسة ارتباط جيني عابرة للأصول تستند إلى تحليل ثانوي للإحصاءات الموجزة المنشورة لدراسات الارتباط على مستوى الجينوم (GWAS). وأُعد التقرير وفقًا لامتداد STREGA (STrengthening the REporting of Genetic Association Studies) لبيان STROBE [23]. وثُبّتت خطة التحليل (السمات، وقواعد المطابقة، والمقاييس، والعتبات، والتوزيع المسبق للتحديد الدقيق) قبل فحص بيانات النتائج، وهي ملخصة في البروتوكول المتاح مع مواد الدراسة.

2.2 مصادر البيانات

جاءت الإحصاءات الموجزة القطرية من دراسة GWAS لبرنامج الجينوم القطري التي شملت 45 سمة سريرية [16]، وأُجريت على المشاركين في البنك الحيوي القطري القائم على السكان [17]. وشمل تحليل الاكتشاف 6,218 بالغًا خضعوا لتسلسل الجينوم الكامل، واستخدم نماذج خطية مختلطة معدلة حسب العمر والجنس والمكونات الرئيسة والقرابة. وحُوّلت السمات تحويلًا طبيعيًا

معكوسًا قائمًا على الرتب، ولذلك يُعبَّر عن التأثيرات بوحدات الانحراف المعياري (SD) لكل أليل [16]. ونزّلنا الملفات المتوافقة وفق GRCh38 لـ LDL-C (GCST90013333؛ n = 5,972) وHDL-C (GCST90013332؛ n = 6,013) وTG (GCST90013334؛ n = 6,017) وTC (GCST90013331؛ n = 6,017) من فهرس NHGRI-EBI GWAS Catalog [24] في 6 نوفمبر 2026، وتحققنا من كل ملف مقابل المجموع الاختباري للمستودع.

جاءت المتغيرات المرجعية لدى ذوي الأصول الأوروبية من تحليلات GLGC لعام 2021 الخاصة بذوي الأصول الأوروبية (n = 1,320,016) [8]، واستُرجعت عبر واجهة GWAS Catalog للأرقام المرجعية GCST90239658 (LDL-C) وGCST90239652 (HDL-C) وGCST90239664 (TG) وGCST90239676 (TC). وقد حُوّلت البواقي في GLGC تحويلًا طبيعيًا معكوسًا، ولذلك فإن التأثيرات الأوروبية أيضًا بوحدات SD [8]. وسجلنا لكل متغير مرجعي موقعه وفق GRCh38، وأليل التأثير، وتكرار أليل التأثير، والتأثير لكل أليل، وفاصل الثقة 95% (CI)، وقيمة P. واشتُق الخطأ المعياري (SE) بقسمة عرض فاصل الثقة على 3.92.

2.3 مطابقة المتغيرات وضبط الجودة

استُبعدت المتغيرات القطرية التي يقل تكرار أليلها الأصغر (MAF) عن 0.01، اتساقًا مع التحليل الأصلي. وطوبقت المتغيرات المرجعية مع المتغيرات القطرية حسب الكروموسوم والموقع والأليل، ووُحّدت التأثيرات والتكرارات القطرية على أليل التأثير الأوروبي. واستُبعدت المتغيرات ملتبسة السلسلة (A/T أو C/G) التي يزيد MAF القطري لها على 0.42. وعند تكرار الموقع نفسه، احتُفظ بالسجل ذي أصغر قيمة P أوروبية. ولأن الملفات القطرية تورد التأثيرات وقيم P دون أخطاء معيارية، فقد استُخلص SE بوصفه $|z/\beta|$، حيث z هو مُكمِّم التوزيع الطبيعي المعياري عند $P/2$ - 1. وعندما تجاوزت P القيمة 0.92 وأصبحت هذه النسبة غير مستقرة، قُرّب SE بالصيغة $1/\sqrt{2f(1 - f)N}$، حيث f تكرار أليل التأثير وN حجم العينة. وقُدّرت معلمة التحكم الجينومي λ من عينة منتظمة بنسبة 1 من كل 100 من جميع المتغيرات القطرية القابلة للتحليل.

2.4 النتائج المدروسة والتحليل الإحصائي

أُجريت جميع التحليلات لكل سمة من سمات الدهون على حدة.

توافق الاتجاه والتكرار. عُدّ المتغير متوافقًا عندما كانت إشارة تأثيره القطري مماثلة لإشارة تأثيره الأوروبي. وعُدّ متكررًا تكرارًا اسميًا عندما كانت، إضافةً إلى ذلك، قيمة P القطرية < 0.05، وعُدّ متكررًا بتصحيح Bonferroni عندما كانت قيمة P القطرية < 0.05/m، حيث m عدد المتغيرات المحللة لتلك السمة. وافترضت القيم المتوقعة انتقالًا كاملًا، أي أن التأثير القطري الحقيقي يساوي التأثير الأوروبي. وبموجب هذا الافتراض، يكون احتمال التوافق للمتغير i هو $\Phi(|\beta_{E,i}|/\mathrm{SE}_{Q,i})$ والقوة الإحصائية للتكرار هي $\Phi(|\beta_{E,i}|/\mathrm{SE}_{Q,i} - z_{crit})$، حيث Φ دالة التوزيع الطبيعي المعياري، و$\beta_{E,i}$ التأثير الأوروبي، و$\mathrm{SE}_{Q,i}$ الخطأ المعياري القطري، وz_{crit} القيمة الحرجة ثنائية الطرف. وقورنت الأعداد المرصودة بتوزيع Poisson-binomial (توزيع بواسون ذي الحدين) لهذه الاحتمالات باستخدام احتمال الذيل الأدنى. وتُعرض النسب مع فواصل ثقة 95% بطريقة Clopper–Pearson.

معايرة حجم التأثير. أُجري انحدار للتأثيرات القطرية على التأثيرات الأوروبية مرورًا بنقطة الأصل، موزونًا بـ $1/\mathrm{SE}_Q^2$ [25]. ويساوي الميل 1 عند الانتقال الكامل، وينخفض عن 1 عندما تكون التأثيرات مخففة. وضُخّم خطؤه المعياري بالجذر التربيعي لمعامل التشتت المتبقي عندما تجاوز هذا المعامل 1. ويُعرض معامل ارتباط Pearson بين التأثيرات مع فاصل ثقة 95% محوَّل بتحويل Fisher.

عدم التجانس بين الأصول. حسبنا لكل متغير إحصاءة Q لـ Cochran للمقارنة بين التأثيرين الأوروبي والقطري (درجة حرية 1)، وإحصاءة I^2 الخاصة بها [26]، وتقديرًا تلويًا بنموذج التأثيرات الثابتة الموزون بمقلوب التباين.

انتقال الدرجة متعددة الجينات من الإحصاءات الموجزة. قُيّمت درجة موزونة بالتأثيرات الأوروبية من الإحصاءات الموجزة بعد تقليم المتغيرات المرجعية بحيث تفصل بينها مسافة 500 kb على الأقل. وبافتراض استقلال المتغيرات وسمة معيارية، يكون التغاير بين الدرجة والسمة $\sum_i w_i H_{Q,i} \beta_{E,i}$ وتباين الدرجة $\sum_i w_i^2 H_i^2$، حيث $w_i = H_i \beta_{E,i}$ و$2f_i H_i = 2f_i(1 - f_i)_{Q,i}$). ويساوي التباين المفسَّر (R^2) مربع التغاير مقسومًا على تباين الدرجة. وطرحنا تباين المعاينة للتغاير، $\sum_i (w_i H_i)^2 \mathrm{SE}^2_{Q,i}$، من مربع التغاير لإزالة التحيز التصاعدي الناجم عن الضوضاء، واشتققنا فواصل الثقة 95% من فاصل الثقة الطبيعي للتغاير. ويساوي R^2 المتوقع عند الانتقال الكامل تباين الدرجة، وقد حسبناه بالتكرارات القطرية وبالتكرارات الأوروبية. وتعزل النسبة بينهما أثر الاختلافات في تكرار الأليلات، في حين تعزل نسبة R^2 المرصود إلى المتوقع (R^2 النسبي) أثر تخفيف أحجام التأثير.

