Lipids in Health and Disease

RESEARCH

Open Access

Regular exercise modifies the effects of 12 LDL-C SNPs on cardiovascular disease risk: KoGES

Yon Chul Park[1†], Ahhyun Cho[1†], Yong-Sang Kim[2], Yonghwan Kim[2], Jae-Min Park[3], Kunhee Han[4], Bomtaeck Kim[5], Kwangyoon Kim[5], Junghwa Shin[6], Yu-Jin Kwon[7], Ja-Eun Choi[8] and Kyung-Won Hong[8*]

Abstract

Background Cardiovascular disease risk is significantly driven by dyslipidemia, making it a crucial target for clinical management; however, the translation of low-density lipoprotein cholesterol (LDL-C)-related genetic susceptibility into clinical events is complex. Although regular physical activity improves cardiometabolic health, its effects on LDL-C concentrations are often modest. The present study investigated whether habitual exercise modifies the clinical expression of established LDL-C-associated genetic variants and their association with both cardiovascular and cerebrovascular outcomes among individuals with dyslipidemia.

Methods Data from 20,748 adults with dyslipidemia from the Korea Genome and Epidemiology Study (KoGES) were analyzed. Twelve LDL-C-associated single-nucleotide polymorphisms (SNPs), which were validated across independent Korean cohorts, were utilised to assess genetic risk. Associations with LDL-C levels and clinical outcomes (specifically including overall cardiovascular and cerebrovascular disease, coronary artery disease, and ischemic stroke) were evaluated by using multivariable regression and propensity score matching to minimise self-selection bias. Gene-exercise interactions were assessed by using multiplicative interaction terms.

Results All of the selected SNPs demonstrated robust, directionally consistent associations with LDL-C levels regardless of exercise status. Exercise was associated with improved high-density lipoprotein cholesterol and triglyceride levels but did not substantially alter LDL-C concentrations. Notably, significant gene-exercise interactions for ischemic stroke were identified at the *DHODH/HP* (interaction odds ratio [OR] = 0.65, $P = 0.029$) and *HPR* (interaction OR = 0.65, $P = 0.026$) loci. These interaction signals remained directionally consistent in the propensity score-matched cohort, thus suggesting that exercise could attenuate genetic stroke risk through nonlipid pathways. Although these findings did not persist after stringent Bonferroni correction, they provide exploratory evidence of locus-specific interplay.

Conclusions Although the genetic architecture of LDL-C appears to be stable across physical activity strata, habitual exercise may modify the clinical manifestations of LDL-related genetic susceptibility, particularly regarding ischemic

[†]Yon Chul Park and Ahhyun Cho contributed equally to this work.

*Correspondence:
Kyung-Won Hong
kyungwon.hong@theragenhealth.com

Full list of author information is available at the end of the article

stroke. These findings suggest that regular exercise reduces cerebrovascular risk even in the presence of inherited lipid-related risk. By highlighting the active buffering effect of lifestyle modifications against genetic vulnerability, these results provide actionable evidence for precision prevention strategies, thereby fundamentally supporting broader public health initiatives to reduce the burden of noncommunicable diseases and foster long-term health and well-being.

Keywords Dyslipidemias, Genetic predisposition to disease, Exercise, Ischemic stroke, Gene-environment interaction, Precision medicine

Introduction

Cardiovascular disease (CVD) continues to rank as the foremost contributor to mortality worldwide, persistently imposing a substantial public health burden despite advances in preventive and therapeutic strategies [1]. Acting as a principal preventable determinant in the development and advancement of atherosclerotic plaques, dyslipidemia typically manifests through a combination of hypertriglyceridemia, raised low-density lipoprotein cholesterol (LDL-C), and diminished high-density lipoprotein cholesterol (HDL-C) [2, 3]. Across diverse cohorts, including both Western and Asian demographics, dyslipidemia exhibits a robust correlation with the onset of coronary artery disease (CAD) and ischaemic stroke (IS), highlighting a critical imperative for targeted risk mitigation in these vulnerable individuals [4].

Regular physical activity is a cornerstone of nonpharmacological management for dyslipidemia and CVD prevention. Extensive epidemiological and interventional evidence has demonstrated that exercise improves cardiometabolic health by enhancing insulin sensitivity and reducing systemic inflammation [5, 6]. With respect to lipid metabolism, although exercise consistently improves HDL-C and triglyceride levels, its effects on LDL-C concentrations are often relatively modest and heterogeneous. Meta-analyses have demonstrated that reductions in LDL-C attributable to exercise alone are generally small and highly variable, particularly among older adults and individuals with established metabolic abnormalities [7–9]. Importantly, the cardiovascular benefits of physical activity cannot be fully explained by quantitative changes in circulating LDL-C levels alone. Even in the absence of a substantial reduction in LDL-C levels, regular exercise has been associated with lower risks of cardiovascular morbidity and mortality through pleiotropic mechanisms. These effects include improved endothelial function, reduced oxidative stress, modulation of inflammatory pathways, and beneficial effects on thrombosis and autonomic regulation [6]. These observations suggest that physical activity may modify cardiovascular risk through biological pathways that extend beyond or act downstream of circulating lipid concentrations.

Interindividual variability in lipid profiles and cardiovascular outcomes is significantly affected by genetic susceptibility. Comprehensive genome-wide association studies (GWASs) have uncovered a multitude of single-nucleotide polymorphisms (SNPs) associated with LDL-C levels, thus implicating genes that are involved in hepatic cholesterol synthesis, lipoprotein transport, and LDL receptor-mediated clearance, including *LDLR*, *PCSK9*, *HMGCR*, *APOE*, and *ABO* [10, 11]. Collectively, these variants account for a substantial proportion of the lifelong variation in LDL-C levels and have been linked to cardiovascular risk across diverse populations [12]. In addition to their main effects, genetic variants may interact with lifestyle factors to influence cardiometabolic outcomes through gene-environment interactions [13, 14]. However, most prior studies have focused on interactions involving intermediate metabolic traits (such as lipid concentrations) rather than clinically meaningful cardiovascular outcomes. Moreover, few studies have investigated whether physical activity affects the association between LDL-C-related genetic variants and cardiovascular disease risk in individuals with dyslipidemia. This knowledge gap is particularly relevant in dyslipidemic populations, where circulating LDL-C levels may be pharmacologically or physiologically buffered (e.g., by lipid-lowering therapy and compensatory metabolic regulation), thereby attenuating observable differences in LDL-C levels between lifestyle strata [9, 15].

The hypothesis of the present study is that regular exercise modifies the translation of LDL-related genetic susceptibility into clinical vascular events through LDL-C-independent pathways rather than by substantially reducing LDL-C levels alone. Compared with previously published literature that has predominantly focused on intermediate lipid trait modifications, the novel aspect of this study involves the direct evaluation of how physical activity alters the clinical expression of established LDL-C-associated genetic variants in relation to concrete cardio-cerebrovascular endpoints among a high-risk dyslipidemic population. Accordingly, an investigation was conducted to determine whether habitual exercise modifies cardiovascular and cerebrovascular outcomes by using data from the Korea Genome and Epidemiology Study (KoGES). Specifically, this study focused on 12 LDL-associated SNPs that were selected a priori based

on robust prior GWAS evidence (and subsequently replicated in independent Korean cohorts) to test gene-environment interactions. The clarification of these mechanisms may support the development of individualised cardiovascular prevention strategies.

Materials and methods

Study population

This study derived its sample from the KoGES, a comprehensive genome-based prospective epidemiological cohort. The South Korean government—specifically the Ministry of Health and Welfare, the Korea Disease Control and Prevention Agency, and the National Institute of Health—finances this extensive initiative. Fundamentally, KoGES aims to elucidate how genetic profiles and environmental exposures interact to drive prevalent complex diseases within the Korean demographic. To establish its baseline, the KoGES recruited men and women aged 40 and older, drawing participants from both general community populations and the national health screening program. The dataset encompasses detailed information on anthropometric measurements, blood biochemical profiles, lifestyle factors, and self-reported medical and medication histories. Genome-wide SNP genotype data were available for 72,299 participants (PMID: 27085081).

