Blood lipids and risk of acute myocardial infarction using Mendelian randomization

Guen hui Kim

Graduate School of Public Health
Yonsei University
Department of Epidemiology and Health Promotion
Division of Epidemiology

Blood lipids and risk of acute myocardial infarction using Mendelian randomization

A Masters Thesis

Submitted to the Department of Epidemiology and Health Promotion,

Division of Epidemiology

and the Graduate School of Yonsei University

in partial fulfillment of the

requirements for the degree of

Master of Public Health

Guen hui Kim

June 2020

This certifies that the Masters thesis of

Guen hui Kim is approved.

Sun Ha Jee

Thesis Supervisor: Sun Ha Jee

Ji-Young Lee

Thesis Committee Member: Ji-Young Lee

Keum Ji Jung

Thesis Committee Member: Keum Ji Jung

Graduate School of Public Health

Yonsei University

June 2020

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Blood lipids and risk of acute myocardial infarction using Mendelian randomization

Guen hui Kim

Graduate School of Public Health

Yonsei University

Directed by Professor Sun Ha Jee

Background: It is widely known that blood lipids have negative and positive associations with ischemic heart disease. However, there was little evidence for an association between acute myocardial infarction (AMI) and blood lipids. Also, it has been controversial whether triglyceride is causal with the risk of AMI. Therefore, in the present study, we conducted a Mendelian randomization analysis using the Korean Cancer Prevention Study-II (KCPS-II) biobank cohort to examine whether blood lipids have a causal effect on risk of AMI.

Materials and Methods: According to the data of KCPS-II biobank, this study included total 16,553 participants who had provided the information of genetic variants with informed consent. Observational multivariable Cox regression of incident cases and one-sample MR Cox regression model approaches were carried out in this paper to clarify the association between blood lipids and AMI. Genome-wide association study (GWAS) was conducted to select highly associated single nucleotide polymorphisms (SNPs) with lipid traits. We calculated weighted genetic risk score (WGRS), including all mutually independent SNPs after excluding all other interrelated SNPs with linkage disequilibrium (LD). Mendelian Randomization analysis (MR), which can attenuate the confounding effect or inverse causality, was used to verify the causality between AMI and lipids using WGRS.

Results: The WGRS of each LDL-C, HDL-C, and triglycerides were very strongly associated with each lipid fractions. Among them, 1mg/dl genetically elevated LDL cholesterol was associated with an increased risk of AMI (HR:1.041; 95% CI:1.015-1.068). In contrast, the result demonstrated no causal associations with HDL-C, TG, and the risk of AMI (HR:0.990; 95% CI:0.905-1.084; HR:1.602; 95% CI:0.321-7.988, respectively).

Conclusion: Using one-sample Mendelian randomization, findings showed consistent evidence for an adverse effect of increased LDL-C on AMI risk. On the other hand, no significant association between HDL-C, TG and risk of AMI were

observed in the Korean population.

Keywords: Acute myocardial infarction, Mendelian randomization, GWAS, genetic risk score, blood lipids

I. INTRODUCTION

1. Study background

Acute myocardial infarction (AMI) is one of the leading causes of heart disease deaths, and its fatality rate and incidence rate have been on the rise every year in Korea (HIRA, 2017; Y. Kim et al., 2019). According to the National Health Insurance statistical yearbook, the incidence of acute myocardial infarction per 1,000 people has been on the rise every year from 1.36 in 2011 to 1.73 in 2015. Chronic diseases such as hypertension, type 2 diabetes, and dyslipidemia have continuously associated with risk factors for myocardial infarction in Korea (Khera & Kathiresan, 2017; Y. Kim et al., 2019).

The characteristics of Korean AMI patients are considerably dissimilar with the Western population by etiologies and risk factors (Sim & Jeong, 2017). Since Korea has been westernized recently, the risk factors of AMI in Korea also has been changed (Hong, Kang, Lee, & Kim, 2009; Ueshima et al., 2008). Therefore, the proportion of dyslipidemia in patients with AMI has been sharply increased from 8.0% in 2006 to 23.0% in 2018 in Korea (Sim & Jeong, 2017). It has raised about 3-fold during the past decades. Low level of high-density lipoprotein cholesterol (HDL-C) and high triglyceride levels are the patterns of dyslipidemia in the Korean population, which is different from those of dyslipidemia in Western population (Sim & Jeong, 2017; Ueshima et al., 2008).

Through several observational studies, low-density lipoprotein cholesterol (LDL-C) and triglycerides (TG) have been discovered for risk factors of AMI, while HDL-C is associated with reduced risk of AMI (Tada, Nohara, & Kawashiri, 2018). Also, several Mendelian randomization studies conducted the causal effect of LDL-C, HDL-C, and triglycerides on ischemic heart disease or cardiovascular disease. However, most of the reviews are mainly conducted in the range of overall heart disease, not the specific subtype of heart disease. Therefore, currently, there is little evidence for an association between acute myocardial infarction (AMI) and blood lipids.

In the preceding MR analyses, indisputable evidence of causal relationship with LDL-C and AMI is existed (Voight et al., 2012). However, the previous MR study revealed no causal relationship between HDL-C and the risk of AMI (Voight et al., 2012). In addition, there has been uncertainty whether triglyceride is causal with acute myocardial infarction (Allara et al., 2019). In Western populations, some studies showed that serum TG level is associated with a higher risk of AMI and other heart diseases (Allara et al., 2019; Holmes et al., 2015; Nordestgaard & Tybjærg-Hansen, 2011). Since triglycerides transported by VLDL, chylomicron, or remnant of metabolism, it has been controversial whether TG is causally associated with AMI risk (Holmes et al., 2015; Jørgensen et al., 2013; Kawashiri, Tada, Nomura, & Yamagishi, 2018). Moreover, most of these studies using Mendelian randomization analysis have focused on Western population. Thus, we conducted

genetic analysis approach among Korean populations to clarify the effect of lipids on AMI in Korean population.

Mendelian randomization (MR) analysis applies the instrumental variable (IV) method for regression analysis to examine the estimates of the causal effect. Genetic variants which are robustly associated with exposures are used as an instrument variant. Therefore, MR analysis is less susceptible to measurement error, reverse causation, and confounding than observational multivariable regression approaches.

We used genetic variants robustly associated with LDL-C, HDL-C, and TG identified in Korean Cancer Prevention Study-II (KCPS-II) biobank cohort to explore whether these three lipid traits have a causal effect on AMI risk. To do this, we conducted a one-sample MR analysis using summed values for multiple independent SNPs as instrumental variables, from which estimates were compared with conventional observational multivariable regression results in the same study.

2. Objectives

In this study, we performed a Mendelian randomization approach to estimate the causal effect of lipids and risk of acute myocardial infarction by using the population-based Korean Cancer Prevention Study-II (KCPS-II biobank) biobank cohort data.

The objectives of this study are as follows:

1) Evaluate the epidemiological association between blood lipids and acute myocardial infarction using Korean Cancer Prevention Study-II (KCPS-II biobank).
2) Examine the genetic variants of low-density lipoprotein, high-density lipoprotein, and Triglyceride levels using Genome-wide association studies (GWAS).
3) Investigate the causality of blood lipids on acute myocardial infarction performing a Mendelian randomization analysis.

II. MATERIALS AND METHODS

1. Study design

The study design of this paper is as below (Figure 1).

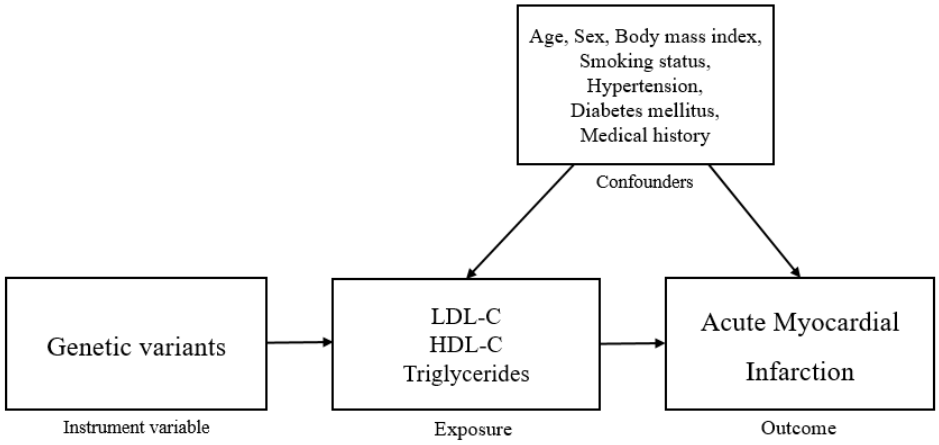


Figure 1. Study design

2. Study population and data sources

A large cohort of the population-based Korean Cancer Prevention Study-II (KCPS-II biobank), which was started in April 2004, is a data source of this study. The blood samples collected from the KCPS-II biobank were at first taken in 2004, from two hospitals and more extensive to eleven hospitals since April 2006 (Jee, Lee, Jung, & Jee, 2016). These eleven health promotion centers located in Seoul and Gyeonggi Province comprise a population of 18,879,351 (Jee et al., 2018). Among the health examination participants from 2004 to 2013, the number of individuals who had contributed informed consent for the study was 159,844 (Jee et al., 2018; Lee et al., 2017). The Severance Medical Ethics Committee of the Korean approved this study (no. 4-2011-0277).

Within this population, 16,553 individuals were available for the study who had genetic information. Two kinds of study population were established in this paper. In the first type of study population, after excepting who had missing or extremely abnormal values for the essential variables including LDL-C, HDL-C, triglycerides, and excluding who had a history of acute myocardial infarction, final 16,257 participants were satisfied (Figure 2). Also, within the 159,844 participants who had given written informed consent, we randomly selected 5,000 subjects for the sub-cohort group (Figure 2). Within the sub-cohort, four thousand eight hundred seventy-nine participants had no missing on major variables (Figure 2). Finally, data for up to 328 cases of acute myocardial infarction and 4,869 AMI-free controls were

available from KCPS-II biobank.

In the second type of study population, the participants using lipid-lowering drugs were excluded, and it is well presented in figure 3. Fifty-two participants were using lipid-lowering drugs in the sub-cohort, while three hundred and sixty-five participants were using lipid-lowering drugs in the main cohort. After excluding participants using lipid-lowering medicine, the final 5,131 participants for the study population were available (Figure 3).