المواقع القطرية ذات الدلالة على مستوى الجينوم والتحديد الدقيق. جُمعت المتغيرات القطرية ذات $P < 5 \times 10^{-8}$ حسب المسافة (500 kb) في متغيرات قائدة، وفُحص كل متغير قائد بحثًا عن متغير مرجعي لأي سمة دهنية ضمن 500 kb. وأُجري تحديد دقيق لكل متغير قائد قطري يقع ضمن 250 kb من متغير مرجعي أوروبي للسمة نفسها، عبر نطاق ±250 kb وفق نموذج المتغير السببي الواحد. واستخدمنا عوامل Bayes التقريبية لـ Wakefield مع انحراف معياري مسبق للتأثير قدره 0.15 [27]،

واشتققنا احتمالات الإدراج اللاحقة (PIP) وكوّنّا مجموعات الموثوقية 95% [28]. واستُخدم انحرافان معياريان مسبقان قدرهما 0.10 و0.20 تحليلًا للحساسية. ووُسمت المواقع التي يوجد ضمن 1 Mb منها متغير قائد قطري آخر ذو دلالة على مستوى الجينوم بوصفها مناطق متعددة الإشارات، واستُبعدت من ملخص التحديد الدقيق.

التعددية والسمات المترابطة. يرتبط LDL-C وTC ارتباطًا قويًا، ويشترك كثير من المتغيرات المرجعية بين السمات. لذلك نعرض كل سمة على حدة ونتعامل مع الأعداد المجمعة عبر السمات بوصفها وصفية فقط. وحُددت عتبات Bonferroni داخل كل سمة، ولا يجمع أي اختبار رسمي السمات الأربع.

تحليلات الحساسية. أُعيد تقدير ميل المعايرة (a) مقتصرًا على المتغيرات المرجعية ذات MAF قطري ≥ 0.05، و(b) مقتصرًا على المتغيرات المرجعية ذات P أوروبية < 5×10^{-12}، مما يحد من آثار لعنة الفائز، و(c) مع استبدال كل SE قطري بالتقريب القائم على التكرار.

أُجري تحليل إضافي لاحق لتقليم المسافة أثناء المراجعة التحريرية؛ انظر النتائج.

البرمجيات. استُخدم Python 3.13.18 مع NumPy 2.5.3 وpandas 2.3.3 وSciPy 1.18.1 وstatsmodels 0.15.0، ورُسمت الأشكال باستخدام matplotlib 3.11.2. ويقرأ نص برمجي واحد قابل لإعادة الإنتاج المستخلصات المجمدة فقط.

3. النتائج

3.1 البيانات ومطابقة المتغيرات

وفّرت دراسة GWAS القطرية 7,731,946 متغيرًا قابلًا للتحليل لكل سمة، وتراوحت معلمة التحكم الجينومي λ بين 0.990 و1.056 (الجدول 1). ومن بين 1,600 متغير مرجعي أوروبي مفهرس (403 لـ LDL-C، و380 لـ HDL-C، و388 لـ TG، و429 لـ TC)، تعذّر استخلاص تقدير التأثير أو فاصل الثقة أو الموضع لـ 4 منها، ولم يكن 494 متغيرًا موجودًا في البيانات القطرية، ويعود ذلك في الغالب إلى ندرتها أو غيابها فيها، واستُبعد 15 متغيرًا ملتبس السلسلة. وتبقّى بذلك 269 و261 و274 و283 متغيرًا على التوالي (1,087 زوجًا من أزواج المتغير والسمة). ومن بين 509 متغيرات مرجعية قابلة للاستخلاص لم تُحلل، كان لدى 250 منها MAF أوروبي أقل من 0.05 ولدى 132 منها MAF أوروبي أقل من 0.01، وبلغ وسيط MAF الأوروبي لها 0.054، مقارنةً بوسيط MAF قطري يبلغ نحو 0.25 بين المتغيرات المحللة. ومن ثم تتكون المجموعة المحللة أساسًا من متغيرات شائعة. وتراوح وسيط MAF القطري للمتغيرات المحللة بين 0.244 و0.274، وكانت تكرارات أليل التأثير مترابطة بقوة بين الأصول (r = 0.92 إلى 0.93؛ الشكل 3A). واقتضى الأمر تقريب SE القطري لـ 4.5% إلى 7.4% من المتغيرات، وحيثما توافر التقديران كلاهما كان التوافق بينهما وثيقًا (r ≥ 0.999). ولأن معظم التأثيرات الأوروبية كانت صغيرة (وسيط القيمة المطلقة للتأثير الأوروبي 0.015 إلى 0.016 SD لكل أليل)، فإن وسيط القوة الإحصائية للتكرار الاسمي لمتغير مفرد في العينة القطرية لم يتجاوز 0.097 إلى 0.115.

الجدول 1. مصادر البيانات ومطابقة المتغيرات ومؤشرات الجودة لسمات الدهون الأربع.

الخاصية	LDL-C	HDL-C	TG	TC
حجم عينة GWAS القطرية	5,972	6,013	6,017	6,017
المتغيرات القطرية المحللة (MAF ≥ 0.01)	7,731,946	7,731,946	7,731,946	7,731,946
معلمة التحكم الجينومي λ (GWAS القطرية)	0.990	1.056	1.016	0.998
المتغيرات المرجعية الأوروبية المفهرسة	403	380	388	429
غير متوافرة في البيانات القطرية	132	111	110	141
مستبعدة لكونها متناظرة ملتبسة السلسلة	1	7	4	3
المتغيرات المرجعية المحللة	269	261	274	283
وسيط MAF القطري للمتغيرات المحللة	0.246	0.274	0.266	0.244
ترابط تكرار الأليلات، r (95% CI)	0.92 (0.93–0.90)	0.92 (0.94–0.90)	0.93 (0.94–0.91)	0.92 (0.93–0.90)
وسيط القوة الإحصائية للتكرار الاسمي	0.099	0.115	0.101	0.097

GWAS، دراسة الارتباط على مستوى الجينوم؛ MAF، تكرار الأليل الأصغر؛ LDL-C، كوليسترول البروتين الدهني منخفض الكثافة؛ HDL-C، كوليسترول البروتين الدهني مرتفع الكثافة؛ TG، الدهون الثلاثية؛ TC، الكوليسترول الكلي؛ CI، فاصل الثقة. المتغيرات المرجعية الأوروبية هي ارتباطات ذوي الأصول الأوروبية في اتحاد علم وراثة الدهون العالمي لعام 2021 (n = 1,320,016). القوة الإحصائية هي احتمال الحصول على P < 0.05 في الاتجاه نفسه إذا كان التأثير القطري مساويًا للتأثير الأوروبي. وتعذّر استخلاص

أربعة متغيرات مرجعية مفهرسة (LDL-C 1، وHDL-C 1، وTC 2) لغياب التأثير أو فاصل الثقة أو الموضع، ولذلك يزيد العدد المفهرس على مجموع الصفوف الثلاثة التالية بمقدار هذه المتغيرات.