Participants in the present study were included from multiple KoGES subcohorts, including the KoGES-HEXA (Health Examinees) cohort, in which participants were recruited between 2004 and 2013 from 39 hospitals in metropolitan and major cities across Korea through the national health examination program. In addition, the KoGES_Ansan and Ansung cohorts (which are community-based prospective cohorts established in 2001–2002) enrolled residents of urban (Ansan) and rural (Ansung) areas and conducted biennial follow-up examinations. The study population also included participants from the KoGES-CAVAS (Cardiovascular Disease Association Study) cohort, in which adults residing in rural regions were recruited to investigate risk factors for cardiovascular and metabolic diseases (PMID: 27085081).

In prior analyses, the reproducibility of LDL cholesterol-associated genetic markers was evaluated across all three subcohorts (KoGES-HEXA, Ansan/Ansung, and CAVAS), and only markers that were consistently validated across cohorts were selected for the present study. For the primary analysis examining the impact of physical activity on cardiocerebrovascular outcomes among individuals with dyslipidemia, 58,701 participants from the KoGES-HEXA cohort were included.

Among these participants, those with missing covariate data ($n = 471$) and those with a history of cancer ($n = 2,202$) were excluded. From the remaining population, individuals meeting the criteria for dyslipidemia were identified, thus yielding 20,770 participants. After further exclusion of participants with missing physical activity data, a total of 20,748 participants with dyslipidemia were included in the final analytical cohort.

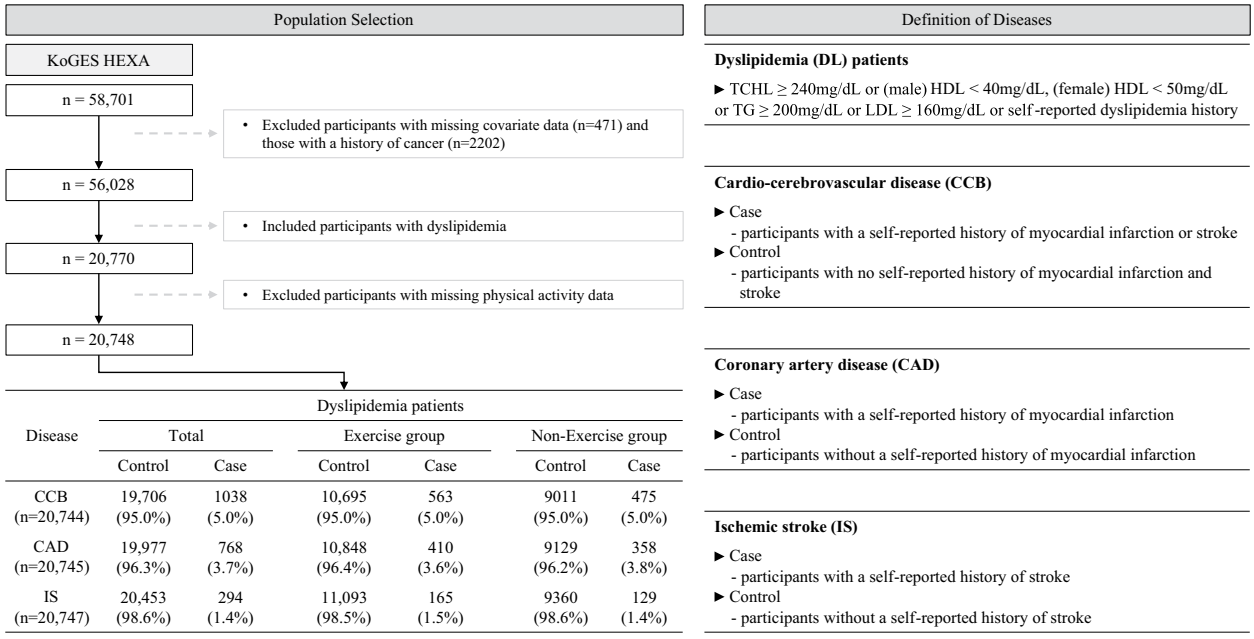
To evaluate the associations between LDL-related SNPs, physical activity status, and cardio-cerebrovascular outcomes, analyses were separately conducted for cardiocerebrovascular disease, coronary artery disease, and ischemic stroke. After excluding outcome-specific missing data, 20,744 participants were included in the CCB analysis, 20,754 participants were included in the CAD analysis, and 20,747 participants were included in the IS analysis. Participants in each outcome-specific cohort were further stratified according to physical activity status for the subgroup analyses.

A schematic overview of the participant selection process and overall study design is presented in Figs. 1 and 2.

All of the study procedures were conducted in accordance with the ethical principles of the Declaration of Helsinki (1975). Written informed consent was obtained from all of the participants prior to enrollment. The study protocol was approved by the Institutional Review Board of TheragenEtex (approval number: 700062–20190819-GP-006–02).

Clinical measurements and disease definitions

Anthropometric measurements were performed by well-trained medical staff according to standardized protocols. Following a minimum 8-h fasting period, qualified medical personnel collected venous blood samples. Body mass index (BMI) was calculated by dividing a participant's weight in kilograms by the square of their height in meters (kg/m^2). To assess waist circumference (WC), measurements were taken precisely at the midpoint between the iliac crest and the inferior margin of the lowest rib. Finally, LDL-C concentrations were estimated via the Friedewald formula: TCHL – HDL-C – (TG/5). Lifestyle information was obtained through interviewer-administered questionnaires. The baseline survey stratified alcohol intake as never, previous, or current consumption. Furthermore, individuals in the current category reported specific details about their drinking patterns over the trailing year. Smoking status was classified as never, former, or current smoker. Physical activity status was assessed by using a self-reported questionnaire, and participants were classified into an exercise or non-exercise group based on regular physical activity. For statistical analyses, exercise status was treated as a binary variable (exercise = 1, non-exercise = 0). To ensure measurement stability, assessors took no fewer than two seated measurements for both systolic (SBP) and diastolic blood pressure (DBP), applying the calculated means to the statistical models. A diagnosis of dyslipidemia for this study was established based on a self-reported history of the condition, or the presence of elevated lipid markers



Disease	Dyslipidemia patients					
	Total		Exercise group		Non-Exercise group	
	Control	Case	Control	Case	Control	Case
CCB (n=20,744)	19,706 (95.0%)	1038 (5.0%)	10,695 (95.0%)	563 (5.0%)	9011 (95.0%)	475 (5.0%)
CAD (n=20,745)	19,977 (96.3%)	768 (3.7%)	10,848 (96.4%)	410 (3.6%)	9129 (96.2%)	358 (3.8%)
IS (n=20,747)	20,453 (98.6%)	294 (1.4%)	11,093 (98.5%)	165 (1.5%)	9360 (98.6%)	129 (1.4%)

Fig. 1 Flowchart of participant selection and phenotype definitions. From 58,701 KoGES HEXA participants, individuals with missing covariates (smoking, drinking, exercise, and BMI parameters; $n = 471$) and those with a history of cancer or no response ($n = 2,202$) were excluded, leaving 56,028 participants. Dyslipidemia (DL) cases were identified ($n = 20,788$), and 18 cases with missing blood pressure data were excluded, leaving 20,770 DL participants for analysis. Right panels summarise operational definitions: DL criteria (TCHL ≥ 200 mg/dL, HDL < 40 mg/dL or < 50 mg/dL, TG ≥ 150 mg/dL, LDL ≥ 130 mg/dL, or prior DL), CVD outcomes (CCb: myocardial infarction or stroke; CAD: myocardial infarction; IS: stroke), and covariates (sex, age, BMI, smoking, drinking, PC1, and PC2). The table at the bottom shows overall case/control counts and categorisation according to exercise status (Exercise O vs. Exercise X)

(TCHL ≥ 240 mg/dL, LDL-C ≥ 160 mg/dL, TG ≥ 200 mg/dL) and reduced HDL-C (under 40 mg/dL in men or under 50 mg/dL in women). CCB was defined based on self-reported medical history. Participants who reported a history of myocardial infarction or stroke were defined as cases, whereas those who reported no history of myocardial infarction or stroke were defined as controls. CAD was defined as follows: cases included participants with a self-reported history of myocardial infarction, and controls involved participants without such a history. IS was defined as follows: cases included participants with a self-reported history of stroke, and controls involved participants without a self-reported history of stroke.