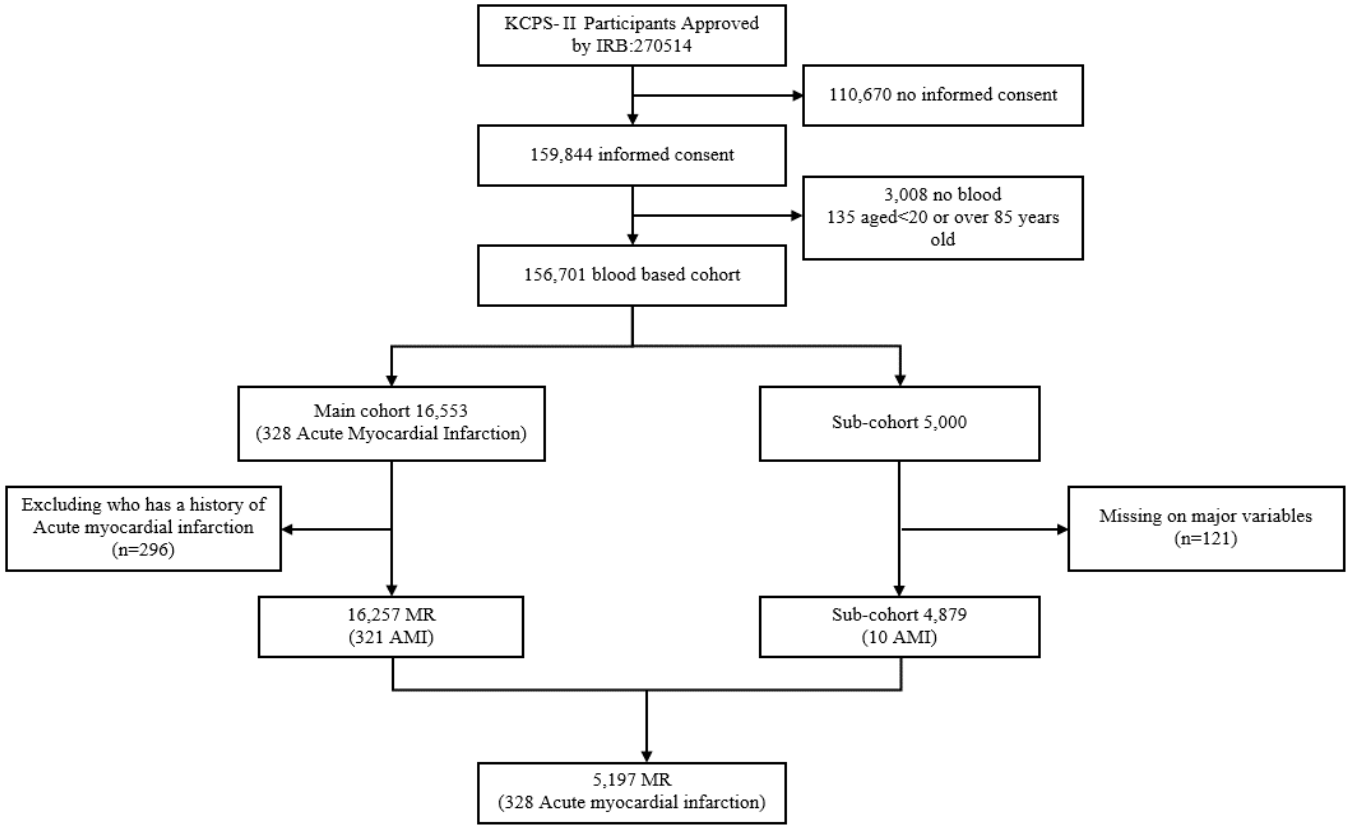


Figure 2. Flow chart of study population (n=5,197)

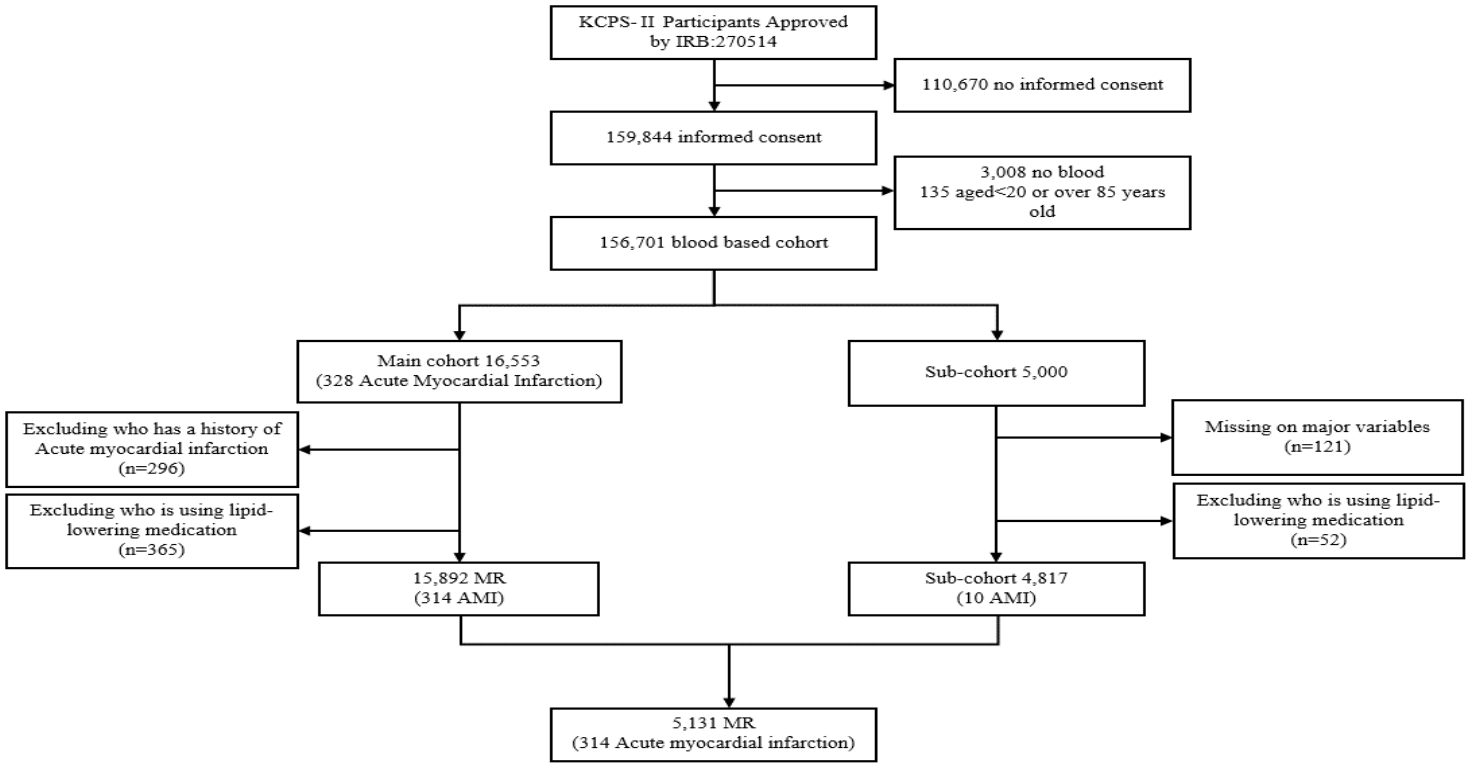


Figure 3. Flow chart of study population excluding lipid-lowering drugs users (n=5,131)

3. Data collection

From 2004 to 2013, the data was collected from health assessment. Each individual was answered the self-completion of a questionnaire to gather information about medical history, social status, drugs use, and behavioral risk factors such as alcohol, smoking, and physical activity (Jee et al., 2018). A medical examination was also undertaken by all participants to collect anthropometric data and blood pressure data, which were measured in a standardized manner. Body mass index (BMI) calculation was measured by dividing weight (kg) by height squared (m^2). Blood pressure was measured using a sphygmomanometer in a seated position.

Clinical chemistry assay serum samples were collected while participants were fasted overnight and were kept at -70℃ (Jee et al., 2018). Glucose, total cholesterol, triglyceride, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and other biomarkers were measured in the hospital laboratory by a COBAS INTEGRA 800 and a 7600 Analyzer (Hitachi, Tokyo, Japan) (Jee et al., 2018). LDL-C was calculated directly based upon total, HDL-C, and triglyceride levels (Jee et al., 2018).

4. Measurement of Outcome

The outcome variable of this study is the incidence of acute myocardial infarction, which was recorded in hospital admission discharge from 2005 to 2014 (median follow-up duration of 7 years) (Lee et al., 2017). These outcome variable data were obtained from health insurance claim data from the National Health Insurance Service. As stated by the international classification of disease - Tenth revision codes, acute myocardial infarction was defined as I21.0—I21.9.

Acute myocardial infarction (I21.0-I21.9) specified as acute or with a stated duration of 4 weeks (28 days) or less from the onset. By using the tenth revision of the international classification of diseases codes (ICD-10), the accuracy for diagnosing AMI in medical insurance claims data was >70%, and reliability was fair to good in 2012 (Kimm, Yun, Lee, Jang, & Jee, 2012).

5. Genotyping and quality control

The genotype data were produced using the Korean chip (K-chip) obtained from the K-chip Consortium. K-chip was designed by the center for genome science, Korea national institute of health, Korea (4845-301, 3000-3031). The Korean chip project is a genome-based research project that identifies the genetic causes of common chronic diseases common to Koreans. It has developed a Korean chip for

gene analysis and created and analyzed genetic information for about 180,000 Koreans based on a large cohort (Moon et al., 2019). The Korean chip contains 830,000 SNPs with no imputation of SNPs. Internal quality control (QC) of genetic data has been carried out at minor allele frequencies (MAF) < 0.05, and SNPs were showing deviation from Hardy-Weinberg equilibrium (HWE) at P < 0.0001 (Jee et al., 2016; Lee et al., 2017).

We selected genetic instruments of each exposure variable, LDL-C, HDL-C, and TG, with genome-wide significance ($P<5*10^{-8}$). GWAS analysis was performed using a linear regression under an additive genetic model adjusted for age and sex. To determine genetic associations with Lipids over all SNPs, PLINK version 1.07 was used. Also, Haploview version 4.1 was used to generate Manhattan plots and linkage disequilibrium (LD) structures. We also conducted imputation analysis after genotyping.

6. Imputation methods for predicting lipids genotyping based on SNPs

In this study, as a reference panel for estimating missing genotypes, imputation analysis was performed using the 1000G Phase 3 dataset for the East Asian population (Japanese in Tokyo JPT, Chinese in Beijing CHB).

Genotype imputation has become a standard tool for genome-wide related research because researchers can economically approximate all genome sequence data from single nucleotide polymorphism sequence data (Marchini & Howie, 2010). Tagging SNP represents the surrounding genetic regions. Thus imputation techniques can predict thousands of SNPs in the genetic areas to expand information and use them for research. This method enhances statistical power, provides a fine mapping of causal variants, and plays an essential role in a meta-analysis of genome-wide association studies (Das, Abecasis, & Browning, 2018).

For years, genotype imputation has significantly benefited from improved genotype technology and increased genotype information in publicly available datasets. Examples of such data sets include the international Hapmap project, the 1000 Genome Project (1000G), the UK10K Project, the Haplotype preferences Consortium (HRC), and the Transomics for Pressure Medicine (TOPMED) program (Das et al., 2018). The larger panels, including a catalog of more detailed genetic variants, increase the probability of causal variant imputation and increase the accuracy of downstream binding analysis power.