3.2 توافق الاتجاه والتكرار المعايَر بالقوة الإحصائية

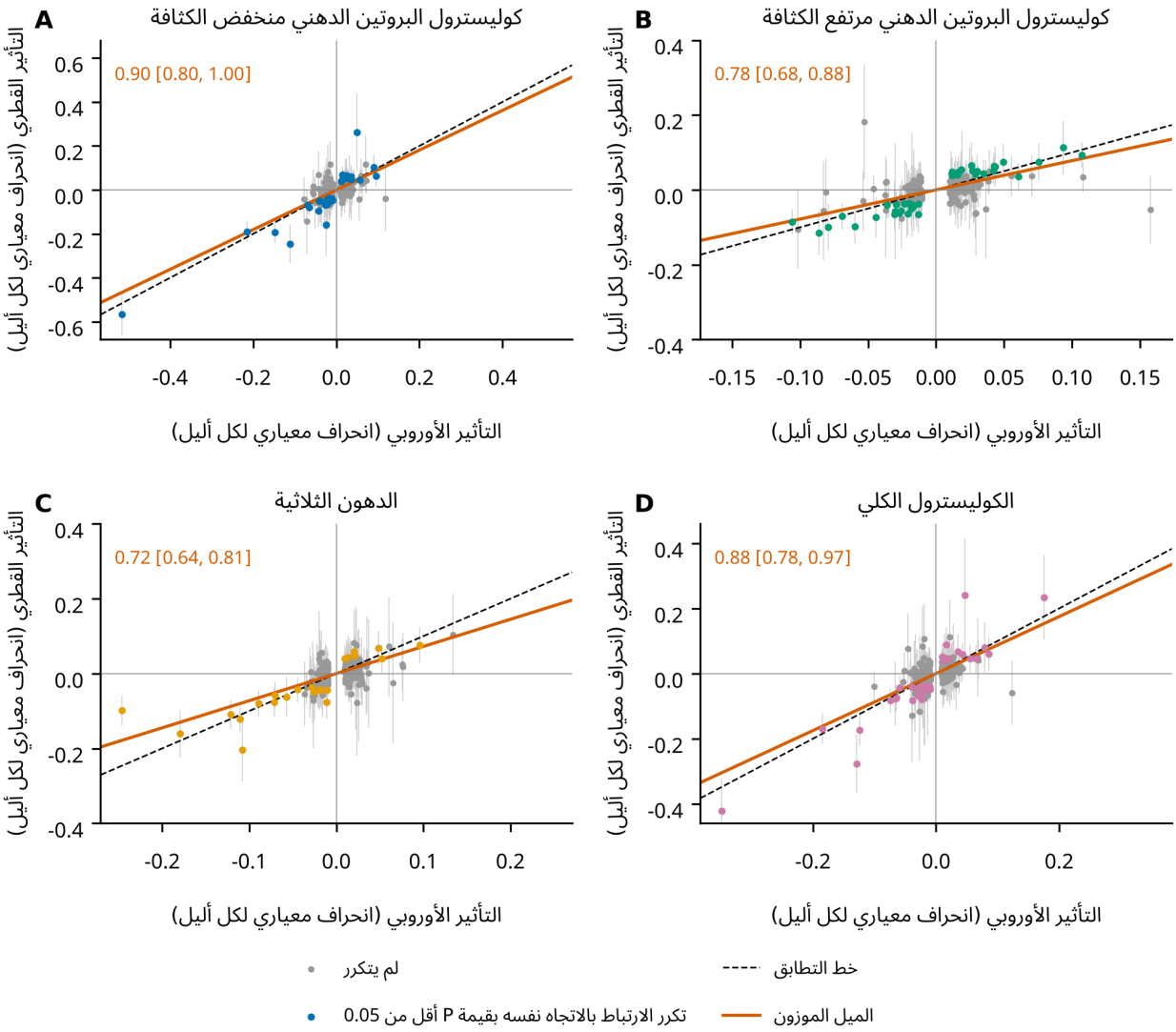
عبر السمات، كان لـ 790 من أصل 1,087 زوجًا من أزواج المتغير والسمة (72.7%) اتجاه التأثير نفسه في الأصلين. وتراوح التوافق بين 69.5% لـ LDL-C و75.1% لـ HDL-C (الجدول 2، الشكل 2A)، وهو أعلى بكثير من الصدفة لكنه أدنى من القيم المتوقعة عند الانتقال الكامل (77.3% إلى 80.2%). وكان هذا العجز قابلًا للكشف إحصائيًا لـ LDL-C (P = 0.001) وHDL-C (P = 0.023) وTG (P = 0.024)، ولكن ليس لـ TC (P = 0.060).

حدث التكرار الاسمي لـ 129 زوجًا إجمالًا مقابل 184.3 متوقعًا. وحسب السمة، تكرر 29 من أصل 269 متغيرًا لـ LDL-C (10.8%) مقابل 42.9 متوقعًا (نسبة المرصود إلى المتوقع 0.68؛ P = 0.004)، و41 من أصل 261 متغيرًا لـ HDL-C (15.7%) مقابل 49.1 (النسبة 0.83؛ P = 0.082)، و23 من أصل 274 متغيرًا لـ TG (8.4%) مقابل 44.7 (النسبة 0.51؛ P < 0.001)، و36 من أصل 283 متغيرًا لـ TC (12.7%) مقابل 47.6 (النسبة 0.76؛ P = 0.017) (الشكل 2B). وفي المقابل، طابق التكرار ذو الدلالة بتصحيح Bonferroni القيمة المتوقعة لكل سمة: 5 مقابل 5.9 لـ LDL-C، و5 مقابل 6.4 لـ HDL-C، و7 مقابل 7.8 لـ TG، و6 مقابل 7.3 لـ TC (جميع قيم P ≥ 0.32؛ الشكل 2C). ومن ثم انتقلت المتغيرات ذات التأثيرات الكبيرة كما هو متوقع، في حين نشأ العجز من المتغيرات الكثيرة ذات التأثيرات الصغيرة.

الجدول 2. قابلية انتقال المتغيرات المرجعية للدهون لدى ذوي الأصول الأوروبية إلى المجموعة السكانية القطرية.

المقياس	LDL-C	HDL-C	TG	TC
الاتجاه نفسه، n (%)	187 (69.5)	196 (75.1)	199 (72.6)	208 (73.5)
الاتجاه نفسه المتوقع، %	77.3	80.2	77.7	77.5
P، الاتجاه المرصود مقابل المتوقع	0.001	0.023	0.024	0.060
التكرار الاسمي، المرصود/المتوقع	29/42.9	41/49.1	23/44.7	36/47.6
نسبة المرصود إلى المتوقع، اسمي	0.68	0.83	0.51	0.76
P، التكرار الاسمي المرصود مقابل المتوقع	0.004	0.082	< 0.001	0.017
التكرار بتصحيح Bonferroni، المرصود/المتوقع	5/5.9	5/6.4	7/7.8	6/7.3
ميل حجم التأثير (95% CI)	0.90 (1.00–0.80)	0.78 (0.88–0.68)	0.72 (0.81–0.64)	0.88 (0.97–0.78)
ترابط حجم التأثير، r (95% CI)	0.77 (0.82–0.72)	0.53 (0.62–0.44)	0.69 (0.75–0.63)	0.74 (0.79–0.68)
عدم التجانس n، P < 0.05 (%)	21 (7.8)	19 (7.3)	18 (6.6)	22 (7.8)
عدم التجانس بعد تصحيح Bonferroni، n	0	1	1	0
R^2 المتوقع للدرجة، بالتكرارات الأوروبية (%)	9.81	5.86	7.32	8.28
R^2 المتوقع للدرجة، بالتكرارات القطرية (%)	7.91	5.90	7.85	7.35
R^2 المرصود للدرجة (%) (95% CI)	6.46 (7.82–5.25)	3.57 (4.52–2.75)	4.13 (5.17–3.23)	5.63 (6.92–4.51)
R^2 النسبي (المرصود/المتوقع) (95% CI)	0.82 (0.99–0.66)	0.60 (0.77–0.47)	0.53 (0.66–0.41)	0.77 (0.94–0.61)

تفترض القيم المتوقعة أن كل تأثير قطري لكل أليل يساوي التأثير الأوروبي، بمعلومية الخطأ المعياري القطري. والميل هو انحدار التأثيرات القطرية على الأوروبية مرورًا بنقطة الأصل والموزون بمقلوب التباين (تشير القيمة 1 إلى الانتقال الكامل). وعدم التجانس هو إحصاءة Q لـ Cochran بدرجة حرية 1 لكل متغير. وR^2 للدرجة هو التباين في كل سمة دهنية المفسَّر بدرجة موزونة بالتأثيرات الأوروبية للمتغيرات المرجعية، مقدرًا من الإحصاءات الموجزة بافتراض استقلال المتغيرات. الاختصارات كما في الجدول 1.