Genotyping

The dataset includes integration of administrative health records with a broad spectrum of environmental exposures and phenotypic traits. Furthermore, it incorporates genome-wide genetic profiles alongside various biospecimens, such as urine, serum, plasma, and DNA. Following standardized protocols, clinical personnel drew peripheral blood specimens and safely deposited them at the National Biobank of Korea for long-term archiving and future research applications. To conduct the genetic analysis, we isolated genomic DNA from the peripheral blood specimens and performed genotyping via the Korea Biobank Array (KoreanChip), the detailed design and content of which have been previously described.

Quality control procedures were applied to ensure the reliability of the genotyping data. To ensure data quality, we filtered out SNPs that demonstrated a minor allele frequency (MAF) < 1%, a genotype missing rate > 1%, or a call rate < 97%. Additionally, we removed variants showing significant deviation from Hardy–Weinberg equilibrium ($p < 1 \times 10^{-6}$) (PMID: 30,718,733).

Preliminary investigation for SNP selection and replication

To identify robust and reproducible genetic variants associated with LDL cholesterol levels, a two-step SNP selection strategy was employed. First, SNPs that have been previously reported to be associated with LDL-C were identified through a comprehensive literature review of GWAS. For each candidate SNP, prior evidence regarding effect direction and magnitude was evaluated to establish biological plausibility.

Second, genome-wide association analyses were conducted in the KoGES-HEXA cohort, which included approximately 7.9 million SNPs. Linear regression models were applied to assess associations between SNPs and LDL-C levels, adjusting for age, sex, and the first two principal components to account for demographic factors and population stratification. A genome-wide significance threshold of $P < 5 \times 10^{-8}$ was applied. SNPs meeting this criterion and those that were previously reported in external GWASs were selected as candidate LDL-C-associated variants.

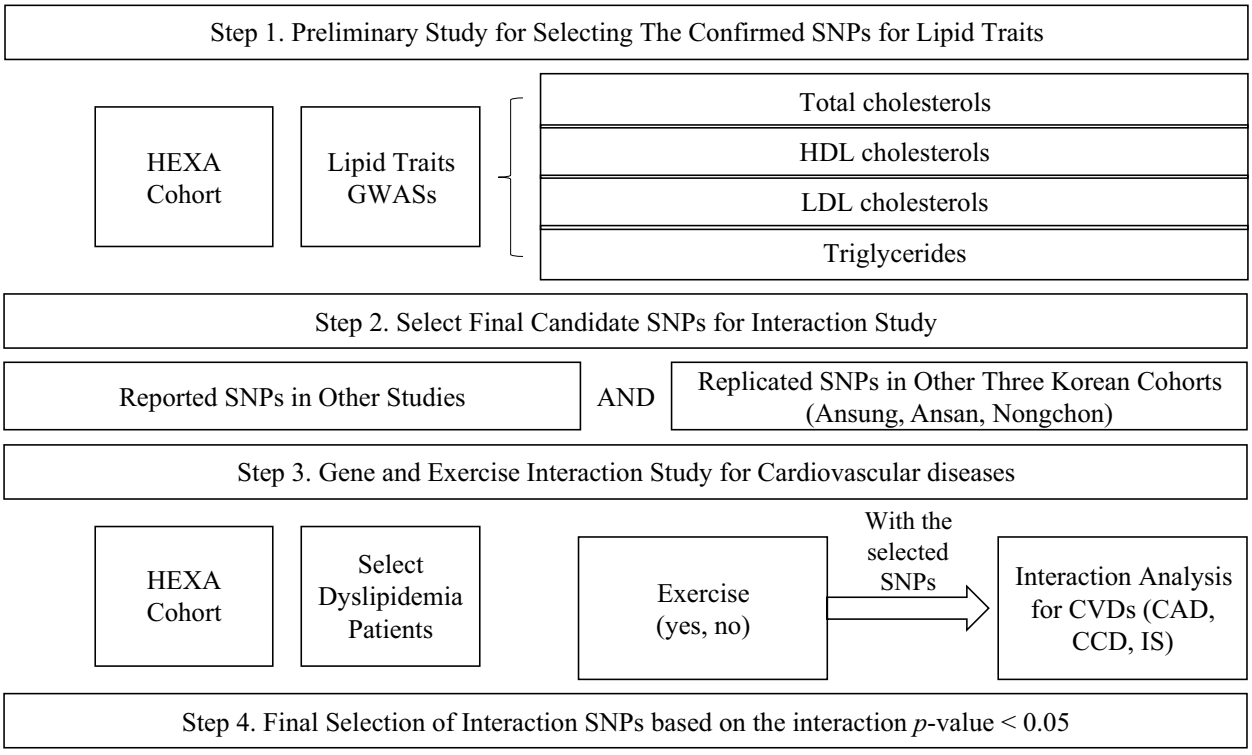


Fig. 2 Analysis pipeline for SNP selection and gene-exercise interaction testing. Step 1: Preliminary identification of lipid-trait SNPs (total cholesterol, HDL-C, LDL-C, and triglycerides) in the HEXA cohort. Step 2: Final sentinel selection based on previously reported SNPs and replication in three Korean cohorts (Ansan, Ansung, and Nongchon). Step 3: Interaction testing among dyslipidemia patients in HEXA utilising regular exercise (yes/no) for the outcomes of CAD, CCb, and IS. Step 4: Final interaction SNPs defined at a two-sided $P < 0.05$. CAD, coronary artery disease; CCb, cardio-cerebrovascular disease; IS, ischemic stroke

Replication analyses were subsequently performed in independent KoGES subcohorts (including the Ansan, Ansung, and Nongchon cohorts) using the same analytical framework. SNPs demonstrating statistically significant associations with LDL-C and consistent effect directions across these cohorts were considered to be successfully replicated and retained for downstream analyses. Detailed information on prior GWAS reports and replication results is provided in the supplementary SNP table.

Construction of the genetic risk score (GRS)

A genetic risk score (GRS) was constructed by using 12 LDL cholesterol-associated SNPs that were identified in prior genome-wide association studies of the KoGES HEXA cohort. For each SNP, the reported effect size (β coefficient) from the external KoGES HEXA GWAS was used as the weight. The GRS was calculated as a weighted sum of risk alleles according to the following formula:

$$\mathrm{GRS} = \sum_{\mathrm{i}=1}^{12} \beta_{\mathrm{i}} \times \text{risk allele count}_{\mathrm{i}}$$

where the risk allele count was coded as 0, 1, or 2 under an additive genetic model. This approach ensured that SNP selection and weighting were based on independent prior evidence, thereby minimising overfitting and avoiding circular inference.

Propensity score matching

To minimize potential confounding arising from baseline disparities between the exercise and non-exercise groups, propensity score matching (PSM) was performed. Propensity scores were estimated through a logistic regression model, incorporating age, sex, BMI, smoking history, and drinking history as covariates. The matching process utilized a one-to-one nearest-neighbor approach without replacement, with the caliper width set at 0.2 standard deviations of the logit of the propensity score. Standardized mean differences (SMDs) served as the metric for assessment; specifically, an absolute SMD < 0.1 signified that the covariates had reached an acceptable balance between groups.