7. Mendelian randomization analysis

Mendelian Randomization (MR), which investigates causality using a genetic variant as an instrument variable (IV), was conducted to examine whether there is

a causal relationship between lipid traits and AMI (Thomas & Conti, 2004). In medical statistic studies or epidemiological studies, the theory of IVs, primarily derived from the field of econometrics, has been commonly used to create counterfactual causal inferences (Lawlor, 2016; Thomas & Conti, 2004). Based on this concept, the idea of Mendelian randomization was introduced as a method using genetic variants as the IV. Fundamental assumptions for a genetic variant to satisfy to be an IV are as follow (Stephen Burgess, 2015):

i. The variant is robustly associated with the exposure.

ii. The variant is not associated with any confounding factors that bias the association of the exposure-outcome association.

iii. The variant affects the outcome only via its association with the modifiable exposure (the assumption also known as ‘exclusion restriction’).

Mendelian randomization is similar to a randomized controlled trial (RCT). As shown in figure 2, in an RCT, individuals are divided into two or more subgroups in a random method (Nitsch et al., 2006). In a similar way, in Mendelian randomization, genetic variant was used to form subgroups (Figure 4). Different treatments are given to each subgroup in randomized trial. In the same way, Mendelian randomization uses a genetic variant to divide individuals into subgroups (Figure 4). These subgroups differ systematically in the exposure, but

not in any other factor except for those causally 'downstream' of the exposure. A difference in outcomes between these subgroups would therefore indicate a causal effect of the exposure on the outcome (Burgess et al., 2019). The way of inferring a causal effect of the exposure on the outcome is similar to inferring an intention-to-treat effect in an RCT (Teumer, 2018).

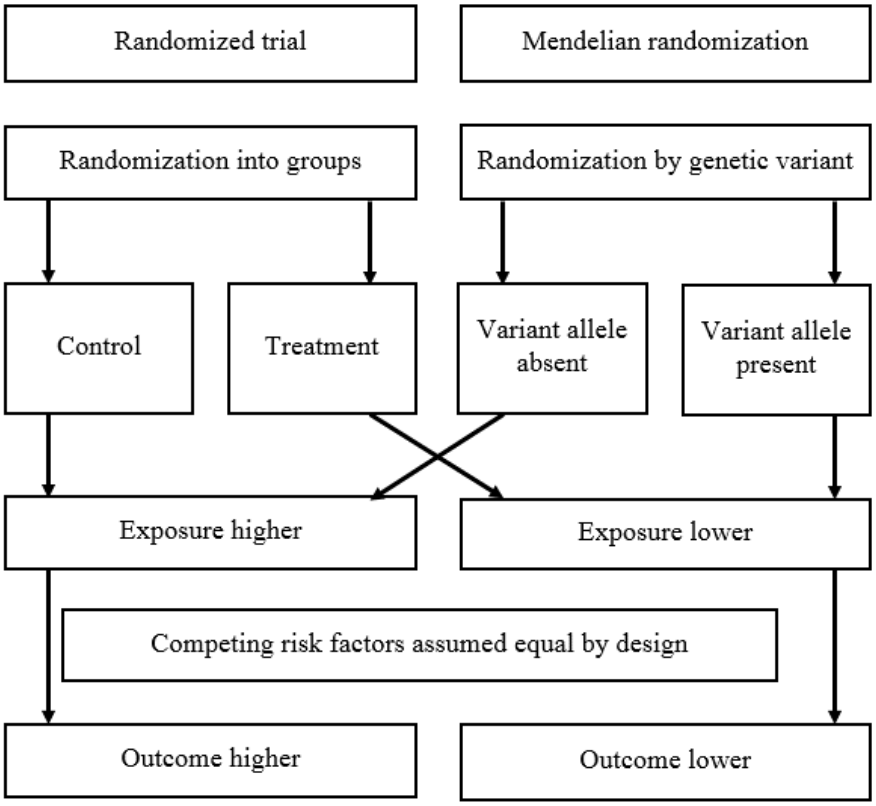


Figure 4. Comparison of a randomized controlled trial and Mendelian randomization

8. Selection of IV and calculating genetic risk scores

After excluding all other interrelated SNPs with linkage disequilibrium (LD), the weighted genetic risk score (WGRS) of each lipid was calculated by using mutually independent SNPs.

We tried two approaches to constructing the genetic risk scores. In the first way of calculating WGRS, increasing the number of risk alleles at each SNP (score of 0, 1, or 2) was weighted by the per-allele regression coefficient on the relevant phenotype and then summed across SNPs (Lee et al., 2017). The second approach is the same as the first method of using the estimate, which analyzed by the regression between the association of the risk allele and lipids. However, under the assumption that the regression is non-linear, made dummy variables of risk alleles at each SNP (0, 1, or 2). After that, the WGRS was calculated by multiplying each estimated coefficient of dummy variables by the number of corresponding risk alleles (0, 1, or 2).

9. Sensitivity analysis

To examine the robustness of the results, we conducted several sensitivity analyses. Since it is a case-cohort design, participants in a sub-cohort randomly selected from the whole original cohort, and the sub-cohort sampling fraction is the

proportion of individuals in the initial cohort. In the case-cohort study design, all incident cases of the disease of interest in the original cohort are to be the case group, and the sub-cohort, randomly selected, is to be the control group. When it comes to performing case-cohort study design, over-representation of cases in the sample should be described. Therefore, we created four types of control groups to perform sensitivity analyses for MR.

First of all, to deal with over-representation issues of sub-cohort, we conducted a sensitivity analysis by making the control group with all participants without AMI. Second, in terms of the proportion of diseases in KCPS-II, the most representative diseases were cancer and ASCVD. Therefore, sensitivity analyses were performed as control groups with participants who were with cancer and those with ASCVD. Sensitivity analyses for the four control groups are likely to attenuate a bias in one control group.

On account of the exogenous effect, participants who are taking lipid-lowering drugs could have controlled blood lipid levels. Therefore, we restricted the analyses to participants, removing participants who were currently using lipid-lowering drugs, and we also compared the results with the original study. Furthermore, we conducted Mendelian randomization analysis, which conducted by using lipid-lowering drugs use as a covariate for adjustment.

10. Statistical methods

We carried out a conventional multivariable cox proportional hazard model with the outcome variable of acute myocardial infarction status while adjusting for covariates of age, sex, and other confounding variables. According to the lipid fractions, we used log-transformed triglycerides level for the statistical tests and regression analyses since it showed skewed distribution. All statistical analysis was implemented in SAS 9.2 (SAS Institute Inc., Cary, NC). Also, the statistical tests were two-sided, and $P<0.05$ was considered statistically significant.

In MR analysis, we performed an instrumental variable analysis of blood lipids and the risk of acute myocardial infarction. Instrumental variable (IV) approach, chance to control confounding bias, is usually performed with the regression model. However, in this paper, under an additive hazard model, we applied the instrumental variable (IV) method for regression analysis in a survival context to examine the estimates of the causal effect (Tchetgen, Walter, Vansteelandt, Martinussen, & Glymour, 2015).

To confirm that the instrumental variable is strongly associated with exposure, we calculated statistics of F-statistic to assess the strength of the association WGRS with each lipid level. By using F-statistics, the significance of genetic variables was assessed. And then, Mendelian randomization analysis was carried out with two different methods: Wald estimator and two-stage least squares method (2SLS). We

used these two methods to assess the estimates of the causal effect of exposure X (blood lipids) on binary outcome Y (risk of acute myocardial infarction). We performed Mendelian randomization analyses with STATA software version 13.1 (Stata Corp. LP, College Station, TX, USA).

III. RESULTS

PART I. Descriptive analysis of the Korean Cancer Prevention Study-II (KCPS-II) biobank

1. Proportion of diseases in KCPS-II biobank

Figure 5 illustrates the proportion of diseases in KCPS-II. As shown in the pie chart, the portion of all cancer makes up the most considerable amount in KCPS-II. Followed by atherosclerotic cardiovascular disease (ASCVD), including stroke and heart disease, accounts for 23%. Moreover, in the case of the percentage of ASCVD, the portion of ischemic heart disease (IHD) and stroke accounts for 43% and 31%, respectively.

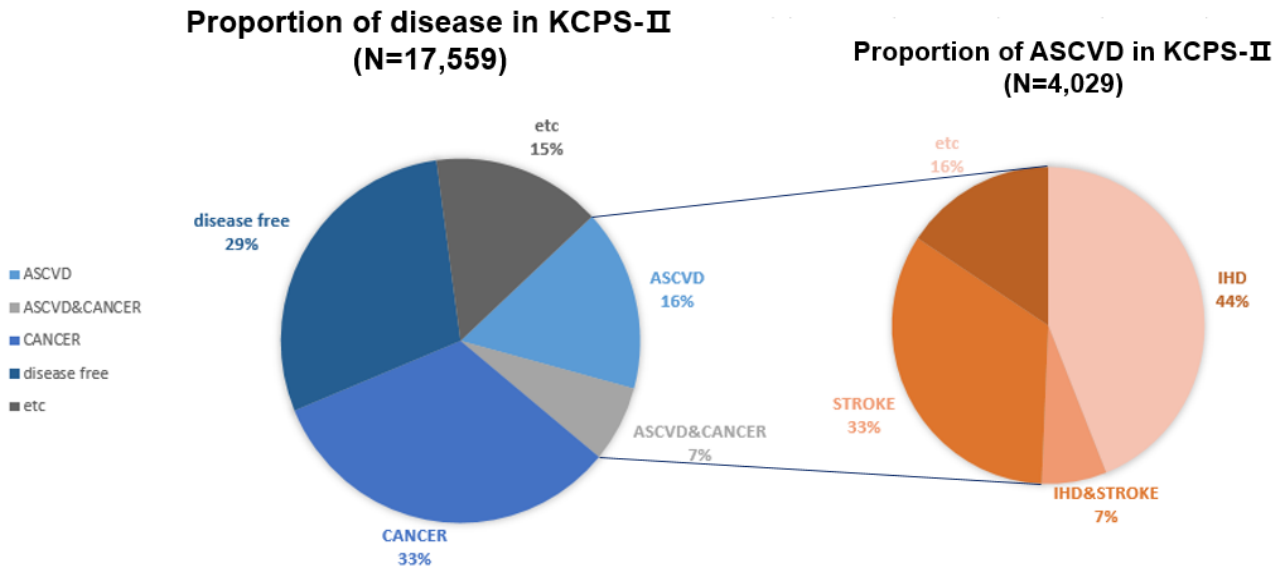


Figure 5. Proportion of diseases in KCPS-II biobank (n=17,559)

Note: ASCVD, atherosclerotic cardiovascular disease; IHD, ischemic heart disease

The number of disease participants in KCPS-II is well described in figure 6. All the number of diseases contain the number of overlapped disease. Without chronic diseases, such as hypertension, diabetes mellitus, and metabolic syndrome, ASCVD showed the top number of incidence, and the incidence of thyroid cancer was higher than those of any other type of cancer.