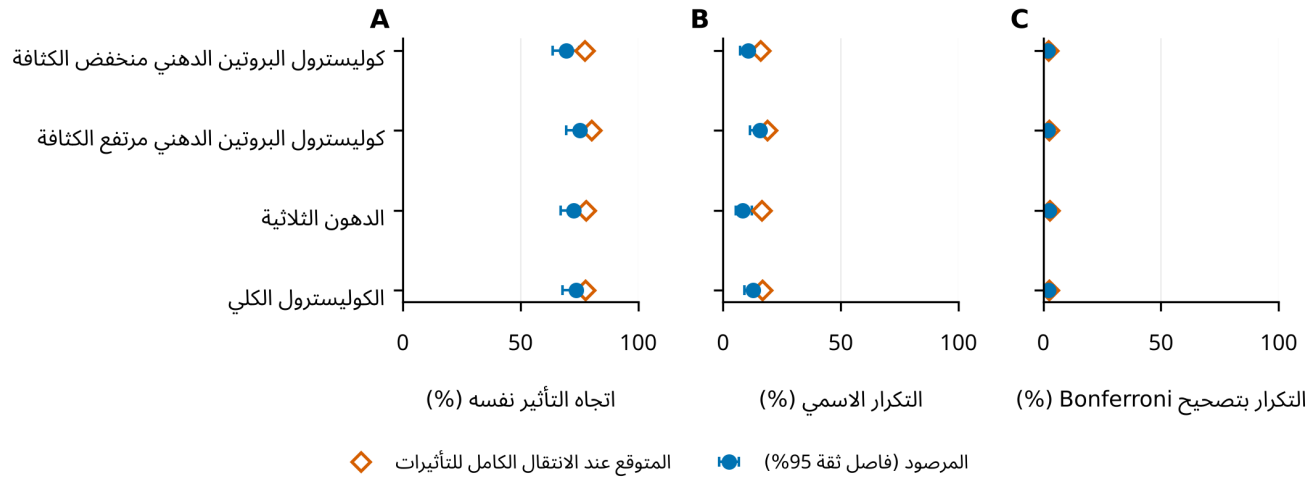


الشكل 1. التأثيرات لكل أليل للمتغيرات المرجعية للدهون لدى ذوي الأصول الأوروبية في الإحصاءات الموجزة الأوروبية (المحور x) والقطرية (المحور y) لكل من (A) كوليسترول LDL، و(B) كوليسترول HDL، و(C) الدهون الثلاثية، و(D) الكوليسترول الكلي. النقاط الملونة تكررت تكرارًا اسميًا ($P < 0.05$، الاتجاه نفسه)، والنقاط الرمادية لم تتكرر. الخطوط الرأسية فواصل ثقة 95% قطرية. الخط المتقطع هو خط التطابق، والخط المتصل هو الميل الموزون بمقلوب التباين مطبوعًا مع فاصل ثقته 95%. التأثيرات بوحدات الانحراف المعياري لمستويات الدهون المحولة تحويلًا طبيعيًا معكوسًا لكل أليل تأثير. LDL، البروتين الدهني منخفض الكثافة؛ HDL، البروتين الدهني مرتفع الكثافة؛ SD، الانحراف المعياري.

3.3 معايرة حجم التأثير وعدم التجانس

كانت التأثيرات القطرية أصغر عمومًا من التأثيرات الأوروبية (الشكل 1)، وإن امتد فاصل الثقة 95% لـ LDL-C حتى 1. فقد بلغ الميل الموزون 0.90 (95% CI 0.80–1.00) لـ LDL-C، و0.88 (0.78–0.97) لـ TC، و0.78 (0.68–0.88) لـ HDL-C، و0.72 (0.64–0.81) لـ TG. وتراوحت معاملات ترابط أحجام التأثير بين 0.53 (HDL-C) و0.77 (LDL-C). وكانت الميول متينة في تحليلات الحساسية؛ إذ أعطى الاقتصار على المتغيرات ذات MAF قطري ≥ 0.05 ميولًا قدرها 0.84 (LDL-C) و0.82 (HDL-C) و0.71 (TG) و0.82 (TC)، وأعطى الاقتصار على المتغيرات الأوروبية قوية الارتباط ($P < 5 \times 10^{-12}$) القيم 0.92 و0.81 و0.74 و0.88، وأعطى SE القائم على التكرار القيم 0.90 و0.78 و0.73 و0.88. ولوحظ عدم تجانس اسمي ($P < 0.05$) لدى 6.6% إلى 7.8% من المتغيرات، وهو قريب من نسبة 5% المتوقعة بالصدفة. ولم يبقَ غير متجانس بعد تصحيح Bonferroni سوى زوجين من أزواج المتغير والسمة. ففي TG، كان للمتغير rs964184 في منطقة ZPR1–APOA5 تأثير لكل أليل قدره -0.098 SD في البيانات القطرية مقابل -0.246 SD في البيانات الأوروبية. وفي HDL-C، يُعد المتغير rs538969346 قرب SMAD7 نادرًا لدى الأوروبيين (تكرار أليل التأثير 0.010) لكنه أكثر شيوعًا لدى القطريين (0.034)، وكان تأثيره في الاتجاه المعاكس. ومن ثم كان التخفيف منتشرًا ومتواضعًا عند المواقع المفردة، ولم يكن مدفوعًا بعدد قليل من المواقع شديدة التعارض.

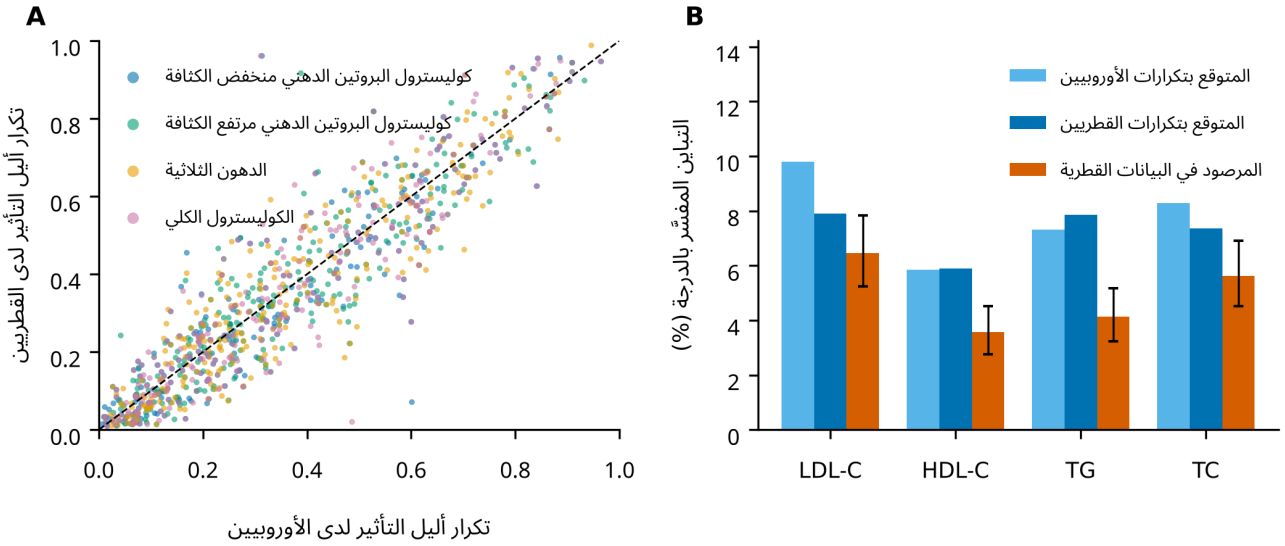
أُضيف أثناء المراجعة تحليل حساسية لاحق بتقليم المسافة عند 1 وMb 5. وعند 5 Mb كانت الميول وفواصل الثقة 95%: LDL-C 0.949 (1.056–0.841)، وHDL-C 0.823 (0.934–0.711)، وTG 0.738 (0.830–0.645)، وTC 0.944 (1.050–0.838). ولا يثبت تقليم المسافة استقلال اختلال الارتباط. وتظل احتمالات بواسون ذي الحدين والفواصل الأصلية مشروطة باستقلال الاختبارات وثبات التأثيرات الأوروبية، وقد تكون مفرطة الدقة. ويضعف التقليم الأوسع دليل تخفيف TC. وهذه تحليلات حساسية لاحقة وليست محددة مسبقًا.



الشكل 2. النسب المئوية المرصودة (دوائر مع فواصل ثقة 95%) والمتوقعة (معينات مفرغة، بافتراض الانتقال الكامل للتأثيرات الأوروبية) للمتغيرات المرجعية التي (A) لها اتجاه التأثير نفسه، و(B) تكررت تكرارًا اسميًا ($P < 0.05$، الاتجاه نفسه)، و(C) تكررت بتصحيح Bonferroni في الإحصاءات الموجزة القطرية. النسب المتوقعة هي متوسط احتمال كل حدث عبر المتغيرات بمعلومية التأثير الأوروبي والخطأ المعياري القطري.

3.4 انتقال الدرجة متعددة الجينات

أزال التقليم متغيرًا واحدًا لـ HDL-C، ولذلك استخدمت تحليلات الدرجة 269 و260 و274 و283 متغيرًا. وعند الانتقال الكامل، كانت الدرجة الموزونة بالتأثيرات الأوروبية ستفسر 7.91% من تباين LDL-C بتكرارات الأليلات القطرية، مقارنةً بـ 9.81% بالتكرارات الأوروبية. ومن ثم فإن اختلافات تكرار الأليلات وحدها تخفض التباين المفسَّر المتوقع إلى 0.81 من القيمة الأوروبية لـ LDL-C وإلى 0.89 لـ TC. أما في HDL-C (1.01) وTG (1.07) فلم يكن لهذه الاختلافات أثر يُذكر (الجدول 2، الشكل 3B). وبلغ التباين المفسَّر المرصود في البيانات القطرية 6.46% (95% CI 5.25–7.82) لـ LDL-C، و5.63% (6.92–4.51) لـ TC، و4.13% (5.17–3.23) لـ TG، و3.57% (4.52–2.75) لـ HDL-C. وبالنسبة إلى القيمة المتوقعة بالتكرارات القطرية، يقابل ذلك قيم R^2 نسبية قدرها 0.82 (0.99–0.66) و0.77 (0.94–0.61) و0.53 (0.66–0.41) و0.60 (0.77–0.47) على التوالي. وكان ارتباط الدرجة ذا دلالة إحصائية عالية لكل سمة (جميع قيم $P < 0.001$)، وساوى تأثيرها لكل وحدة ميل المعايرة (على سبيل المثال 0.90، 95% CI 0.81–0.99، لـ LDL-C).