Statistical analysis

Power analysis to justify the sample size was conducted by using Quanto (version 1.2.4). Estimation of the required sample size relied on an additive genetic model

incorporating a two-sided significance level of 0.05, an interaction odds ratio of 1.2, and a minor allele frequency of 0.02. This calculation accounted for a cardiovascular disease incidence of approximately 2% among those with dyslipidemia, a figure aligned with the *Dyslipidemia Fact Sheet in Korea 2024* by the Korean Society of Lipid and Atherosclerosis. This analysis was conducted to assess the feasibility of detecting gene-exercise interaction effects in patients with dyslipidemia. To evaluate the associations between genetic risk and cardiovascular outcomes, multivariable logistic regression analyses were performed for CCB, CAD, and IS. The primary predictors included the GRS, exercise status, and their interaction term (GRS × exercise). Statistical adjustment for age, sex, BMI, and histories of smoking and drinking was applied across all models. To evaluate the predictive efficacy of the GRS-based models, the analysis utilized receiver operating characteristic (ROC) curves. The predicted probabilities were obtained through fivefold stratified cross-validation, and the area under the ROC curve (AUC) was calculated to quantify model discrimination. The optimal probability threshold for classification was determined by using the Youden index, which was defined as the point that maximises the sum of sensitivity and specificity minus one. ROC curves were generated by using the roc_curve function from the sklearn.metrics package. For analyses at the individual SNP level, 12 LDL-related SNPs were examined. Linear regression models examined the potential associations between individual SNPs and LDL cholesterol levels. These assessments incorporated stratification by exercise status and accounted for several covariates, including age, sex, BMI, and status of smoking and alcohol, alongside the first two principal components (PC1 and PC2). Logistic regression analyses were used to examine associations between SNPs and cardiovascular outcomes (CCB, CAD, and IS). To mitigate potential bias from population stratification, the study utilized principal component analysis (PCA) on the genome-wide genotype data. Accordingly, the first two principal components (PC1 and PC2) served as essential covariates within the regression frameworks to ensure genomic stability. Genotype × exercise interaction terms were included to test for multiplicative interaction effects. All of the SNP-level analyses were performed by using PLINK v1.9.

Interaction analyses were conducted for the 12 LDL-associated SNPs that were selected a priori based on prior GWAS evidence. To control for multiple testing, Bonferroni correction was applied as a confirmatory threshold ($\alpha = 0.05/12$). In addition, nominal P values (< 0.05) were examined within an exploratory framework to identify potential interaction signals, given the limited statistical power for interaction testing (particularly for low-frequency variants). Stratified analyses by exercise status were performed to facilitate interpretation of interaction effects and to examine the direction and consistency of the SNP associations within each stratum. These analyses were descriptive; therefore, they were evaluated by using nominal P values. Given the hypothesis-driven nature of this targeted SNP analysis, nominal P values are presented for interpretability, with Bonferroni-corrected results being additionally provided as a sensitivity analysis to control for the familywise error rate.

Results

Study population and baseline characteristics

A total of 20,748 participants with dyslipidemia from the KoGES-HEXA cohort were included in the final analysis (Fig. 1). Among them, 11,258 participants (54.3%) reported engaging in regular exercise, whereas 9,490 participants (45.7%) were classified as non-exercisers.

At baseline, participants in the exercise group were slightly older than those in the non-exercise group (55.7 ± 7.4 vs. 54.9 ± 7.9 years, $P < 0.001$), and a lower proportion of participants in the exercise group were women (55.9% vs. 61.3%, $P < 0.001$). Body mass index and waist circumference were modestly but significantly lower among exercisers (BMI: 24.6 ± 2.7 vs. 24.7 ± 2.9 kg/m^2; waist circumference: 83.0 ± 8.5 vs. 83.7 ± 8.8 cm; both $P < 0.001$). With respect to biochemical traits, exercisers exhibited lower triglyceride levels (171.6 ± 109.0 vs. 182.1 ± 119.8 mg/dL, $P < 0.001$) and higher HDL cholesterol levels (49.0 ± 14.1 vs. 47.7 ± 13.4 mg/dL, $P < 0.001$), whereas LDL cholesterol levels did not significantly differ between the two groups (128.3 ± 41.6 vs. 129.2 ± 41.9 mg/dL, $P = 0.142$). The prevalence of cardio-cerebrovascular disease, coronary artery disease, and ischemic stroke was comparable between exercisers and non-exercisers (all $P > 0.05$) (Table 1).

Baseline characteristics after propensity score matching

The propensity score matching process, which accounted for sex, age, body mass index, smoking status, and drinking status, yielded 9,104 matched pairs of exercisers and non-exercisers were generated. Most of the baseline characteristics were well balanced between the two groups, including age (55.1 ± 7.3 vs. 55.2 ± 7.8 years, $P = 0.475$), body mass index (24.7 ± 2.8 vs. 24.7 ± 2.9 kg/m^2, $P = 0.804$), and systolic and diastolic blood pressure (both $P > 0.1$). Lipid profiles were also largely comparable after matching, except for the observation that HDL cholesterol levels remained higher in the exercise group (49.4 ± 14.0 vs. 47.8 ± 13.4 mg/dL, $P < 0.001$) and that triglyceride levels remained lower among exercisers (171.0 ± 109.1 vs. 181.5 ± 119.4 mg/dL, $P < 0.001$). The prevalence of CCB, CAD, and IS did not significantly differ between groups in the matched cohort (all $P > 0.1$),

Table 1 Baseline characteristics of the study population

Variables	Dyslipidemia Patients			*P*
	Total (*n* = 20,748)	Exercise group (*n* = 11,258)	Non-exercise group (*n* = 9490)	
Female (n, %)	12,110 (58.4)	6293 (55.9)	5817 (61.3)	< 0.001
Age (mean ± sd)	55.3 ± 7.6	55.7 ± 7.4	54.9 ± 7.9	< 0.001
Cardio-cerebrovascular disease (n, %)	1038 (5.0)	563 (5.0)	475 (5.0)	1.000
Coronary artery disease (n, %)	768 (3.7)	410 (3.6)	358 (3.8)	0.643
Ischemic stroke (n, %)	294 (1.4)	165 (1.5)	129 (1.4)	0.558
Anthropometric traits				
Body mass index (kg/m^2)	24.7 ± 2.8	24.6 ± 2.7	24.7 ± 2.9	< 0.001
Waist circumference (cm)	83.3 ± 8.7	83.0 ± 8.5	83.7 ± 8.8	< 0.001
Systolic blood pressure (mmHg)	125 ± 14.8	125.3 ± 14.8	124.7 ± 14.9	0.004
Diastolic blood pressure (mmHg)	77.3 ± 9.7	77.4 ± 9.7	77.2 ± 9.8	0.074
Biochemical traits				
Fasting plasma glucose (mg/dL)	96.1 ± 28.3	96.1 ± 27.9	96.1 ± 28.7	0.946
Total cholesterol (mg/dL)	212.2 ± 43.9	211.5 ± 43.9	213.0 ± 44.0	0.011
HDL cholesterol (mg/dL)	48.4 ± 13.8	49.0 ± 14.1	47.7 ± 13.4	< 0.001
Triglyceride (mg/dL)	176.4 ± 114.2	171.6 ± 109.0	182.1 ± 119.8	< 0.001
LDL cholesterol (mg/dL)	128.7 ± 41.7	128.3 ± 41.6	129.2 ± 41.9	0.142
Lifestyle factors				
Smoking status (N, %):				
Never	13,870 (66.9)	7472 (66.4)	6398 (67.4)	< 0.001
Quit	3847 (18.5)	2397 (21.3)	1450 (15.3)	
Current	3031 (14.6)	1389 (12.3)	1642 (17.3)	
Drinking status (N, %):				
Never	10,578 (51.0)	5449 (48.4)	5129 (54.0)	< 0.001
Quit	890 (4.3)	509 (4.5)	381 (4.0)	
Current	9280 (44.7)	5300 (47.1)	3980 (41.9)	

Values are presented as the means ± standard deviations for continuous variables and as numbers (percentages) for categorical variables. Dyslipidemia was defined based on lipid profiles (TCHL ≥ 240 mg/dL, HDL-C < 40 mg/dL in men or < 50 mg/dL in women, TG ≥ 200 mg/dL, or LDL-C ≥ 160 mg/dL) or a self-reported history of the condition. Exercise status was operationalised as a binary variable based on self-reported regular physical activity. *P* values were calculated by using independent t tests for continuous variables and chi-square tests for categorical variables to compare the exercise and non-exercise groups

thus supporting a reasonable covariate balance for the subsequent interaction analyses (Table S1).