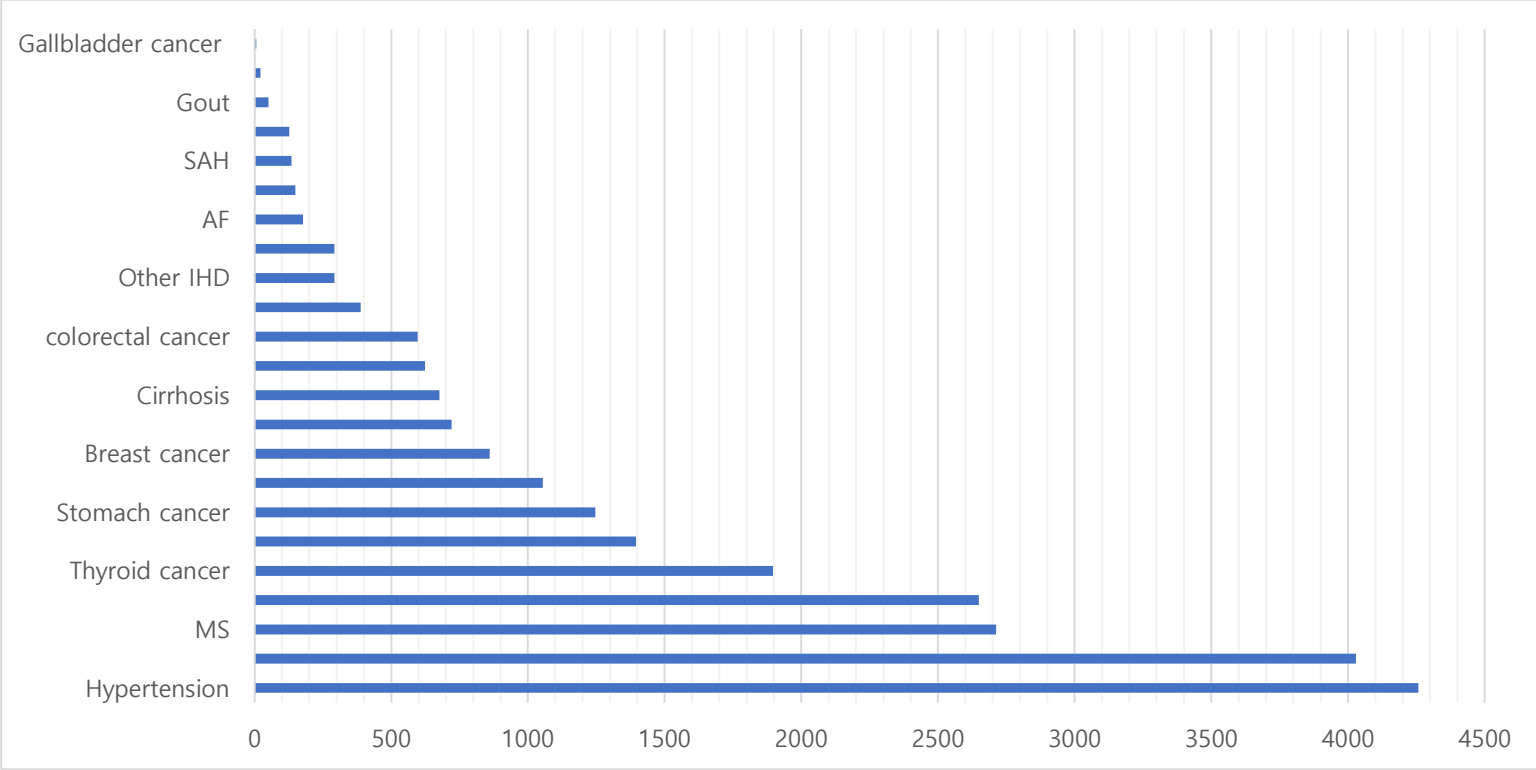


Figure 6. Number of participants with disease in KCPS-II (n=17,559)

Note: SAH, subarachnoid hemorrhage; ICH, intracerebral hemorrhage; AF, atrial fibrillation; AMI, acute myocardial infarction; MS, metabolic syndrome; ASCVD, atherosclerotic cardiovascular disease

2. Baseline characteristics of KCPS-II biobank participants according to different study population

A total of 159,844 participants were included in the KCPS-II biobank. We divided KCPS-II data into several subgroups. The general characteristics of KCPS-II by subgroups are well summarized in table 1. The first group includes all participants of KCPS-II. The second group consists of 16,553 participants who have genetic information. The third group is a sub-cohort that comprises randomly selected 5,000 subjects. The rest of the groups were grouped by the most common type of disease in the biobank. The mean age of 16,553 participants in KCPS-II biobank was 46.1, which is much higher than that of 159,844 participants in KCPS-II biobank (41.5) and sub-cohort (42.5). The triglyceride (TG) level of ASCVD group shows the highest among the groups (159.6), while the HDL-C level was the lowest. Also, ASCVD group especially took a large proportion of lipid-lowering drugs users among groups (9.4%, 5.3%, respectively), while a total of 159,844 participants group only had 1.8% of current users of lipid-lowering drugs.

Table 1. Baseline characteristics of KCPS-II biobank according to different study population

	159,844	16,553	Sub-cohort 5,000	ASCVD		Cancer					
				IHD	STROKE	Thyroid	Stomach	Lung	Breast	colon	liver
Number of participants	159,844	16,553	5,000	2,046	1,564	1,896	1,247	1,055	860	597	292
Female, N (%)	62,944(39.6)	5,782(35.0)	1,659(33.4)	467(22.8)	507(32.4)	1,108(58.4)	299(24.0)	130(26.5)	853(99.2)	116(27.8)	32(11.0)
Age (yrs)	41.5±10.5	46.1±11.0	42.5±9.0	54.1±10.9	53.9±11.7	41.5±9.2	51.7±11.2	56.0±10.8	45.7±9.6	50.7±11.2	52.2±10.2
LDL-C (mg/dl)	112.3±31.5	114.4±32.2	114.0±31.5	114.3±39.8	114.0±36.9	111.7±32.2	111.7±35.7	115.5±37.6	107.0±36.6	113.3±38.7	104.1±35.2
HDL-C (mg/dl)	52.2±10.9	51.0±10.7	50.9±10.0	45.8±14.8	48.1±15.1	51.8±12.3	48.4±15.3	49.3±15.3	54.1±17.1	49.2±15.2	48.0±15.2
TG (mg/dl)	134.8±89.3	141.9±94.7	140.8±91.1	159.6±104.9	151.7±97.5	126.0±82.3	146.6±99.4	143.6±97.7	106.1±65.6	149.2±97.2	122.3±80.0
FBS (mg/dl)	90.7±15.0	94.3±21.7	90.1±15.1	99.3±22.4	97.2±22.0	89.4±14.4	95.8±19.6	95.0±19.8	89.4±17.9	96.1±21.8	96.3±21.1
MLPD, N (%)	2,797(1.8)	430(2.6)	52(1.0)	193(9.4)	83(5.3)	19(1.0)	36(2.9)	25(5.1)	29(3.4)	13(2.2)	5(1.7)
AMI, N (%)	1,002(0.6)	624(3.8)	10(0.2)	624(30.5)	77(4.9)	10(0.5)	12(1.0)	13(2.7)	3(0.4)	10(1.7)	5(1.7)

Note: LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides; FBS, fasting blood sugar; AMI, acute myocardial infarction; MLPD, lipid-lowering drugs use status; ASCVD, Atherosclerotic cardiovascular disease
Data are expressed as mean (SD) unless otherwise indicated

3. Results of multivariable Cox regression analysis

During a followed-up period, 159,844 participants in KCPS-II biobank, a total of 449 men and women were hospitalized on account of acute myocardial infarction (AMI). Also, 323 incidents of acute myocardial infarction events were recorded in 16,553 participants in KCPS-II biobank. The associations between blood lipids level and the risk of acute myocardial infarction were explored using Cox proportional hazard models. In the analyses, adjusted for age, sex, other lipid profiles, smoking status, diabetes mellitus, hypertension, and body mass index (BMI), we found significant positive associations between LDL-C and risk of acute myocardial infarction, which remained similar in the adjusted model (Model 5) both 159,844 participants and 16,553 participants (hazard ratio 1.008; 95% confidence interval 1.005 to 1.010, hazard ratio 1.010; 95% confidence interval 1.007 to 1.013, respectively).

Moreover, a positive association between log-transformed TG and risk of AMI were observed in relation to both groups (Model 5) (hazard ratio 1.369; 95% confidence interval 1.128 to 1.662, hazard ratio 1.463; 95% confidence interval 1.161 to 1.843, respectively). Cox regression estimates were similar after excluding participants who reported using treatment of hyperlipidemia (Model 6).

On the other hand, a significant inverse association was observed between HDL-C and the risk of acute myocardial infarction in the multivariable adjusted model

(Model 5) both 159,844 participants and 16,553 participants (hazard ratio 0.971; 95% confidence interval 0.960 to 0.982, hazard ratio: 0.971; 95% confidence interval 0.958 to 0.985, respectively). Both groups showed the similar results of relationships between blood lipids and the risk of AMI (Table 2-4).

Table 2. Association of acute myocardial infarction with LDL-C using Multivariable Cox regression in KCPS-II biobank

Exposure variable	Model	159,844 participant in KCPS-II biobank (AMI cases, n=449)	16,553 participant in KCPS-II biobank (AMI cases, n=323)
		HR (95% CI)	HR (95% CI)
LDL-C (mg/dl)	Model 1	1.012(1.010-1.015)	1.014(1.011-1.017)
	Model 2	1.009(1.006-1.012)	1.012(1.009-1.016)
	Model 3	1.007(1.004-1.009)	1.009(1.006-1.012)
	Model 4	1.008(1.005-1.010)	1.010(1.007-1.013)
	Model 5	1.008(1.005-1.010)	1.010(1.007-1.013)
	Model 6	1.008(1.005-1.011)	1.010(1.007-1.013)

Note: HR, Hazard ratio; CI, Confidence interval; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides
Model 1: Crude model
Model 2: Adjusted for age and sex
Model 3: Model 2 plus additional adjustments for other lipid profiles
Model 4: Model 3 plus additional adjustments for smoking status, diabetes, hypertension, and body mass index
Model 5: Model 4 plus additional adjustments for treatment of hyperlipidemia
Model 6: Eliminated lipid lowering drugs users and adjusted for age, sex, other lipid profiles, smoking status, diabetes, hypertension, alcohol consumption and body mass index

Table 3. Association of acute myocardial infarction with HDL-C using Multivariable Cox regression in KCPS-II biobank

Exposure variable	Model	159,844 participant in KCPS-II biobank (AMI cases, n=449)	16,553 participant in KCPS-II biobank (AMI cases, n=323)
		HR (95% CI)	HR (95% CI)
HDL-C (mg/dl)	Model 1	0.933(0.923-0.943)	0.935(0.924-0.947)
	Model 2	0.955(0.945-0.965)	0.951(0.939-0.963)
	Model 3	0.969(0.958-0.980)	0.969(0.955-0.982)
	Model 4	0.971(0.960-0.982)	0.971(0.958-0.985)
	Model 5	0.971(0.960-0.982)	0.971(0.958-0.985)
	Model 6	0.970(0.958-0.982)	0.972(0.958-0.986)