الشكل 3. (A) تكرارات أليل التأثير للمتغيرات المرجعية المحللة في البيانات الأوروبية والقطرية حسب السمة؛ والخط المتقطع هو خط التطابق. (B) التباين في كل سمة دهنية المفسَّر بدرجة موزونة بالتأثيرات الأوروبية للمتغيرات المرجعية: المتوقع عند الانتقال الكامل بتكرارات الأليلات الأوروبية والقطرية، والمرصود في الإحصاءات الموجزة القطرية مع فواصل ثقة 95%. LDL-C، كوليسترول البروتين الدهني منخفض الكثافة؛ HDL-C، كوليسترول البروتين الدهني مرتفع الكثافة؛ TG، الدهون الثلاثية؛ TC، الكوليسترول الكلي.

3.5 المواقع القطرية ذات الدلالة على مستوى الجينوم والتحديد الدقيق

أسفر التجميع حسب المسافة عن 22 موقعًا قطريًا ذا دلالة على مستوى الجينوم: 5 لـ LDL-C، و5 لـ HDL-C، و7 لـ TG، و5 لـ TC. ووقعت المواقع الـ 22 جميعها ضمن 500 kb من متغير مرجعي أوروبي للسمة نفسها، ومن ثم لم يكن أي منها خاصًا بالأصل عند هذه الدقة. ووقع واحد وعشرون منها ضمن 250 kb وخضعت للتحديد الدقيق. وكانت خمسة منها، جميعها على الكروموسوم 11q23 (Mb 118.0–116.5) لـ HDL-C وTG، مناطق متعددة الإشارات، وكانت متغيراتها القائدة متغيرات منخفضة التكرار (تكرار أليل التأثير 0.011 إلى 0.014) ذات تأثيرات كبيرة، منتشرة على مسافة تقارب 1.3 Mb. ويتسق هذا النمط مع نمط فرداني ممتد واحد تحمله أقلية من المشاركين القطريين، لا مع خمسة مواقع مستقلة. ولم يتسنَّ اختبار هذا التفسير لعدم توافر مرجع LD عربي.

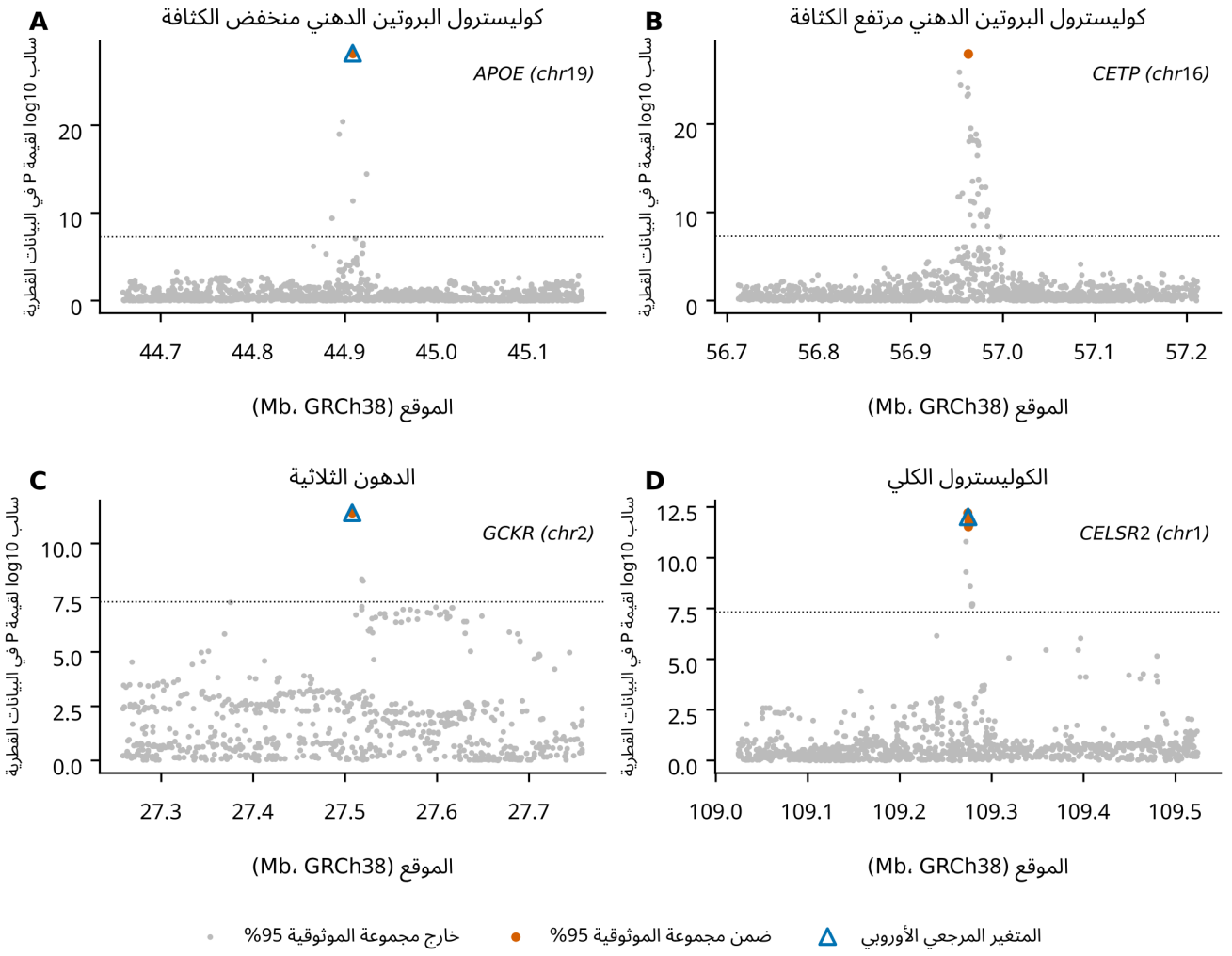
من بين المواقع الـ 16 أحادية الإشارة، احتوت مجموعة الموثوقية 95% في الوسيط على 6 متغيرات (المدى 1 إلى 53)، واحتوت 7 مجموعات على 5 متغيرات أو أقل (الجدول 3، الشكل 4). وانحصرت ثلاث مجموعات في متغير واحد: rs7412 APOE (الأليل المحدِّد للنمط ε2) لـ LDL-C ولـ TC، وrs1260326 GCKR لـ TG. وكان المتغير المرجعي الأوروبي موجودًا في البيانات القطرية عند 15 من هذه المواقع، ووقع داخل مجموعة الموثوقية القطرية 95% عند 12 منها. وعند CELSR2، احتوت مجموعة الموثوقية القطرية لـ LDL-C على المتغير التنظيمي الوظيفي لـ rs12740374 SORT1 [32]. وعند موقع TM6SF2، بلغت الإشارة القطرية ذروتها عند rs111234557 على بعد نحو 57 kb من المتغير المرجعي الأوروبي مغيّر المعنى rs58542926 [33]. ووقع المتغير المرجعي الأوروبي خارج مجموعة الموثوقية القطرية عند APOB (LDL-C وTC) وLPL (HDL-C). ولم تتغير أحجام مجموعات الموثوقية إلا قليلًا مع انحرافين معياريين مسبقين قدرهما 0.10 أو 0.20 (البيانات التكميلية).

الجدول 3. التحديد الدقيق البايزي بنموذج المتغير السببي الواحد للمواقع القطرية ذات الدلالة على مستوى الجينوم المتداخلة مع المتغيرات المرجعية الأوروبية.