Selection and replication of LDL cholesterol-associated SNPs

A total of 12 LDL cholesterol-associated SNPs were selected based on prior GWAS evidence and genome-wide association analyses conducted in the KoGES-HEXA cohort (Fig. 2). All of the selected variants exceeded the genome-wide significance threshold ($P < 5 \times 10^{-8}$) in the HEXA cohort and demonstrated directionally consistent associations with LDL cholesterol levels. Replication analyses across independent Korean cohorts (including Ansan, Ansung, and Nongchon cohorts) confirmed statistically significant associations for all of the selected SNPs with consistent effect estimates across cohorts. These variants were mapped to biologically relevant loci, including *PCSK9, CELSR2, HMGCR, LDLR, APOE, TOMM40, DHODH/HP, HPR*, and *ABO*, thereby supporting their robustness and suitability for downstream gene-exercise interaction analyses (Table S2).

Associations between LDL-related SNPs and ldl cholesterol levels

Among patients with dyslipidemia, most of the selected SNPs demonstrated strong and directionally consistent associations with LDL cholesterol levels in both the exercise and non-exercise groups (Table 2). The *PCSK9* variant rs151193009 exhibited the greatest effect size, with the T allele being associated with marked LDL reductions in both exercisers ($\beta = -22.65 \pm 2.58$ mg/dL, $P = 1.71 \times 10^{-18}$) and non-exercisers ($\beta = -15.55 \pm 3.09$ mg/dL, $P = 4.99 \times 10^{-7}$). Similarly, variants in *CELSR2, HMGCR, LDLR, TOMM40*, and *APOE* exhibited significant associations with LDL cholesterol in both exercise strata, thus indicating that the genetic architecture of LDL cholesterol was largely preserved regardless of exercise status. In contrast, ABO rs532436 was associated with higher LDL cholesterol levels in both exercisers ($\beta = +3.70 \pm 0.61$ mg/dL, $P = 1.10 \times 10^{-9}$) and non-exercisers ($\beta = +3.16 \pm 0.67$ mg/dL, $P = 2.14 \times 10^{-6}$). Notably, *DHODH/HP* rs77303550 and *HPR* rs217184 demonstrated statistically significant LDL-lowering associations only in the non-exercise group, thereby suggesting possible heterogeneity according to exercise status.

Table 2 Genetic effects of LDL cholesterol-associated variants on LDL levels and their interactions with exercise habits on CCB, IS, and CAD risk in patients with dyslipidemia