Note: HR, Hazard ratio; CI, Confidence interval; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides
Model 1: Crude model
Model 2: Adjusted for age and sex
Model 3: Model 2 plus additional adjustments for other lipid profiles
Model 4: Model 3 plus additional adjustments for smoking status, diabetes, hypertension, and body mass index
Model 5: Model 4 plus additional adjustments for treatment of hyperlipidemia
Model 6: Eliminated lipid lowering drugs users and adjusted for age, sex, other lipid profiles, smoking status, diabetes, hypertension, alcohol consumption and body mass index

Table 4. Association of acute myocardial infarction with TG using Multivariable Cox regression in KCPS-II biobank

Exposure variable	Model	159,844 participant in KCPS-II biobank (AMI cases, n=449)	16,553 participant in KCPS-II biobank (AMI cases, n=323)
		HR (95% CI)	HR (95% CI)
Log-transformed TG (mg/dl)	Model 1	2.492(2.144-2.897)	2.325(1.938-2.789)
	Model 2	1.887(1.598-2.227)	1.900(1.563-2.309)
	Model 3	1.507(1.250-1.816)	1.512(1.214-1.884)
	Model 4	1.380(1.137-1.675)	1.462(1.161-1.841)
	Model 5	1.369(1.128-1.662)	1.463(1.161-1.843)
	Model 6	1.460(1.193-1.788)	1.539(1.214-1.951)

Note: HR, Hazard ratio; CI, Confidence interval; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides
Model 1: Crude model
Model 2: Adjusted for age and sex
Model 3: Model 2 plus additional adjustments for other lipid profiles
Model 4: Model 3 plus additional adjustments for smoking status, diabetes, hypertension, and body mass index
Model 5: Model 4 plus additional adjustments for treatment of hyperlipidemia
Model 6: Eliminated lipid lowering drugs users and adjusted for age, sex, other lipid profiles, smoking status, diabetes, hypertension, alcohol consumption and body mass index

4. Comparisons of current lipid-lowering drugs users and non-users

At the study baseline, there were 2,797 of 159,844 participants (1.7%), and 434 of 16,955 participants (2.5%) were using lipid-lowering drugs. Table 5 shows the baseline characteristics of the study participants stratified by lipid-lowering treatment status (users versus non-users).

Mean (SD) age in lipid-lowering drugs users was 53.8 (10.7) years and in non-users it was 41.3 (10.4) years; $P<.0001$. Different characteristics were found between users and non-users group (Table 5). In 159,844 participants, users had a higher LDL-C (114.0 ± 42.1), higher TG level (179.0 ± 118.0), more likely to lower HDL-C level (49.7 ± 12.3) than non-users. Depending on whether participants were users or non-users of lipid-lowering drugs, using independent t-test to examine the characteristics of participants. Lipid-lowering drugs were significantly associated with TG, total cholesterol, and age. However, no significant association between lipid-lowering drugs and LDL-C was found.

Table 5. General characteristics of lipid-lowering drugs users and non-users

159,844 participant in KCPS-II biobank	Lipid-lowering drugs non users (n=157,047)	Lipid-lowering drugs users (n=2,797)	*P*-value
	Mean ± SD	Mean ± SD	
Age (yrs)	41.3 ± 10.4	53.8 ± 10.7	<.0001
LDL-C (mg/dl)	112.2 ± 31.3	114.0 ± 42.1	0.0261
HDL-C (mg/dl)	52.2 ± 10.9	49.7 ± 12.3	<.0001
TG (mg/dl)	134.0 ± 88.5	179.0 ± 118.0	<.0001
Total Cholesterol (mg/dl)	188.7 ± 32.5	191.0 ± 44.2	0.007
16,553 participant in KCPS-II biobank	lipid -lowering drugs non users (n=16,561)	lipid -lowering drugs users (n=434)	*P*-value
	Mean ± SD	Mean ± SD	
Age (yrs)	46.0 ± 11.1	54.8 ± 10.5	<.0001
LDL-C (mg/dl)	114.3 ± 31.7	117.7 ± 43.4	0.1009
HDL-C (mg/dl)	51.0 ± 10.7	49.2 ± 12.2	0.0029
TG (mg/dl)	140.3 ± 90.7	180.9 ± 119.4	<.0001
Total Cholesterol (mg/dl)	190.8 ± 33.1	195.7 ± 45.4	0.0274

Abbreviations: LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides
Data are expressed as mean(SD) unless otherwise indicated. *P*-value by chi-square test for continuous variables.

PART II. Mendelian randomization analysis of blood lipids and risk of acute myocardial infarction

1. Characteristics of individuals in case-cohort design

This study includes sub-cohort, which is randomly collected from the original cohort with all the cases. Figure 7 shows the illustration of the case-cohort study design. Ten subjects were both in sub-cohort and AMI incidence cases (Figure 7). A total 5,131 participants were available for the case-cohort design study population.

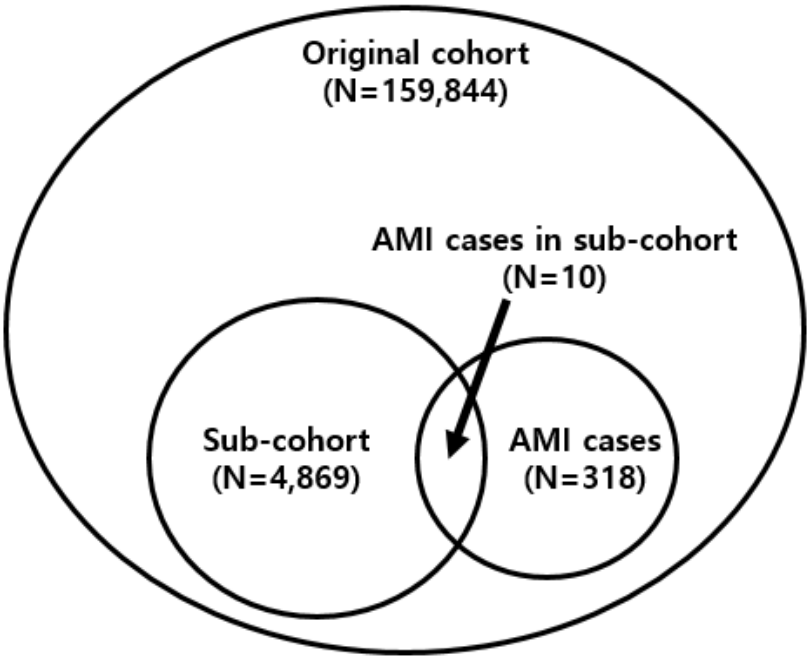


Figure 7. Graphic description of a case-cohort design (n=5,197)

Note: AMI; acute myocardial infarction

A total 5,197 participants (4,869 without acute myocardial infarction in sub-cohort and 328 Acute myocardial infarctions for whole cohort) from KCPS-II biobank, and measurements of their LDL-C, HDL-C, and TG levels as well as the genotypes, were included in the analysis. The characteristics of the study participants are presented in table 6. The mean age of AMI cases was older than control subjects (AMI 52.0 ± 11.5 years vs. control subjects 42.5 ± 9.0 years of age) (Table 6). Likewise, the proportion of women among AMI cases (13.4%) was lower than that of control subjects (33.4%). Cases had higher systolic blood pressure (mean ± SD, 127.4 ± 15.1 mmHg) than did control subjects (mean ± SD, 118.2 ± 14.2 mmHg), but levels of height, physical activity and education were almost similar between cases and control subjects (Table 6). Cases were more likely to have self-reported diabetes and hypertension at baseline, respectively (35.1%, 46.7%).

For serum lipids, AMI cases had higher mean concentrations of LDL-C (cases 130.4 ± 42.4 mg/dl) than did control subjects (114.0 ± 31.4 mg/dl; $P < .0001$). A similar pattern was observed for TG (cases 188.2 ± 129.9mg/dl, controls 141.0 ± 91.3mg/dl; $P < .0001$). In contrast, HDL-C concentrations were higher in control subjects (50.8 ± 10.0 mg/dl) than did AMI cases (45.0 ± 9.6 mg/dl; $P < .0001$).

Table 6. General characteristics of study subjects in KCPS-II study

	Incident Acute myocardial infarction*	Sub-cohort	*P*-value
N	328	4,869	
Female (%)	13.4	33.4	<.0001
Age and socioeconomic factors			
Age (yrs)	52.0 ± 11.5	42.5 ± 9.0	<.0001
College/university education (%)	59.3	64.8	0.1830
Income	32.9	45.9	0.0025
Lifestyle factors			
Smoke status (%)	45.8	31.3	<.0001
Physical activity status (%)	35.7	36.6	0.7440
Anthropometry and blood pressure			
Height (m)	167.3 ± 7.3	167.4 ± 8.4	0.7742
BMI (kg/m^2)	25.1 ± 2.9	23.8 ± 3.1	<.0001
Waist circumference (cm)	87.0 ± 7.7	81.3 ± 9.1	<.0001
Systolic BP (mmHg)	127.4 ± 15.1	118.2 ± 14.2	<.0001
Lipid profiles			
LDL cholesterol (mg/dl)	130.4 ± 42.4	114.0 ± 31.4	<.0001
HDL cholesterol (mg/dl)	45.0 ± 9.6	50.8 ± 10.0	<.0001
Triglycerides (mg/dl)	188.2 ± 129.9	141.0 ± 91.3	<.0001
Self-reported disease			

HTN (%)	46.7	17.7	<.0001
DM (%)	35.1	8.9	<.0001
Treatment of hyperlipidemia (%)	41.0	1.1	<.0001

Data are expressed as mean ± SD unless otherwise indicated. BMI indicates body mass index; BP, blood pressure; CI, confidence interval; HR, hazard ratio; LDL, low-density lipoprotein; HDL, high-density lipoprotein; HTN, hypertension and DM, diabetes mellitus
Analyzed using independent t-test for categorical variables. Analyzed using chi-square test for continuous variables.
*AMI incidence cases in sub-cohort included incident AMI group.

2. Genome-Wide Association with exposure variables

A genome-wide association with LDL-C, HDL-C, and TG in 16,553 participants from KCPS-II biobank identified several SNPs affecting these intermediate phenotypes. As shown in table 7-9, the list of SNPs relevant to LDL-C, HDL-C and TG, respectively were selected after excluding highly interrelated SNPs with linkage disequilibrium (LD) ($r^2 < 0.001$) from the GWAS. Finally, we screened out 8, 10, and 9 lipid-specific genetic variants for LDL-C, HDL-C, and TG, examining their associations with AMI risk. Manhattan plots of genome-wide analyses are presented in appendix 2-4.