السمة	الموقع (الجين)	Chr	المتغيرات	المتغير الأعلى (PIP)	حجم مجموعة الموثوقية 95%	المتغير المرجعي الأوروبي	PIP للمتغير المرجعي	المتغير المرجعي ضمن المجموعة
LDL C-	APOE	19	1,622	rs7412 (1.00)	1	rs7412	1.00	نعم
LDL C-	LDLR	19	1,252	rs112736558 (0.06)	28	rs73015024	0.04	نعم
LDL C-	CELSR2	1	1,268	rs12740374 (0.23)	6	rs12740374	0.23	نعم

السمة	الموقع (الجين)	Chr	المتغيرات	المتغير الأعلى (PIP)	حجم مجموعة الموثوقية 95%	المتغير المرجعي الأوروبي	PIP للمتغير المرجعي	المتغير المرجعي ضمن المجموعة
LDL-C	APOB	2	1,325	rs1800481 (0.33)	9	rs934197	0.00	لا
LDL-C	TM6SF2	19	1,230	rs111234557 (0.67)	8	rs58542926	0.10	نعم
HDL-C	CETP	16	1,560	rs708272 (0.56)	2	rs183130	NA	غير موجود في البيانات القطرية
HDL-C	LPL	8	2,137	rs74869266 (0.36)	27	rs15285	0.00	لا
HDL-C	ALDH1A2	15	1,842	rs1077834 (0.80)	6	rs261290	0.05	نعم
TG	GCKR	2	719	rs1260326 (1.00)	1	rs1260326	1.00	نعم
TG	MLXIPL	7	1,032	rs33951980 (0.13)	53	rs3812316	0.05	نعم
TG	TM6SF2	19	1,230	rs111234557 (0.83)	3	rs58542926	0.03	نعم
TC	APOE	19	1,622	rs7412 (1.00)	1	rs7412	1.00	نعم
TC	CELSR2	1	1,269	rs7528419 (0.38)	5	rs12740374	0.24	نعم
TC	LDLR	19	1,272	rs8106503 (0.11)	27	rs73015024	0.04	نعم
TC	TM6SF2	19	1,230	rs111234557 (0.92)	2	rs58542926	0.03	نعم
TC	APOB	2	1,325	rs1800481 (0.20)	36	rs934197	0.00	لا

Chr، الكروموسوم؛ PIP، احتمال الإدراج اللاحق المحسوب من عوامل Bayes التقريبية لـ Wakefield (انحراف معياري مسبق 0.15) ضمن ±250 kb من المتغير القطري القائد؛ NA، غير منطبق. استُبعدت خمسة مواقع على الكروموسوم 11q23 فيها عدة إشارات ذات دلالة على مستوى الجينوم ضمن 1 Mb لأن افتراض المتغير السببي الواحد غير قابل للتطبيق هناك. الاختصارات كما في الجدول 1.



الشكل 4. الارتباط الإقليمي والتحديد الدقيق في الإحصاءات الموجزة القطرية عند APOE (A) لكوليسترول LDL، وCETP (B) لكوليسترول HDL، وGCKR (C) للدهون الثلاثية، وCELSR2 (D) للكوليسترول الكلي. تشكل النقاط البرتقالية مجموعة الموثوقية 95%. ويشير المثلث المفرغ إلى المتغير المرجعي الأوروبي، ويشير الخط المنقط إلى $P = 5 \times 10^{-8}$. المواقع وفق GRCh38. تفترض مجموعات الموثوقية وجود متغير سببي واحد في كل منطقة (عوامل Bayes التقريبية لـ Wakefield، انحراف معياري مسبق 0.15). والمتغير المرجعي الأوروبي عند CETP (rs183130) غائب عن البيانات القطرية.

4. المناقشة

4.1 النتائج الرئيسة

في 1,087 زوجًا من المتغيرات والسمات، توافق اتجاه التأثير في نحو ثلاثة أرباع الأزواج. وتكررت المتغيرات الكبيرة التأثير بمعدل قريب من المتوقع، بينما كان التكرار الاسمي أقل من توقع الانتقال الكامل. وتعتمد المقارنات على التقديرات النقطية الأوروبية واستقلال اختبارات المتغيرات وقابلية مقارنة مقاييس السمات؛ ولا تثبت سببًا بيولوجيًا للتخفيف، ولا تفصل الاختلاف الحقيقي في التأثير عن اختلال الارتباط والانتقاء والقياس ولعنة الفائز.

4.2 المقارنة بالدراسات السابقة

خلص Hebbar وزملاؤه إلى أن المتغيرات الأيضية المحددة بدراسات GWAS أظهرت قابلية انتقال منخفضة إلى شبه الجزيرة العربية [18]. وتقيّد معايرتنا بالقوة الإحصائية هذا الاستنتاج؛ فمع وسيط قوة إحصائية يبلغ نحو 10% لكلٍ متغير، فإن مجموعة قطرية تضم 6,000 شخص لن تكرر سوى أقلية صغيرة من المتغيرات الأوروبية حتى لو انتقل كل تأثير انتقالًا تامًا. ومن ثم فإن عدّ حالات التكرار ذات الدلالة يبخس قابلية الانتقال حقها، والمقارنة الملائمة هي مع العدد المتوقع. وأفاد Ismail Umlai وزملاؤه بترابط متوسط في أحجام تأثير LDL-C بين البيانات القطرية وبيانات GLGC باستخدام بيانات فردية من البنك الحيوي القطري [19]. ويتفق ميل المعايرة لـ LDL-C في بياناتنا المفتوحة (0.90) وترابط التأثيرات (0.77) معه في الاتجاه، ونوسّع المقارنة لتشمل HDL-C وTG وTC. وبيّن Shim وزملاؤه أن الدرجات متعددة الجينات للأمراض القلبية الأيضية، بعد تحسينها للعرب باستخدام بيانات عامة، تفوقت على الدرجات المنشورة وجاء أداؤها مماثلًا لأدائها لدى ذوي الأصول الأوروبية في مرضى عرب سعوديين

[20]. وتقيس تقديراتنا لـ R^2 النسبي الفقدان الذي يحدث عند تطبيق الأوزان الأوروبية دون هذا التحسين، على مستوى كل سمة دهنية، باستخدام بيانات يمكن لأي شخص إعادة إنتاجها.

تدعم الدراسات العابرة للأصول اشتراك الإشارات الجينية وانخفاض قابلية انتقال الدرجات بين الأصول [9–11,29–31]. وتتفق الميول المنخفضة هنا مع اختلاف اختلال الارتباط وتعريف السمات والانتقاء والبيئة. ولا يحدد هذا التصميم الآلية المسؤولة ولا يثبت انخفاض التأثيرات السببية.

تضيف نتائج التحديد الدقيق أدلة على الآلية. فقد احتوت مجموعة الموثوقية عند SORT1–CELSR2 على rs12740374، وهو المتغير التنظيمي الذي ثبت تجريبيًا أنه يغيّر تعبير SORT1 في الكبد [32]. وبلغت إشارة TM6SF2 في البيانات القطرية ذروتها قرب المتغير مغيّر المعنى المعروف E167K وليس عنده [33]، وهو ما يسوّغ المتابعة باستخدام معلومات LD القطرية. وعند APOB وLPL، وقعت المتغيرات المرجعية الأوروبية خارج مجموعات الموثوقية القطرية. وهذه مواقع مرشحة يمكن فيها لبنية LD العربية أن تساعد على التمييز بين المتغيرات السببية والمتغيرات الواسمة، وهي الحجة الرئيسة التي قدمها GLGC لصالح التحديد الدقيق في مجموعات متنوعة [8]. وتتسق الإشارات منخفضة التكرار المتجمعة على الكروموسوم 11q23، قرب عنقود APOA5–APOA4–APOC3–APOA1، مع نمط فرداني ممتد، كما هو متوقع في مجموعة سكانية ذات امتدادات طويلة من التماثل الزيجوتي [13,15]. وتتداخل أيضًا مع مناطق وُصفت فيها في قطر متغيرات دهنية عائلية ومتغيرات متزايدة الشيوع سكانيًا [34]. وكان أقوى عدم تجانس على مستوى الموقع في الدراسة في هذه المنطقة أيضًا؛ إذ كان للمتغير المرجعي الأوروبي rs964184 أقل من نصف تأثيره الأوروبي في الدهون الثلاثية في البيانات القطرية. ويتسق ذلك مع الرأي القائل إن البنية السببية لمنطقة APOA5 تختلف بين المجموعتين السكانيتين، ربما لأن نمطًا فردانيًا قطريًا منخفض التكرار يحمل جزءًا من الإشارة التي تسِمها متغيرات شائعة لدى الأوروبيين.