Target SNPs of LDL cholesterol	Chr:bp	Alleles		Positional candidate gene(s)	ALT freq.	Other Ethnics			Coding Allele	Coding Allele Frequency	LDL cholesterol Association results in Dyslipidemia Patients only				*Second Disease*	Cardio-cerebrovascular disease							
											Exercise group only (n=10,730)		Non Exercise group only (n=10,040)			Candidate gene Association Study (Main effect)						Interaction	
																Exercise group only (n=10,730)			Non Exercise group only (n=10,040)			Gene and Excercise Interaction test	
		REF	ALT			ASN	EUR	AFR			Beta (SE)	*p*	Beta(SE)	*p*		*case*	OR (95% CI)	*p*	*case*	OR (95% CI)	*p*	OR (95% CI)	*p*
rs151193009	1:55509585	C	T	PCSK9	0.02	0.01	0.00	0.00	T	0.02	-22.65 (2.577)	**1.71E-18**	-15.55 (3.091)	**4.99E-07**	CCb	548	1.373 (0.828-2.276)	2.19E-01	491	1.187 (0.631-2.232)	5.94E-01	1.072 (0.466-2.468)	8.70E-01
															IS	150	0.808 (0.256-2.551)	7.16E-01	145	1.201 (0.376-3.835)	7.57E-01	-	-
															CAD	410	1.558 (0.898-2.705)	1.15E-01	358	1.452 (0.749-2.814)	2.70E-01	1.052 (0.429-2.578)	9.12E-01
rs12740374	1:109817590	G	T	CELSR2	0.06	0.04	0.21	0.20	T	0.06	-7.509 (1.206)	**4.99E-10**	-9.466 (1.367)	**4.63E-12**	CCb	548	0.897 (0.681-1.181)	4.37E-01	491	1.141 (0.855-1.523)	3.72E-01	0.77 (0.512-1.158)	2.09E-01
															IS	150	0.708 (0.411-1.221)	2.14E-01	145	0.592 (0.291-1.203)	1.47E-01	-	-
															CAD	410	0.976 (0.716-1.329)	8.75E-01	358	1.344 (0.986-1.832)	6.13E-02	0.724 (0.464-1.129)	1.54E-01
rs660240	1:109817838	T	C	CELSR2	0.93	0.95	0.80	0.61	T	0.06	-7.179 (1.206)	**2.69E-09**	-9.263 (1.353)	**7.99E-12**	CCb	548	0.886 (0.672-1.167)	3.88E-01	491	1.142 (0.857-1.521)	3.64E-01	0.757 (0.503-1.137)	1.80E-01
															IS	150	0.700 (0.406-1.206)	1.99E-01	145	0.579 (0.285-1.177)	1.31E-01	-	-
															CAD	410	0.965 (0.708-1.315)	8.19E-01	358	1.352 (0.995-1.838)	5.40E-02	0.709 (0.454-1.106)	1.29E-01
rs3846661	5:74639178	A	G	HMGCR	0.51	0.50	0.42	0.93	A	0.48	-2.569 (0.547)	**2.72E-06**	-2.791 (0.601)	**3.45E-06**	CCb	548	0.981 (0.867-1.108)	7.52E-01	491	0.903 (0.79-1.032)	1.33E-01	1.083 (0.902-1.299)	3.93E-01
															IS	150	0.845 (0.677-1.054)	1.35E-01	145	0.872 (0.681-1.117)	2.79E-01	-	-
															CAD	410	1.034 (0.898-1.191)	6.43E-01	358	0.933 (0.802-1.086)	3.72E-01	1.098 (0.891-1.354)	3.79E-01
rs12654264	5:74648603	T	A	HMGCR	0.48	0.51	0.39	0.40	A	0.48	-2.622 (0.548)	**1.74E-06**	-2.761 (0.601)	**4.36E-06**	CCb	548	0.981 (0.868-1.109)	7.59E-01	491	0.903 (0.791-1.032)	1.35E-01	1.082 (0.902-1.299)	3.96E-01
															IS	150	0.853 (0.683-1.064)	1.58E-01	145	0.883 (0.689-1.131)	3.24E-01	-	-
															CAD	410	1.031 (0.895-1.188)	6.76E-01	358	0.929 (0.799-1.082)	3.44E-01	1.10 (0.892-1.355)	3.74E-01
rs532436	9:136149830	G	A	ABO	0.26	0.19	0.19	0.12	A	0.26	3.696 (0.606)	**1.10E-09**	3.16 (0.666)	**2.14E-06**	CCb	548	**0.863 (0.75-0.993)**	**4.01E-02**	491	0.962 (0.829-1.116)	6.05E-01	0.884 (0.718-1.087)	2.41E-01
															IS	150	0.813 (0.63-1.051)	1.14E-01	145	1.102 (0.842-1.442)	4.79E-01	0.727 (0.500-1.057)	9.53E-02
															CAD	410	0.879 (0.748-1.034)	1.20E-01	358	0.891 (0.75-1.058)	1.88E-01	0.976 (0.768-1.24)	8.43E-01
rs77303550	16:72079657	C	T	DHODH,HP	0.23	0.24	0.20	0.15	T	0.25	-1.212 (0.642)	5.90E-02	-3.725 (0.708)	**1.46E-07**	CCb	548	0.921 (0.794-1.067)	2.74E-01	491	1.054 (0.902-1.232)	5.09E-01	0.896 (0.722-1.113)	3.21E-01
															IS	150	0.808 (0.613-1.065)	1.31E-01	145	1.266 (0.959-1.671)	9.59E-02	**0.645 (0.435-0.955)**	**2.87E-02**
															CAD	410	0.943 (0.795-1.118)	4.97E-01	358	0.993 (0.829-1.189)	9.36E-01	0.988 (0.77-1.267)	9.22E-01
rs217184	16:72105965	T	C	HPR	0.26	0.24	0.21	0.35	C	0.26	-1.176 (0.632)	6.28E-02	-3.417 (0.695)	**9.07E-07**	CCb	548	0.906 (0.783-1.048)	1.84E-01	491	1.034 (0.886-1.206)	6.72E-01	0.898 (0.725-1.113)	3.26E-01
															IS	150	0.828 (0.633-1.084)	1.71E-01	145	1.294 (0.986-1.699)	6.31E-02	**0.646 (0.439-0.949)**	**2.60E-02**
															CAD	410	0.913 (0.771-1.081)	2.91E-01	358	0.956 (0.800-1.143)	6.23E-01	0.993 (0.776-1.272)	9.58E-01
rs2738464	19:11242307	G	C	LDLR	0.69	0.72	0.87	0.62	G	0.29	-3.723 (0.611)	1.16E-09	-3.942 (0.665)	**3.16E-09**	CCb	548	0.870 (0.757-1.000)	5.05E-02	491	0.966 (0.832-1.121)	6.49E-01	0.898 (0.731-1.104)	3.07E-01
															IS	150	**0.76 (0.587-0.984)**	**3.72E-02**	145	1.026 (0.780-1.348)	8.56E-01	-	-
															CAD	410	0.948 (0.809-1.111)	5.10E-01	358	0.957 (0.807-1.135)	6.10E-01	1.001 (0.791-1.266)	9.94E-01
rs7254892	19:45389596	G	A	TOMM40	0.06	0.08	0.03	0.03	A	0.06	-14.33 (1.27)	**2.29E-29**	-16.22 (1.401)	**9.04E-31**	CCb	548	0.734 (0.533-1.01)	5.73E-02	491	0.905 (0.657-1.246)	5.41E-01	0.847 (0.540-1.330)	4.71E-01
															IS	150	0.714 (0.399-1.278)	2.56E-01	145	0.791 (0.420-1.489)	4.67E-01	-	-
															CAD	410	0.753 (0.521-1.088)	1.30E-01	358	0.986 (0.692-1.404)	9.38E-01	0.804 (0.483-1.337)	4.00E-01
rs769446	19:45408628	T	C	APOE	0.05	0.08	0.07	0.05	C	0.07	-15.31 (1.174)	**1.46E-38**	-16.18 (1.304)	**4.53E-35**	CCb	548	0.773 (0.577-1.036)	8.47E-02	491	0.869 (0.642-1.178)	3.67E-01	0.927 (0.609-1.411)	7.23E-01
															IS	150	0.736 (0.430-1.262)	2.65E-01	145	0.608 (0.312-1.185)	1.44E-01	-	-
															CAD	410	0.794 (0.567-1.112)	1.79E-01	358	1.001 (0.720-1.39)	9.97E-01	0.834 (0.522-1.331)	4.46E-01
rs7412	19:45412079	C	T	APOE	0.06	0.10	0.06	0.05	T	0.06	-16.47 (1.204)	**3.17E-42**	-17.46 (1.326)	**3.05E-39**	CCb	548	0.759 (0.562-1.026)	7.32E-02	491	0.855 (0.627-1.167)	3.24E-01	0.926 (0.603-1.423)	7.26E-01
															IS	150	0.773 (0.452-1.325)	3.49E-01	145	0.628 (0.322-1.223)	1.71E-01	-	-
															CAD	410	0.761 (0.536-1.08)	1.26E-01	358	0.976 (0.697-1.367)	8.87E-01	0.822 (0.508-1.330)	4.23E-01

Plus-minus values represent regression coefficients ± standard errors (SEs) for LDL-cholesterol associations or odds ratios (ORs) with 95% confidence intervals for cardio-cerebrovascular outcomes. Analyses were stratified by exercise status (regular exercise vs. no exercise) and adjusted for age, sex, and body mass index. The interaction column shows multiplicative interaction terms between genotype and exercise in logistic regression models. *P* values < 0.05 were considered to be statistically significant

CCb Cardio-cerebrovascular disease, *CAD* Coronary artery disease, *IS* Ischemic stroke, *BP* Blood pressure

Main genetic effects on cardio-cerebrovascular outcomes

Among the 12 LDL-associated variants, most SNPs were not significantly associated with CCB, CAD, or IS after stratification by exercise status (Table 2). Nonetheless, several locus-specific associations were observed. For *ABO* rs532436, the A allele was associated with a lower risk of CCB among exercisers (OR = 0.863, 95% CI: 0.723–0.999, *P* = 0.048), whereas no corresponding association was observed among non-exercisers. In addition, *LDLR* rs2738464 demonstrated a protective association with ischemic stroke in the exercise group (OR = 0.760, 95% CI: 0.587–0.984, *P* = 0.037), whereas no significant effect was detected in the non-exercise group. Other variants, including those in *PCSK9, HMGCR, APOE*, and *TOMM40*, did not exhibit statistically significant associations with CCB, CAD, or IS in either exercise stratum.

Gene-exercise interaction effects on cardio-cerebrovascular risk

Formal interaction analyses revealed nominally significant SNP × exercise interactions for ischemic stroke at two loci. For *DHODH/HP* rs77303550, a significant interaction was observed for IS risk (interaction OR = 0.645, *P* = 0.0287), thus indicating that regular exercise attenuated the genetic risk associated with the T allele. A similar interaction was observed for *HPR* rs217184, with an interaction OR of 0.646 (*P* = 0.0260) being reported for IS. No significant interaction effects were detected for CCB or CAD at these loci in the primary analysis.

These findings remained directionally consistent in the propensity score-matched cohort. In the matched analysis, the interaction *P* values for ischemic stroke were 0.022 for rs77303550 and 0.0185 for rs217184, thereby supporting the robustness of these signals despite attenuation after multiple-testing correction (Table S3). Additional GRS-based analyses revealed modest associations with CCB and IS in the exercise group; however, the GRS × exercise interaction itself was not statistically significant for any of the three outcomes (Table S4).

In the age-stratified analyses, the nominal interaction signals for ischemic stroke were most apparent in participants in their 50 s for both rs77303550 and rs217184, whereas no consistent interaction pattern was observed across all age groups (Table S5). Sex-stratified analyses suggested that several interaction signals were more evident in women than in men, particularly for *DHODH/HP* and *HPR*-related loci, although the overall pattern remained exploratory (Table S6). In obesity-stratified analyses, nominally significant interactions for ischemic stroke at rs77303550 and rs217184 were more

pronounced in the obese subgroup, whereas corresponding associations were not consistently observed in the non-obese subgroup (Table S7). Additional sensitivity analyses adjusting for baseline HDL-C and triglyceride levels revealed that the SNP-exercise interactions for ischemic stroke remained significant for rs77303550 and rs217184, thus suggesting that these associations were not solely explained by exercise-associated lipid changes (Table S8).