We also extracted the SNPs associated with each lipid through GWAS among the participants of lipid-lowering drugs non-users. Compared with SNPs obtained from participants, including lipid-lowering drug users, there were no substantial differences, while the *P*-value has been lowered (Table 7-9).

Table 7. List of top SNPs strongly relevant to LDL-C before and after eliminating lipid-lowering drugs users

LDL-C GWAS															
Before eliminating participants who is taking lipid-lowering drugs								After eliminating participants who is taking lipid-lowering drugs							
SNP (n=8)	chr	Nearby gene	A1	A2	MAF	Beta	*P*-value	SNP (n=8)	chr	Nearby gene	A1	A2	MAF	Beta	*P*-value
rs157584	19	TOMM40	C	T	0.2891	-3.208	1.90E-16	rs157584	19	TOMM40	C	T	0.2891	-3.098	2.03E-15
rs600038	9	ABO	C	T	0.2633	2.891	1.72E-13	rs600038	9	ABO	C	T	0.2633	2.836	4.57E-13
rs7466988	9		C	T	0.2393	2.923	5.42E-12	rs7466988	9		C	T	0.2393	2.911	6.41E-12
rs2112653	5		T	C	0.4325	-2.525	6.40E-12	rs2112653	5		T	C	0.4325	-2.51	8.45E-12
rs77303550	16	TXNL4B	T	C	0.2423	-2.674	2.02E-10	rs77303550	16	TXNL4B	T	C	0.2423	-2.617	4.64E-10
rs12983082	19	LDLR	C	A	0.1348	3.244	5.16E-10	rs12983082	19	LDLR	C	A	0.1348	3.151	1.57E-09
rs138785751	2		T	C	0.1218	-3.105	1.92E-08	rs79272326	10		G	A	0.2758	-2.145	3.49E-08
rs58875530	10		C	T	0.2523	-2.24	3.52E-08	rs138785751	2		T	C	0.1218	-3.031	3.97E-08

Abbreviations: Beta indicated the effect of A1 allele compared with A2 allele; HDL-C, high-density lipoprotein; A1, reference allele; A2, coded allele; MAF, minor allele frequency; SNP, single nucleotide polymorphism; chr, chromosome; GWAS adjusted for age and sex.

Table 8. List of top SNPs strongly relevant to HDL-C before and after eliminating lipid-lowering drugs users

HDL-C GWAS														
Before eliminating participants who is taking lipid-lowering drugs								After eliminating participants who is taking lipid-lowering drugs						
SNP (n=10)	chr	Nearby gene	A1	A2	MAF	Beta	*P*-value	SNP (n=10)	chr	Nearby gene	A1	A2	MAF	Beta	*P*-value
rs7170361	15	LIPC	C	T	0.4643	1.029	5.40E-18	rs9958734	18	LIPG	C	T	0.4514	0.8394	1.77E-12
rs9958734	18	LIPG	C	T	0.4514	0.858	3.78E-13	rs1601935	15	ALDH1A2	G	T	0.3672	0.7834	3.01E-11
rs1240776	11		T	C	0.3118	0.7948	8.20E-11	rs1240776	11		T	C	0.3118	0.8063	5.72E-11
rs1601935	15	ALDH1A2	G	T	0.3672	0.7553	1.05E-10	rs157592	19		C	A	0.07412	-1.372	4.32E-10
rs157592	19		C	A	0.07412	-1.386	1.93E-10	rs2740488	9	ABCA1	C	A	0.251	-0.8007	5.18E-10
rs2740488	9	ABCA1	C	A	0.251	-0.7738	1.54E-09	rs60526148	8		T	C	0.1456	0.957	2.76E-09
rs4767014	12	RPH3A	C	T	0.2311	-0.9037	2.40E-09	rs4767014	12	RPH3A	C	T	0.2311	-0.8837	6.39E-09
rs404935	19	NECTIN2	A	G	0.1325	-0.9627	5.05E-09	rs11862052	16		T	C	0.1182	-1.016	1.55E-08
rs11862052	16		T	C	0.1182	-1.039	5.64E-09	rs76898656	15		A	G	0.2788	-0.7137	2.56E-08
rs60526148	8		T	C	0.1456	0.8786	3.89E-08	rs404935	19	NECTIN2	A	G	0.1325	-0.9195	3.21E-08

Abbreviations: Beta indicated the effect of A1 allele compared with A2 allele; HDL-C, high-density lipoprotein; A1, reference allele; A2, coded allele; MAF, minor allele frequency; SNP, single nucleotide polymorphism; chr, chromosome; GWAS adjusted for age and sex.

Table 9. List of top SNPs strongly relevant to TG before and after eliminating lipid-lowering drugs users

TG GWAS															
Before eliminating participants who is taking lipid-lowering drugs								After eliminating participants who is taking lipid-lowering drugs							
SNP (n=9)	chr	Nearby gene	A1	A2	MAF	Beta	*P*-value	SNP (n=8)	chr	Nearby gene	A1	A2	MAF	Beta	*P*-value
rs4665972	2	SNX17	C	T	0.4403	-8.548	1.94E-16	rs4665972	2	SNX17	C	T	0.4403	-8.069	1.02E-14
rs2954017	8		T	C	0.4098	7.53	1.31E-14	rs2954017	8		T	C	0.4098	7.325	6.67E-14
rs42132	7		T	C	0.0758	-12.61	2.78E-10	rs60526148	8	LOC107986 921	T	C	0.1456	-9.004	1.46E-10
rs60526148	8	LOC107986 921	T	C	0.1456	-8.482	1.56E-09	rs42132	7		T	C	0.0758	-12.29	7.76E-10
rs112269866	6		A	G	0.1937	7.522	2.24E-09	rs7480302	11		G	A	0.4196	5.5	1.77E-08
rs283811	19	NECTIN2	G	A	0.1788	7.401	5.36E-09	rs12663103	6		C	T	0.05835	11.51	2.84E-08
rs60500353	1	LOC105378 768	T	C	0.1899	-7.623	1.19E-08	rs283811	19	NECTIN2	G	A	0.1788	6.993	3.55E-08
rs12803249	11		G	A	0.1741	7.061	2.49E-08	rs60500353	1	LOC105378 768	T	C	0.1899	-7.296	4.64E-08
rs9267653	6	SLC44A4	T	C	0.3281	-5.691	3.96E-08								

Abbreviations: Beta indicated the effect of A1 allele compared with A2 allele; HDL-C, high-density lipoprotein; A1, reference allele; A2, coded allele; MAF, minor allele frequency; SNP, single nucleotide polymorphism; chr, chromosome; GWAS adjusted for age and sex.

3. Results of conventional Cox proportional hazard model and Mendelian randomization analyses in case-cohort design

Comparison with nested case-control design, the case-cohort study design is necessary to consider the oversampling of the cases with an event (Sharp, Poulaliou, Thompson, White, & Wood, 2014). Since the participants in a sub-cohort were randomly selected from the whole original cohort, the sub-cohort sampling fraction is the proportion of individuals in the original cohort. Thus, we conducted the Cox proportional hazard regression model with weighting to explore the association between blood lipids and the risk of AMI. We compared the result from the observational multivariable cox proportional hazard model with those from one-sample Mendelian randomization analysis. In these analyses, five sets of models were used. The basic model (Model 1) was a crude model, and model 2 was adjusted for age and sex. The multivariable model (Model 3) was adjusted for additional variables, including other lipids. Model 4 was adjusted for additional variables involving smoking status, diabetes, hypertension, and body mass index (BMI). Finally, model 5 additionally includes the use of lipid-lowering drugs.

Each weighted genetic risk score (WGRS) was composed using eight representative SNPs for LDL-C, ten representative SNPs for HDL-C, nine representative SNPs for TG. The F-statistic for the association of WGRS of each lipid with AMI was significantly high to conduct the Mendelian randomization after

adjusting for age, sex, smoking status, diabetes mellitus, hypertension, BMI, and use of lipid-lowering drugs (LDL-C F_{WGRS}=64.37; HDL-C F_{WGRS}=234.85; TG F_{WGRS}=301.81) (Table 10). These indicate that created instrument variables are strongly associated with exposures.

The LDL-C levels significantly increased the risk of AMI in the observational multivariate hazard ratio (HR_{LDL-C}=1.009; 95% confidence interval, 1.005-1.012) even after adjusting for confounding variables (Table 10). Also, serum TG suggested evidence for an adverse effect of AMI risk. However, HDL-C levels showed an inverse association with AMI risk in model 5 (HR_{HDL}=0.963; 95% confidence interval, 0.949-0.978).

In one-sample MR analysis, some evidence suggested a causal role of increased LDL-C on AMI risk (HR_{WGRS}=1.046; 95% confidence interval, 1.019-1.073). Other evidence revealed no causal association between serum HDL-C levels and risk of AMI (HR_{WGRS}=0.998; 95% confidence interval, 0.911-1.093). Moreover, our research showed no evidence of a causal relationship between serum TG level and AMI risk (HR_{WGRS}=1.493; 95% confidence interval, 0.310-7.188) (Table 10).

Moreover, these relationships were consistent in Mendelian randomization analysis using WGRS2, which estimates were calculated in the non-linear model. In table 10, the second method of calculating WGRS showed similar F-statistics as in the first method of calculating WGRS. According to MR analysis, the association

seemed stronger for WGRS1 than for WGRS2. However, the results of each lipid with WGRS1 were relatively similar to those of WGRS2.

Table 10. Association of acute myocardial infarction with lipids using Mendelian randomization (n=5,197)

Exposure variable	Model	Mendelian Randomization Analysis						Multivariable Cox regression Analysis	
		F-statistic WGRS1-X	HR (95% CI)	*P*-value	F-statistic WGRS2-X	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
LDL-C (mg/dl)	Model 1	96.48	1.041(1.016-1.067)	0.001	91.48	1.040(1.014-1.067)	0.000	1.015(1.012-1.018)	<.0001
	Model 2	100.18	1.043(1.019-1.068)	0.001	98.52	1.041(1.016-1.067)	0.000	1.012(1.009-1.015)	<.0001
	Model 3	116.92	1.042(1.016-1.069)	0.001	116.19	1.040(1.013-1.068)	0.003	1.008(1.005-1.011)	<.0001
	Model 4	70.73	1.047(1.022-1.073)	0.000	70.35	1.044(1.016-1.072)	0.001	1.009(1.005-1.012)	<.0001
	Model 5	64.37	1.046(1.019-1.073)	0.001	64.03	1.044(1.017-1.073)	0.001	1.009(1.005-1.012)	<.0001
HDL-C (mg/dl)	Model 1	90.63	1.000(0.921-1.086)	1.000	105.03	1.008(0.933-1.088)	0.847	0.929(0.916-0.941)	<.0001
	Model 2	387.51	1.008(0.927-1.096)	0.852	395.32	1.011(0.936-1.092)	0.778	0.942(0.929-0.956)	<.0001
	Model 3	457.79	1.004(0.925-1.111)	0.764	459.54	1.031(0.943-1.126)	0.502	0.963(0.949-0.977)	<.0001
	Model 4	260.99	0.999(0.912-1.094)	0.978	261.80	1.016(0.930-1.110)	0.719	0.963(0.949-0.978)	<.0001
	Model 5	234.85	0.998(0.911-1.093)	0.959	235.59	1.015(0.929-1.109)	0.739	0. 963(0.949-0.978)	<.0001
log-transformed TG (mg/dl)	Model 1	98.29	1.116(0.266-4.684)	0.881	123.41	1.220(0.338-4.409)	0.761	2.356(1.959-2.833)	<.0001
	Model 2	382.40	1.064(0.271-4.169)	0.930	390.45	1.088(0.307-3.864)	0.896	2.051(1.675-2.512)	<.0001
	Model 3	400.24	1.067(0.234-4.873)	0.933	405.40	1.051(0.257-4.304)	0.945	1.618(1.294-2.022)	<.0001
	Model 4	335.03	1.496(0.312-7.178)	0.615	337.69	1.363(0.309-6.005)	0.683	1.222(0.964-1.549)	0.0982
	Model 5	301.81	1.493(0.310-7.188)	0.617	304.21	1.374(0.312-6.060)	0.674	1.216(0.959-1.542)	0.1063