قد تسهم آليات عدة في انخفاض ميول المتغيرات الواسمَة، ومنها اختلال الارتباط والعلاقات بين الجينات والبيئة وتعديل العلاج وقياس السمات. كما تُقدَّر تأثيرات الاكتشاف الأوروبية بخطأ وتُنتقى بحسب الدلالة؛ ولا يزيل القصر على المتغيرات شديدة الارتباط أثر لعنة الفائز. لذلك تُفسَّر النتائج بوصفها وصفًا لانتقال الارتباطات المنشورة للمتغيرات الواسمَة، لا تقديرات للتأثيرات السببية الخاصة بكل أصل.

وللإطار المعايَر بالقوة الإحصائية قيمة منهجية تتجاوز الدهون. فدراسات التكرار في المجموعات السكانية الممثلة تمثيلًا ناقصًا صغيرة في الغالب، والأحكام الثنائية بالتكرار أو عدمه يسهل أن تُفهم خطأً على أنها دليل على غياب قابلية الانتقال. أما مقارنة الأعداد المرصودة بالمتوقعة وتقدير ميل المعايرة مع فاصل ثقته فيحوّلان الإحصاءات الموجزة نفسها إلى تقديرات كمية للانتقال. ولأن هذا الإطار لا يحتاج إلا إلى أحجام التأثير والأخطاء المعيارية وتكرارات الأليلات، فيمكن تطبيقه على أي سمة تُتاح لها إحصاءات موجزة عربية، ومنها سمات سكر الدم وضغط الدم وتراكيز فيتامين D.

4.3 الآثار بالنسبة إلى المملكة العربية السعودية والخليج

تؤكد الإرشادات السعودية الكشف عن الخطر [3]، وتُدرس الدرجات متعددة الجينات لدى العرب [20]. ولا يثبت هذا التحليل بالإحصاءات الموجزة منفعة سريرية أو خطرًا قلبيًا وعائيًا مطلقًا. وضرب الأوزان كلها في ميل مشترك يغير مقياس الدرجة ولا يحسن ترتيبها أو ارتباطها أو التباين الذي تفسره. ويتطلب التطبيق مرجعًا عربيًا لاختلال الارتباط ومواءمة السمات وتحققًا مستقلًا ببيانات فردية. ولا يجوز افتراض انتقال النتائج القطرية مباشرةً إلى السكان السعوديين [13,14].

4.4 نقاط القوة والقيود

تتمتع الدراسة بعدة نقاط قوة؛ فهي تعتمد كليًا على إحصاءات موجزة عامة جرى التحقق منها بالمجموع الاختباري وعلى نص برمجي واحد قابل لإعادة الإنتاج، وتقارن التكرار بقيم متوقعة صريحة مبنية على القوة الإحصائية بدلًا من عدّ النتائج ذات الدلالة، وتفصل أثر اختلافات تكرار الأليلات عن أثر تخفيف أحجام التأثير في أداء الدرجة، وتغطي سمات الدهون القياسية الأربع، وكانت معايرة حجم التأثير فيها مستقرة عبر تحليلات الحساسية.

وثمة قيود أيضًا. أولًا، لم يتوافر لدينا مرجع LD عربي، ولذلك افترض التحديد الدقيق وجود متغير سببي واحد في كل منطقة، واستُبعدت المناطق متعددة الإشارات بدلًا من حلها. ثانيًا، تفترض تقديرات R^2 استقلال المتغيرات المرجعية، ويظل بعض LD المتبقي بين متغيرات GLGC المرجعية ضمن نوافذ 500 kb ممكنًا، وإن كان التقليم قد أزال متغيرًا واحدًا فقط. ثالثًا، أُعيد بناء الأخطاء المعيارية القطرية من قيم P، وكان التوافق مع التقريب القائم على التكرار مرتفعًا جدًا، لكن تقريب قيم P الصغيرة جدًا قد يضيف خطأً طفيفًا. رابعًا، لم يتوافر نحو ثلث المتغيرات المرجعية الأوروبية في البيانات القطرية، ويعود ذلك في الغالب إلى ندرتها أو غيابها، ومن ثم تعذّر تقييم قابلية انتقال المتغيرات النادرة. خامسًا، قد تحمل التأثيرات الأوروبية أثرًا متبقيًا للعنة الفائز يحيّز الميول نحو الانخفاض ويضخّم الأعداد المتوقعة للتكرار، غير أن الميل المقتصر على المتغيرات قوية الارتباط كان مماثلًا، مما يرجح عدم وجود تحيز كبير. سادسًا، قيست الدهون في مجموعتي البيانات دون مواءمة كاملة لحالة الصيام أو تعديل العلاج، وقد عدّل GLGC قيم LDL-C وTC بحسب الأدوية الخافضة للدهون، وقد تسهم هذه الاختلافات في التخفيف. سابعًا، استُخدم المشاركون القطريون بديلًا تقريبيًا للعرب في الشرق الأوسط، ولم تُختبر قابلية الانتقال إلى السكان العرب في السعودية ودول

الخليج الأخرى مباشرةً. وأخيرًا، فإن العينة القطرية متوسطة الحجم، ولذلك تكون فواصل الثقة للتقديرات الخاصة بكل سمة واسعة.

نطاق الاستدلال: تعتمد تقديرات R^2 والفواصل واحتمالات التكرار على استقلال المتغيرات وثبات التأثيرات الأوروبية. ولا يثبت تقليم المسافة استقلال LD. لذا تفسر القيم الأصلية كملخصات استكشافية مشروطة، ولا تستخدم دلالتها لإثبات اختلاف سببي بين الأصول أو منفعة سريرية.

4.5 البحوث المستقبلية

ينبغي أن تجمع البحوث المستقبلية بين لوحات LD عربية والتحديد الدقيق متعدد الأصول وتحليل التموضع المشترك [40] لحل إشارات APOB وLPL وTM6SF2 والكروموسوم 11q23. وينبغي التحقق من الدرجات متعددة الجينات المعايَرة محليًا ومتعددة الأصول مقابل الأحداث القلبية الوعائية الحادثة في المجموعات السعودية والخليجية [36,38]. كما ينبغي توسيع التحليل ليشمل سمات البروتينات والمستقلبات، التي أُبلغ فيها عن اختلافات في البنية الجينية بين العرب والأوروبيين [39].

5. الاستنتاجات

أظهرت ارتباطات الدهون الأوروبية المنشورة توافقًا كبيرًا في الاتجاه لدى العرب القطريين. وظلت ميول HDL-C وTG أقل من 1 مع تقليم أوسع حسب المسافة، بينما تأثر الاستدلال بشأن LDL-C وTC بدرجة أكبر. وتبقى تقديرات تباين الدرجة والمجموعات المرشحة وفق افتراض الإشارة الواحدة استكشافية ومشروطة. ويلزم تحقق ببيانات فردية ومعلومات عربية عن اختلال الارتباط قبل التطبيق السريري.

الإقرارات

التمويل: NA

تضارب المصالح: NA

بيان مجلس المراجعة المؤسسية: استخدمت هذه الدراسة فقط إحصاءات موجزة إجمالية متاحة للعموم وغير محددة للهوية لدراسات الارتباط على مستوى الجينوم من NHGRI-EBI GWAS Catalog؛ ولم يُطّلع على أي بيانات على المستوى الفردي، ولا يُدّعى الحصول على موافقة أخلاقية.

بيان الموافقة المستنيرة: غير منطبق (بيانات إجمالية عامة غير محددة للهوية).

بيان إتاحة البيانات: تتوافر الإحصاءات الموجزة لبرنامج الجينوم القطري في NHGRI-EBI GWAS Catalog (الأرقام المرجعية GCST90013331 وGCST90013332 وGCST90013333 وGCST90013334؛ https://www.ebi.ac.uk/gwas/)، وتتوافر ارتباطات ذوي الأصول الأوروبية في GLGC لعام 2021 في المورد نفسه (الأرقام المرجعية GCST90239652 وGCST90239658 وGCST90239664 وGCST90239676). وتتوافر الشيفرة البرمجية للتحليل والنتائج المشتقة على مستوى المتغيرات مع مواد الدراسة.