The results of the receiver operating characteristic analyses of the GRS-based models are presented in Figure S1. Overall, these supplementary analyses supported the main finding that the exercise-related modification of genetic susceptibility was more evident for ischemic stroke than for CCB or CAD and appeared to be locus specific (rather than being generalised across all LDL-related variants).

Discussion

The present study provides novel evidence demonstrating that regular exercise mitigates the downstream cerebrovascular consequences of inherited lipid-related risk (independent of circulating LDL-C levels). By evaluating a large population-based cohort of Korean adults with dyslipidemia, the current investigation offers a more comprehensive view of gene-lifestyle interplay beyond intermediate lipid traits. Consistent with prior genome-wide association studies, all of the selected SNPs exhibited robust and directionally consistent associations with LDL-C, thus confirming the reproducibility of established lipid-related loci in an East Asian population [10–12].

Despite strong genetic effects on LDL-C, most variants were not independently associated with cardiovascular or cerebrovascular disease outcomes. This finding aligns with prior evidence indicating that the translation of lipid-related genetic effects into clinical cardiovascular events is often modest and context dependent, particularly in populations with established metabolic risk factors or in those individuals receiving lipid-lowering therapy [15, 16]. These observations underscore the multifactorial nature of cardiovascular risk beyond circulating LDL-C levels alone. Importantly, the hypothesis of the present study concerns the modification of the clinical expression of genetic susceptibility (rather than the actual alteration of LDL-C concentrations).

Habitual exercise did not substantially modify the direct genetic regulation of LDL-C. Effect estimates for LDL-associated SNPs were largely comparable between exercisers and non-exercisers, thus suggesting that the genetic architecture underlying LDL-C is relatively stable across physical activity strata. This results is consistent with the findings of intervention and observational studies demonstrating that exercise exerts relatively modest and heterogeneous effects on LDL-C levels compared with its more pronounced influence on HDL cholesterol and triglyceride levels [5, 17–19]. The observation that cardiovascular benefits may occur without a substantial reduction in LDL-C levels supports the idea that physical activity can counteract genetic vulnerability through LDL-C-independent mechanisms.

Notably, although the main genetic effects on cardiovascular outcomes were largely negligible, nominally significant gene-exercise interactions for ischemic stroke were identified at the DHODH/HP and HPR loci. These findings suggest that regular exercise may attenuate genetically mediated susceptibility to IS among individuals with dyslipidemia, even when LDL-C levels themselves are not substantially altered. Importantly, these interaction signals remained directionally consistent in the propensity score-matched cohort, thereby diminishing concerns about residual confounding due to baseline imbalances.

The biological plausibility of these observations is supported by established vascular and mitochondrial mechanisms. Exercise improves endothelial function, reduces oxidative stress, enhances mitochondrial efficiency, and modulates inflammatory pathways, which are processes that are particularly relevant to cerebrovascular integrity [5, 20–22]. *DHODH* plays a key role in mitochondrial respiration and reactive oxygen species generation, whereas *HPR* is involved in heme metabolism and oxidative balance within vascular tissues [23, 24]. Thus, it is conceivable that physical activity modulates the downstream vascular effects of risk alleles at these loci by optimising mitochondrial redox balance and vascular oxidative stability.

Similar gene-lifestyle interactions have been described for other cardiometabolic traits, thus reinforcing the concept that genetic susceptibility is not deterministic but may be modified by environmental exposures [25–27]. However, most prior studies have focused on intermediate metabolic traits rather than clinically meaningful cardiovascular outcomes. By examining cardiovascular and cerebrovascular events in individuals with dyslipidemia, the present study extends research on gene-environment interactions to a more clinically relevant domain.

From a clinical perspective, these findings highlight the importance of lifestyle interventions even among individuals with elevated inherited lipid-related risks. Although exercise did not substantially alter genetically driven LDL-C levels, it may modify the association between specific genetic variants and IS risk. LDL-C represents only one dimension of cardiovascular risk, whereas exercise exerts pleiotropic effects on vascular function, insulin sensitivity, blood pressure regulation, and thrombosis [20, 28, 29]. Therefore, the present findings complement current guideline recommendations emphasising regular physical activity as a cornerstone of cardiovascular

prevention, irrespective of baseline lipid levels or genetic predisposition [30]. Furthermore, these results directly impact patient care by providing actionable evidence that lifestyle modifications can buffer genetic vulnerability. This scenario supports a shift towards precision prevention, in which clinicians can use genetic susceptibility data to motivate high-risk patients to adhere to tailored exercise regimens, thereby directly reducing the incidence of cerebrovascular events in real-world clinical settings.

Strengths and limitations

The primary strength of the current investigation involves the utilisation of a large, well-characterised population-based cohort, augmented by rigorous genetic replication across independent subcohorts. Furthermore, the analysis strategically focused on clinically meaningful cardiovascular outcomes rather than intermediate lipid traits alone. Residual confounding was meticulously minimised through propensity score matching, and robust subgroup and sensitivity analyses were conducted. Moreover, in the secondary analyses, fivefold stratified cross-validation was applied to the GRS-based models to reduce overfitting and ascertain model stability.

Several limitations should also be acknowledged. First, the statistical power to detect gene-exercise interaction effects was limited, particularly for low-frequency variants and relatively infrequent cardio-cerebrovascular events. Although the observed interactions did not persist after a stringent Bonferroni correction, they should be interpreted as exploratory, hypothesis-generating signals that require independent replication [31]. Second, physical activity was assessed by using a binary, self-reported measure, which is subject to recall and misclassification bias [32]. The absence of dose- and intensity-specific information precluded the evaluation of potential dose–response relationships. Third, the inherent constraints of a cross-sectional design prevent the establishment of definitive causal relationships. The possibility of reverse causation remains a significant concern, especially for individuals with a history of IS; such patients might have diminished their physical activity levels as a direct consequence of their clinical condition [33]. Finally, residual confounding due to lipid-lowering medication use cannot be ruled out. Because detailed information on statin type, dosage, and treatment duration was unavailable, stratified analyses by medication status were not feasible [34]. Future studies incorporating quantitative physical activity assessments, larger numbers of cardiovascular events, and prospective longitudinal designs will be essential to confirm these findings and to more definitively characterise gene-exercise interactions in dyslipidemic populations.

Conclusions

In conclusion, the genetic architecture of LDL-C remains stable across physical activity strata, whereas habitual exercise may mitigate the risk of ischemic stroke associated with specific LDL-related genetic loci. Established LDL-associated SNPs demonstrated consistent effects on LDL-C levels regardless of exercise status, thus reinforcing the robust genetic contribution of these SNPs to lipid metabolism. Although nominal evidence of gene-exercise interactions was observed for selected variants in relation to ischemic stroke, these findings should be interpreted as exploratory, given the loss of significance after stringent correction for multiple testing.

Collectively, these findings indicate that regular physical activity contributes to cerebrovascular protection even in the presence of unmodifiable genetic susceptibility. These results directly impact patient care by providing actionable evidence that the prescription of regular physical activity offers tangible cerebrovascular benefits and actively buffers the detrimental effects of inherited genetic burdens for patients with lipid abnormalities. Rather than viewing genetic susceptibility as an inevitable determinant of disease, these findings support an active clinical approach in which tailored behavioral interventions form the foundation of precision prevention strategies. By promoting active lifestyles among high-risk individuals, this approach fundamentally aligns with broader public health objectives to reduce premature mortality from non-communicable diseases and foster long-term health and well-being. With respect to future perspectives, efforts should prioritise the integration of polygenic risk profiling into routine clinical workflows, thus enabling clinicians to use genetic susceptibility data to motivate and individualise intensive lifestyle modifications, thereby optimising patient management and reducing the real-world incidence of cerebrovascular events.