Note: HR, Hazard ratio; CI, Confidence interval; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides; WGRS1, weighted genetic risk score in linear modeling; WGRS2, weighted genetic risk score in non-linear modeling
Model 1: Crude model

Model 2: adjusted for age and sex
Model 3: model 2 plus additional adjustments for other lipid profiles
Model 4: model 3 plus additional adjustments for smoking status, diabetes, hypertension, and body mass index
Model 5: model 4 plus additional adjustments for treatment of hyperlipidemia

4. Association of intermediate phenotype and potential confounders with genetic variables

We tested WGRS used as the IV with risk factors of AMI to explain for potential confounders. Concerning age, sex, BMI, Waist circumference, blood lipids, smoking status, hypertension, diabetes mellitus, there were no significant differences with WGRS. On the whole, WGRS showed strong relationships with LDL-C, but these were not related to confounders (Table 11). Also, correlation analysis between WGRS of LDL-C and continuous variables was performed. Within the variables, WGRS of LDL-C showed a significant correlation with HDL-C (*P*-value 0.030) as well as LDL-C (*P*-value <.0001) (Table 11).

In table 12, WGRS of HDL-C showed a significant correlation with TG (*P*-value 0.0134) and smoking status (*P*-value 0.0030). In addition, WGRS of TG showed a significant association with HDL-C (*P*-value 0.0070) as well as TG (*P*-value <.0001) (Table 13).

Table 11. Associations between two types of LDL-C WGRSs and potential confounders in KCPS-II biobank

	WGRS		WGRS2	
	r or Mean(SD)	*P*-value	r or Mean(SD)	*P*-value
LDL-C (mg/dl)	0.13521	<.0001	0.13219	<.0001
Age (yrs)	-0.01463	0.2917	-0.01555	0.2623
BMI (kg/m2)	0.0084	0.5449	0.00473	0.7333
Waist circumference (cm)	0.00112	0.9359	-0.00544	0.6954
HDL-C (mg/dl)	-0.03066	0.0272	-0.02708	0.0511
TG (mg/dl)	0.01982	0.1533	0.01213	0.3823
Smoking status (%)				
Nonsmoker	-3.0 ± 5.5	0.9889	0.5 ± 8.7	0.5995
Smoker	-3.0 ± 5.6		0.3 ± 8.7	
Hypertension (%)				
No	-3.0 ± 5.5	0.9324	0.5 ± 8.8	0.7976
Yes	-3.0 ± 5.5		0.4 ± 8.7	
Diabetes mellitus (%)				
No	-3.0 ± 5.5	0.6281	0.5 ± 8.8	0.9774
Yes	-2.9 ± 5.4		0.4 ± 8.6	
Female (%)	-3.1 ± 5.6	0.2154	0.2 ± 8.8	0.2331

Note: WGRS, weighted genetic risk score; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides; WGRS1, weighted genetic risk score in linear modeling; WGRS2, weighted genetic risk score in non-linear modeling
P-value from ANOVA for continuous variables or from χ^2 test for categorical variables.

Table 12. Associations between two types of HDL-C WGRSs and potential confounders in KCPS-II biobank

	WGRS		WGRS2	
	r or Mean(SD)	P-value	r or Mean(SD)	P-value
HDL-C (mg/dl)	0.13087	<.0001	0.14043	<.0001
Age (yrs)	0.0122	0.3792	0.01371	0.3230
BMI (kg/m^2)	0.00797	0.5658	-0.00531	0.7022
Waist circumference (cm)	0.01826	0.1890	0.01203	0.3870
LDL-C (mg/dl)	-0.00628	0.6509	-0.0199	0.1517
TG (mg/dl)	-0.03432	0.0134	-0.05019	0.0003
Smoking status (%)				
Nonsmoker	0.7 ± 1.7	0.0030	2.3 ± 2.6	0.0021
Smoker	0.6 ± 1.7		2.0 ± 2.6	
Hypertension (%)				
No	0.7 ± 1.7	0.6777	2.2 ± 2.6	0.9420
Yes	0.7 ± 1.7		2.2 ± 2.6	
Diabetes mellitus (%)				
No	0.7 ± 1.7	0.6280	2.2 ± 2.6	0.7386
Yes	0.7 ± 1.6		2.2 ± 2.6	
Female (%)	0.7 ± 1.6	0.7092	2.2 ± 2.6	0.8746

Note: WGRS, weighted genetic risk score; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides; WGRS1, weighted genetic risk score in linear modeling; WGRS2, weighted genetic risk score in non-linear modeling
P -value from ANOVA for continuous variables or from χ^2 test for categorical variables.

Table 13. Associations between two types of TG WGRSs and potential confounders in KCPS-II biobank

	WGRS		WGRS2	
	r or Mean(SD)	P-value	r or Mean(SD)	P-value
TG (mg/dl)	0.13568	<.0001	0.15194	<.0001
Age (yrs)	-0.02028	0.1438	-0.01619	0.2431
BMI (kg/m2)	-0.00705	0.6114	0.00487	0.7259
Waist circumference (cm)	-0.00841	0.5452	0.00015	0.9912
HDL-C (mg/dl)	-0.03741	0.0070	-0.04499	0.0012
LDL-C (mg/dl)	0.00258	0.8525	0.00724	0.6018
Smoking status (%)				
Nonsmoker	-0.0 ± 0.1	0.0890	0.0 ± 0.1	0.0120
Smoker	-0.0 ± 0.1		0.0 ± 0.1	
Hypertension (%)				
No	-0.0 ± 0.1	0.3681	0.0 ± 0.1	0.1822
Yes	-0.0 ± 0.1		0.0 ± 0.1	
Diabetes mellitus (%)				
No	-0.0 ± 0.1	0.2742	0.0 ± 0.1	0.1367
Yes	-0.0 ± 0.1		0.0 ± 0.1	
Female (%)	-0.0 ± 0.1	0.9650	0.0 ± 0.1	0.3384

Note: WGRS, weighted genetic risk score; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides; WGRS1, weighted genetic risk score in linear modeling; WGRS2, weighted genetic risk score in non-linear modeling

P -value from ANOVA for continuous variables or from χ^2 test for categorical variables.

5. Sensitivity analyses

To examine the robustness of the results, we investigated several sensitivity analyses. In table 14, we performed both multivariable regression and one-sample MR analysis to assess the causal effect of lipid traits in all participants of KCPS-II with genetic information, excluding AMI prevalent cases (n=16,257). After adjusted all confounding variables, LDL-C significantly increased the risk of AMI in the observational multivariable hazard ratio ($HR_{LDL\text{-}C}$=1.010; 95% confidence interval, 1.007-1.013) (Table 14). These positive associations were coincident with Mendelian randomization analysis using WGRS (HR_{WGRS}=1.042; 95% confidence interval, 1.015-1.070) (Table 14). Also, in serum TG, it showed a significantly positive association with the risk of AMI in the observational analysis (HR_{TG}=1.325; 95% confidence interval, 1.057-1.661) while this result was not significant with Mendelian randomization (HR_{WGRS}=0.980; 95% confidence interval, 0.184-5.203) (Table 14).

On the other hand, in HDL-C level, a protective effect on the risk of AMI ($HR_{HDL\text{-}C}$=0.969; 95% confidence interval, 0.956-0.982) was found in multivariable Cox regression analysis (Table 14). In Mendelian randomization analysis, however, a protective effect wasn't significant using the IV method (HR_{WGRS}=0.992; 95% confidence interval, 0.915-1.074) (Table 14).

We also restricted the analyses to participants with cancer and atherosclerotic

cardiovascular disease (ASCVD). We found consistent evidence for the association between lipid traits and AMI risk (Table 15, 16). Further sensitivity analysis was undertaken after excluding participants who reported taking lipid-lowering drugs in the one-sample MR analysis (Table 17).

In table 17, LDL-C levels significantly increased the risk of AMI in the observational multivariate hazard ratio (HR_{LDL-C}=1.009; 95% confidence interval, 1.005-1.012) even after adjusting for confounding variables (Model 4). Also, serum TG showed a positive association with AMI risk (HR_{TG} =1.281; 95% confidence interval, 1.006-1.661). However, HDL-C levels showed an inverse association with AMI risk in model 4 (HR_{HDL}=0.969; 95% confidence interval, 0.956-0.982).

In one-sample MR analysis, the consistent results were shown after excluding lipid-lowering drug users (Table 17). Some evidence suggested a causal role of increased LDL-C on AMI risk (HR_{WGRS}=1.046; 95% confidence interval, 1.019-1.073), while other evidence revealed no causal association between serum HDL-C levels and risk of AMI (HR_{WGRS}=0.998; 95% confidence interval, 0.911-1.093). Moreover, our research showed no evidence of a causal relationship between serum TG level and AMI risk (HR_{WGRS}=1.493; 95% confidence interval, 0.310-7.188).