مساهمات المؤلفين: NA

المراجع

[1] Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. J Am Coll Cardiol. 2020;76(25):2982-3021. doi:10.1016/j.jacc.2020.11.010

[2] Ference BA, Ginsberg HN, Graham I, Ray KK, Packard CJ, Bruckert E, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. Eur Heart J. 2017;38(32):2459-2472. doi:10.1093/eurheartj/ehx144

[3] AlRahimi J, AlSaif S, Alasnag M, Awan Z, Almutairi F, Al Mudaiheem H, et al. 2022 Saudi guidelines for the management of dyslipidemia. Heart Views. 2023;24(2):67-92. doi:10.4103/heartviews.heartviews_102_22

[4] AlMuhaidib S, AlBuhairan F, Tamimi W, AlDubayee M, AlAqeel A, Babiker A, et al. Prevalence and factors associated with dyslipidemia among adolescents in Saudi Arabia. Sci Rep. 2022;12:16888. doi:10.1038/s41598-022-21262-9

[5] Popejoy AB, Fullerton SM. Genomics is failing on diversity. Nature. 2016;538(7624):161-164. doi:10.1038/538161a

[6] Sirugo G, Williams SM, Tishkoff SA. The missing diversity in human genetic studies. Cell. 2019;177(1):26-31. doi:10.1016/j.cell.2019.02.048

[7] Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x

[8] Graham SE, Clarke SL, Wu KHH, Kanoni S, Zajac GJM, Ramdas S, et al. The power of genetic diversity in genome-wide association studies of lipids. Nature. 2021;600(7890):675-679. doi:10.1038/s41586-021-04064-3

[9] Wang Y, Guo J, Ni G, Yang J, Visscher PM, Yengo L. Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations. Nat Commun. 2020;11:3865. doi:10.1038/s41467-020-17719-y

[10] Privé F, Aschard H, Carmi S, Folkersen L, Hoggart C, O'Reilly PF, et al. Portability of 245 polygenic scores when derived from the UK Biobank and applied to 9 ancestry groups from the same cohort. Am J Hum Genet. 2022;109(1):12-23. doi:10.1016/j.ajhg.2021.11.008

[11] Ding Y, Hou K, Xu Z, Pimplaskar A, Petter E, Boulier K, et al. Polygenic scoring accuracy varies across the genetic ancestry continuum. Nature. 2023;618(7966):774-781. doi:10.1038/s41586-023-06079-4

[12] Kachuri L, Chatterjee N, Hirbo J, Schaid DJ, Martin I, Kullo IJ, et al. Principles and methods for transferring polygenic risk scores across global populations. Nat Rev Genet. 2024;25(1):8-25. doi:10.1038/s41576-023-00637-2

[13] Hunter-Zinck H, Musharoff S, Salit J, Al-Ali KA, Chouchane L, Gohar A, et al. Population genetic structure of the people of Qatar. Am J Hum Genet. 2010;87(1):17-25. doi:10.1016/j.ajhg.2010.05.018

[14] Fakhro KA, Staudt MR, Ramstetter MD, Robay A, Malek JA, Badii R, et al. The Qatar genome: a population-specific tool for precision medicine in the Middle East. Hum Genome Var. 2016;3:16016. doi:10.1038/hgv.2016.16

[15] Mbarek H, Devadoss Gandhi G, Selvaraj S, Al-Muftah W, Badji R, Al-Sarraj Y, et al. Qatar genome: insights on genomics from the Middle East. Hum Mutat. 2022;43(4):499-510. doi:10.1002/humu.24336

[16] Thareja G, Al-Sarraj Y, Belkadi A, Almotawa M, Qatar Genome Program Research (QGPR) Consortium, Suhre K, et al. Whole genome sequencing in the Middle Eastern Qatari population identifies genetic associations with 45 clinically relevant traits. Nat Commun. 2021;12:1250. doi:10.1038/s41467-021-21381-3

[17] Al Kuwari H, Al Thani A, Al Marri A, Al Kaabi A, Abderrahim H, Afifi N, et al. The Qatar Biobank: background and methods. BMC Public Health. 2015;15:1208. doi:10.1186/s12889-015-2522-7

[18] Hebbar P, Abu-Farha M, Abubaker J, Channanath AM, Al-Mulla F, Thanaraj TA. Generalizability of GWA-identified genetic risk variants for metabolic traits to populations from the Arabian Peninsula. Genes (Basel). 2021;12(10):1637. doi:10.3390/genes12101637

[19] Ismail Umlai UK, Toor SM, Al-Sarraj YA, Mohammed S, Al Hail MS, Ullah E, et al. A multi-ancestry genome-wide association study and evaluation of polygenic scores of LDL-C levels. J Lipid Res. 2025;66(3):100752. doi:10.1016/j.jlr.2025.100752

[20] Shim I, Kuwahara H, Chen N, Hashem MO, AlAbdi L, Abouelhoda M, et al. Clinical utility of polygenic scores for cardiometabolic disease in Arabs. Nat Commun. 2023;14:6535. doi:10.1038/s41467-023-41985-1

[21] Wakil SM, Ram R, Muiya NP, Mehta M, Andres E, Mazhar N, et al. A genome-wide association study reveals susceptibility loci for myocardial infarction/coronary artery disease in Saudi Arabs. Atherosclerosis. 2016;245:62-70. doi:10.1016/j.atherosclerosis.2015.11.019

[22] Ram R, Wakil S, Muiya N, Andres E, Mazhar N, Hagos S, et al. A common variant association study in ethnic Saudi Arabs reveals novel susceptibility loci for hypertriglyceridemia. Clin Genet. 2017;91(3):371-378. doi:10.1111/cge.12859

[23] Little J, Higgins JPT, Ioannidis JPA, Moher D, Gagnon F, von Elm E, et al. STrengthening the REporting of Genetic Association Studies (STREGA): an extension of the STROBE statement. PLoS Med. 2009;6(2):e1000022. doi:10.1371/journal.pmed.1000022

[24] Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. Nucleic Acids Res. 2023;51(D1):D977-D985. doi:10.1093/nar/gkac1010

[25] Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37(7):658-665. doi:10.1002/gepi.21758

[26] Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-560. doi:10.1136/bmj.327.7414.557

[27] Wakefield J. Bayes factors for genome-wide association studies: comparison with P-values. Genet Epidemiol. 2009;33(1):79-86. doi:10.1002/gepi.20359

[28] Maller JB, McVean G, Byrnes J, Vukcevic D, Palin K, Su Z, et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. Nat Genet. 2012;44(12):1294-1301. doi:10.1038/ng.2435

[29] Marigorta UM, Navarro A. High trans-ethnic replicability of GWAS results implies common causal variants. PLoS Genet. 2013;9(6):e1003566. doi:10.1371/journal.pgen.1003566

[30] Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, et al. Genetic analyses of diverse populations improves discovery for complex traits. Nature. 2019;570(7762):514-518. doi:10.1038/s41586-019-1310-4

[31] Kanai M, Akiyama M, Takahashi A, Matoba N, Momozawa Y, Ikeda M, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. Nat Genet. 2018;50(3):390-400. doi:10.1038/s41588-018-0047-6

[32] Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs KV, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature. 2010;466(7307):714-719. doi:10.1038/nature09266

[33] Kozlitina J, Smagris E, Stender S, Nordestgaard BG, Zhou HH, Tybjærg-Hansen A, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. Nat Genet. 2014;46(4):352-356. doi:10.1038/ng.2901

[34] Diboun I, Al-Sarraj Y, Toor SM, Mohammed S, Qureshi N, Al Hail MSH, et al. The prevalence and genetic spectrum of familial hypercholesterolemia in Qatar based on whole genome sequencing of 14,000 subjects. Front Genet. 2022;13:927504. doi:10.3389/fgene.2022.927504

[35] Duncan L, Shen H, Gelaye B, Meijsen J, Ressler K, Feldman M, et al. Analysis of polygenic risk score usage and performance in diverse human populations. Nat Commun. 2019;10:3328. doi:10.1038/s41467-019-11112-0

[36] Ruan Y, Lin YF, Feng YCA, Chen CY, Lam M, Guo Z, et al. Improving polygenic prediction in ancestrally diverse populations. Nat Genet. 2022;54(5):573-580. doi:10.1038/s41588-022-01054-7

[37] Kanoni S, Graham SE, Wang Y, Surakka I, Ramdas S, Zhu X, et al. Implicating genes, pleiotropy, and sexual dimorphism at blood lipid loci through multi-ancestry meta-analysis. Genome Biol. 2022;23:268. doi:10.1186/s13059-022-02837-1

[38] Urbut SM, Chen Q, Sui Y, Misra A, Koyama S, Hornsby W, et al. Genetics of cholesterol and coronary disease risk across six global ancestries. NEJM Evid. 2025;4(11):EVIDoa2500105. doi:10.1056/EVIDoa2500105

[39] Thareja G, Belkadi A, Arnold M, Albagha OME, Graumann J, Schmidt F, et al. Differences and commonalities in the genetic architecture of protein quantitative trait loci in European and Arab populations. Hum Mol Genet. 2023;32(6):907-916. doi:10.1093/hmg/ddac243

[40] Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014;10(5):e1004383. doi:10.1371/journal.pgen.1004383