Abbreviations

AUC	Area under the receiver operating characteristic curve
BMI	Body mass index
BP	Blood pressure
CAD	Coronary artery disease
CCB	Cardio-cerebrovascular disease
CVD	Cardiovascular disease
DL	Dyslipidemia
GRS	Genetic risk score
GWAS	Genome-wide association study
HDL-C	High-density lipoprotein cholesterol
HWE	Hardy–Weinberg equilibrium
IRB	Institutional Review Board
IS	Ischemic stroke
KoGES	Korea Genome and Epidemiology Study
LDL-C	Low-density lipoprotein cholesterol
OR	Odds ratio
PC1, PC2	Principal components 1 and 2 (population stratification)
PSM	Propensity score matching
ROC	Receiver operating characteristic
SNP	Single-nucleotide polymorphism
TCHL	Total cholesterol
TG	Triglycerides

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12944-026-02963-w.

Supplementary Material 1.

Acknowledgements
We thank the Korea Biobank Network and the KoGES project for providing genomic and epidemiologic data used in this research.

Authors' contributions
YCP and AHC contributed equally to study conception, data analysis, and manuscript drafting. YSK, YHK, JMP, KHH, BTK, KYK, JHS, JEC, and YJK contributed to data interpretation and manuscript revision. KWH supervised the study design, data quality control, and final approval. All of the authors have read and approved the final manuscript.

Funding
This research received no specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability
The datasets analyzed during the current study are available from the Korea Genome and Epidemiology Study (KoGES) repository (https://nih.go.kr/contents.es?mid=a50401010400) upon reasonable request and approval by the KoGES data access committee.

Declarations

Ethics approval and consent to participate
This study was approved by the Institutional Review Board of TheragenEtex (approval number: 700062–20190819-GP-006–02). All of the participants provided informed consent.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Family Medicine, Wonju Severance Christian Hospital, Yonsei University, Wonju, Republic of Korea
[2]Department of Family Medicine, CHA Bundang Medical Center, CHA University, Seongnam, Republic of Korea
[3]Department of Family Medicine, Uijeongbu Eulji Medical Center, Eulji University, Uijeongbu, Republic of Korea
[4]Department of Family Medicine, Seoul Medical Center, Seoul, Republic of Korea
[5]Department of Family Practice and Community Health, Ajou University Hospital, Suwon, Republic of Korea
[6]Department of Family Practice and Community Health, Korea University Ansan Hospital, Ansan, Republic of Korea
[7]Department of Family Medicine, Severance Hospital, Yonsei University, Seoul, Republic of Korea
[8]Theragen Health Co., Ltd., Seongnam, Republic of Korea

Received: 23 October 2025 / Accepted: 23 April 2026
Published online: 27 April 2026

References

1. World Health O. Global Health Estimates: leading causes of death 2019. Geneva: World Health Organization; 2024.
2. Ference BA, Ginsberg HN, Graham I. Low-density lipoproteins cause atherosclerotic cardiovascular disease: evidence from genetic, epidemiologic, and clinical studies. Eur Heart J. 2017;38:2459–72.
3. Libby P, Buring JE, Badimon L. Atherosclerosis. Nat Rev Dis Primers. 2019;5:56.
4. Noh H, Lee J, Kim S. Cause-of-death statistics in 2020 in the Republic of Korea. J Korean Med Assoc. 2023;66:97–107.
5. Thompson PD, Buchner D, Piña IL. Exercise and physical activity in the prevention and treatment of atherosclerotic cardiovascular disease. Circulation. 2003;107:3109–16.
6. Pedersen BK, Saltin B. Exercise as medicine - evidence for prescribing exercise as therapy in 26 different chronic diseases. Scand J Med Sci Sports. 2015;25:1–72.
7. Kraus WE, Houmard JA, Duscha BD. Effects of the amount and intensity of exercise on plasma lipoproteins. N Engl J Med. 2002;347:1483–92.
8. Mann S, Beedie C, Jimenez A. Differential effects of aerobic and resistance training on lipid profile: a systematic review and meta-analysis. Sports Med. 2014;44:211–21.
9. Leon AS, Sanchez OA. Response of blood lipids to exercise training alone or combined with dietary intervention. Med Sci Sports Exerc. 2001;33:S502–15.
10. Teslovich TM, Musunuru K, Smith AV. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010;466:707–13.
11. Willer CJ, Schmidt EM, Sengupta S. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013;45:1274–83.
12. Ference BA, Yoo W, Alesh I. Effect of long-term exposure to lower LDL cholesterol beginning early in life on the risk of coronary heart disease. J Am Coll Cardiol. 2012;60:2631–9.
13. Qi Q, Cho YA. Gene-environment interaction and obesity. Nutr Rev. 2008;66:684–94.
14. Hall KT, Reilly MP, Rimm EB. Genetic variation at the GUCY1A3 locus modifies cardiovascular prevention effects of aspirin. Eur Heart J. 2019;40:3385–92.
15. Boekholdt SM, Hovingh GK, Mora S. Very low levels of atherogenic lipoproteins and the risk for cardiovascular events. Lancet. 2012;380:581–90.
16. Kathiresan S, Melander O, Guiducci C. Polymorphisms associated with cholesterol and risk of cardiovascular events. N Engl J Med. 2008;358:1240–9.
17. Kodama S, Tanaka S, Saito K. Effect of aerobic exercise training on serum levels of high-density lipoprotein cholesterol: a meta-analysis. Arch Intern Med. 2007;167:999–1008.
18. Sarzynski MA, Rankinen T, Bouchard C. Genetics of exercise-induced changes in blood lipids. Curr Opin Lipidol. 2015;26:123–30.
19. Leon AS, Rice T, Mandel S. Blood lipid response to 20 weeks of supervised exercise in a large biracial population: the HERITAGE Family Study. Circulation. 2000;102:3137–44.
20. Green DJ, Hopman MT, Padilla J. Vascular adaptation to exercise in humans: role of hemodynamic stimuli. J Physiol. 2017;595:5005–17.
21. Hambrecht R, Wolf A, Gielen S. Effect of exercise on coronary endothelial function in patients with coronary artery disease. Circulation. 2000;102:454–60.
22. Powers SK, Radak Z, Ji LL. Exercise-induced oxidative stress: past, present and future. Physiol Rev. 2020;100:781–814.
23. Löffler M, Fairbanks LD, Zameitat E. The role of DHODH in mitochondrial metabolism. Biochim Biophys Acta. 2015;1847:646–53.
24. Schaer DJ, Buehler PW, Alayash AI. Heme metabolism and vascular biology. Nat Rev Cardiol. 2013;10:338–52.
25. Ordovas JM. Gene-environment interactions: implications for the metabolic syndrome. Am J Clin Nutr. 2007;85:64–8.
26. Franks PW, Pearson ER. Lifestyle modification and genetic risk. BMJ. 2016;354:i3831.
27. Kilpeläinen TO, Qi L, Brage S. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. PLoS Med. 2011;8:e1001116.
28. Lee IM, Shiroma EJ, Lobelo F. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet. 2012;380:219–29.
29. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ. 2006;174:801–9.
30. Arnett DK, Blumenthal RS, Albert MA. 2019 ACC/AHA guideline on primary prevention of cardiovascular disease. Circulation. 2019;140:e596–646.
31. Aschard H. A perspective on interaction effects in genetic association studies. Am J Epidemiol. 2014;180:203–10.
32. Prince SA, Adamo KB, Hamel ME. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. Int J Behav Nutr Phys Act. 2008;5:56.
33. Lee CD, Folsom AR, Blair SN. Physical activity and stroke risk: a meta-analysis. Stroke. 2003;34:1710–6.
34. Baigent C, Blackwell L, Emberson J. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet. 2010;376:1670–81.

Publisher's Note