Table 14. Association of acute myocardial infarction with lipids using Mendelian randomization in all study population (n=16,257)

Exposure variable	Model	Mendelian Randomization Analysis						Multivariable Cox regression Analysis	
		F-statistic WGRS1-X	HR (95% CI)	P-value	F-statistic WGRS2-X	HR (95% CI)	P-value	HR (95% CI)	P-value
LDL-C (mg/dl)	Model 1	271.68	1.044(1.018-1.070)	0.001	266.92	1.042(1.016-1.069)	0.001	1.014(1.011-1.017)	<.0001
	Model 2	191.64	1.044(1.019-1.070)	0.001	189.69	1.042(1.016-1.068)	0.001	1.012(1.009-1.016)	<.0001
	Model 3	232.75	1.039(1.012-1.066)	0.004	232.06	1.037(1.010-1.065)	0.007	1.009(1.006-1.013)	<.0001
	Model 4	140.69	1.042(1.014-1.070)	0.002	140.31	1.040(1.007-1.013)	0.004	1.010(1.007-1.013)	<.0001
	Model 5	126.85	1.042(1.015-1.070)	0.002	126.53	2.927(1.415-6.057)	0.004	1.010(1.007-1.013)	<.0001
HDL-C (mg/dl)	Model 1	310.93	0.995(0.926-1.070)	0.895	326.11	1.003(0.933-1.079)	0.932	0.935(0.924-0.947)	<.0001
	Model 2	900.23	0.995(1.019-1.070)	0.884	903.85	1.005(0.933-1.081)	0.904	0.951(0.939-0.963)	<.0001
	Model 3	1213.23	1.003(0.927-1.086)	0.940	1211.32	1.016(0.937-1.103)	0.694	0.969(0.956-0.982)	<.0001
	Model 4	678.82	0.992(0.915-1.074)	0.834	677.56	1.004(0.925-1.090)	0.923	0.969(0.956-0.982)	<.0001
	Model 5	611.33	0.992(0.915-1.074)	0.836	610.23	1.004(0.925-1.090)	0.921	0.969(0.956-0.982)	<.0001
log-transformed TG (mg/dl)	Model 1	287.18	1.08(0.246-4.736)	0.919	326.35	1.132(0.281-4.551)	0.862	2.325(1.938-2.789)	<.0001
	Model 2	805.22	0.987(0.226-4.303)	0.986	865.24	1.047(0.262-4.190)	0.948	1.905(1.568-2.315)	<.0001
	Model 3	1087.99	0.811(0.157-4.182)	0.802	1097.56	0.880(0.185-4.175)	0.872	1.516(1.217-1.888)	0.0002

Model 4	796.62	0.811(0.157-4.182)	0.802	801.67	1.057(0.214-5.220)	0.945	1.325(1.057-1.661)	0.0145
Model 5	718.82	0.980(0.184-5.203)	0.981	723.38	1.054(0.213-5.217)	0.948	1.325(1.057-1.661)	0.0147

Note: HR, Hazard ratio; CI, Confidence interval; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides; WGRS1, weighted genetic risk score in linear modeling; WGRS2, weighted genetic risk score in non-linear modeling
Model 1: Crude model
Model 2: adjusted for age and sex
Model 3: model 2 plus additional adjustments for other lipid profiles
Model 4: model 3 plus additional adjustments for smoking status, diabetes, hypertension, and body mass index
Model 5: model 4 plus additional adjustments for treatment of hyperlipidemia

Table 15. Association of acute myocardial infarction with lipids using Mendelian randomization in all cancer participants (n=6,258)

Exposure variable	Model	Mendelian Randomization Analysis		
		HR(95% CI)	P-value	F-statistic WGRS-X
LDL-C (mg/dl)	Model 1	1.040(1.015-1.066)	0.002	107.77
	Model 2	1.041(1.016-1.067)	0.001	57.67
	Model 3	1.035(1.008-1.062)	0.010	74.17
	Model 4	1.037(1.008-1.062)	0.010	42.70
	Model 5	1.040(1.013-1.069)	0.004	38.53
HDL-C (mg/dl)	Model 1	1.007(0.952-1.065)	0.804	181.62
	Model 2	1.012(0.954-1.073)	0.691	310.49
	Model 3	1.021(0.959-1.088)	0.519	446.65
	Model 4	1.020(0.959-1.088)	0.519	243.83
	Model 5	1.010(0.946-1.077)	0.771	220.06
log-transformed TG (mg/dl)	Model 1	1.347(0.298-6.081)	0.698	104.32
	Model 2	1.166(0.255-5.333)	0.844	304.39
	Model 3	0.931(0.165-5.242)	0.935	410.79
	Model 4	0.966(0.166-5.614)	0.969	281.61
	Model 5	0.960(0.163-5.643)	0.964	254.28

Note: HR, Hazard ratio; CI, Confidence interval; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides
Model 1: Crude model
Model 2: adjusted for age and sex
Model 3: model 2 plus additional adjustments for other lipid profiles
Model 4: model 3 plus additional adjustments for smoking status, diabetes, hypertension, and body mass index
Model 5: model 4 plus additional adjustments for treatment of hyperlipidemia

Table 16. Association of acute myocardial infarction with lipids using Mendelian randomization in ASCVD participants (n=3,723)

Exposure variable	Model	Mendelian Randomization Analysis		
		HR(95% CI)	P-value	F-statistic WGRS-X
LDL-C (mg/dl)	Model 1	1.056(1.027-1.087)	0.000	40.01
	Model 2	1.054(1.024-1.084)	0.000	14.43
	Model 3	1.050(1.018-1.082)	0.002	27.26
	Model 4	1.054(1.021-1.088)	0.001	17.46
	Model 5	1.054(1.021-1.088)	0.001	15.73
HDL-C (mg/dl)	Model 1	0.971(0.917-1.028)	0.309	109.79
	Model 2	0.969(0.916-1.026)	0.278	133.97
	Model 3	0.971(0.914-1.032)	0.342	224.87
	Model 4	0.963(0.907-1.022)	0.216	128.65
	Model 5	0.964(0.908-1.023)	0.229	115.83
log-transformed TG (mg/dl)	Model 1	0.917(0.200-4.205)	0.911	63.95
	Model 2	0.794(0.164-3.851)	0.775	66.95
	Model 3	0.573(0.090-3.651)	0.556	158.60
	Model 4	0.754(0.115-4.930)	0.768	129.26
	Model 5	0.771(0.117-5.068)	0.786	116.34

Note: HR, Hazard ratio; CI, Confidence interval; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides
Model 1: Crude model
Model 2: adjusted for age and sex
Model 3: model 2 plus additional adjustments for other lipid profiles
Model 4: model 3 plus additional adjustments for smoking status, diabetes, hypertension, and body mass index
Model 5: model 4 plus additional adjustments for treatment of hyperlipidemia

Table 17. Association of acute myocardial infarction with lipids using Mendelian randomization, excluding lipid-lowering drugs users (n=5,131)

Exposure variable	Model	Mendelian Randomization Analysis						Multivariable Cox regression Analysis	
		F-statistic WGRS1-X	HR (95% CI)	P-value	F-statistic WGRS2-X	HR (95% CI)	P-value	HR (95% CI)	P-value
LDL-C (mg/dl)	Model 1	90.12	1.046(1.020-1.072)	0.000	87.32	1.047(1.021-1.073)	0.000	1.015(1.012-1.018)	<.0001
	Model 2	100.04	1.042(1.017-1.068)	0.001	99.12	1.044(1.018-1.070)	0.001	1.012(1.009-1.015)	<.0001
	Model 3	117.68	1.043(1.016-1.071)	0.002	117.38	1.044(1.015-1.073)	0.002	1.008(1.005-1.011)	<.0001
	Model 4	70.83	1.042(1.014-1.071)	0.003	70.66	1.043(1.014-1.073)	0.003	1.009(1.005-1.012)	<.0001
HDL-C (mg/dl)	Model 1	69.74	0.989(0.901-1.086)	0.821	80.67	0.989(0.899-1.089)	0.823	0.928(0.916-0.941)	<.0001
	Model 2	373.90	0.991(0.903-1.087)	0.843	377.68	0.995(0.907-1.091)	0.913	0.942(0.929-0.956)	<.0001
	Model 3	443.02	0.998(0.897-1.111)	0.970	441.92	1.011(0.899-1.137)	0.859	0.963(0.949-0.978)	<.0001
	Model 4	252.67	0.983(0.882-1.095)	0.751	251.84	0.994 (0.881-1.121)	0.917	0.963(0.948-0.978)	<.0001
log-transformed TG (mg/dl)	Model 1	90.11	0.489(0.111-2.149)	0.344	114.98	0.474(0.127-1.776)	0.268	2.400(1.987-2.899)	<.0001
	Model 2	373.39	0.619(0.157-2.435)	0.493	381.78	0.518(0.149-1.804)	0.301	2.105(1.712-2.588)	<.0001
	Model 3	388.30	0.426(0.088-2.066)	0.290	393.05	0.339(0.080-1.440)	0.143	1.661(1.322-2.086)	<.0001
	Model 4	328.68	0.531(0.101-2.806)	0.456	331.17	0.380(0.080-1.806)	0.224	1.281(1.006-1.632)	0.044

Note: HR, Hazard ratio; CI, Confidence interval; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein; TG, triglycerides; WGRS1, weighted genetic risk score in linear modeling; WGRS2, weighted genetic risk score in non-linear modeling
Model 1: Crude model
Model 2: adjusted for age and sex
Model 3: model 2 plus additional adjustments for other lipid profiles
Model 4: model 3 plus additional adjustments for smoking status, diabetes, hypertension, and body mass index

IV. DISCUSSION

Our study, using Mendelian randomization approach to minimize the possibility of reverse causality and confounding effect, could overcome the potential limitation of conventional epidemiology study. Also, in this study, observational estimates from multivariable Cox regression were compared with those from MR analyses to infer a causal relationship in which three lipid traits may affect AMI risk.

In multivariable Cox regression analysis using data on AMI incidence in the KCPS-II biobank study, LDL-C and TG were positively associated with AMI, whereas there was evidence for a protective effect of HDL-C. One-sample MR analysis in KCPS-II biobank, which used genetic variants associated with lipid fractions, provided some evidence for adverse effect of increased LDL-C on AMI. However, there was little evidence for an association with HDL-C and TG. Moreover, sensitivity analyses showed no substantial change while we examine in several study populations. These results were supported by sensitivity analyses accounting for consistent evidence for LDL-C.

1. Comparison with other studies

Previous studies have confirmed that SNPs affecting LDL-C as a causal risk factor for risk of AMI and have cast doubt on whether SNPs associated with HDL-C directly influence risk for AMI. Otherwise, findings for the associations between AMI risk and serum TG has been conflicting with two showing no conclusive evidence that TG is associated with AMI risk and showing evidence of an adverse effect. One meta-analysis, outcome assessed for AMI showed no significant association between serum TG and myocardial infarction risk (Sarwar et al., 2010). Findings from other Mendelian randomization, genetically raised TG due to *APOA5* genetic variants were associated with increased risk of myocardial infarction (Jørgensen et al., 2013).

Using the same approach as other lipid traits have made it difficult to define whether or not the serum triglyceride levels are causal in AMI. This is because, in contrast to genetic variants associated with LDL-C and HDL-C, almost all SNPs determined for serum triglycerides have extra effects on either LDL-C or HDL-C. This phenomenon called “pleiotropy”, refers to a genetic variant being associated with multiple phenotypic traits (Allara et al., 2019; Do et al., 2013). For lipid transport, cholesterol is mainly carried in LDL or HDL, while triglycerides are mostly transported in VLDL, chylomicron, and their metabolism (Nordestgaard & Varbo, 2014).

Recently, fewer studies have investigated causal associations between blood lipids with AMI risk. Most studies elucidate that causative role of TG in overall heart disease, such as ischemic heart disease (IHD), coronary artery disease (CAD), and other cardiovascular diseases (Do et al., 2013; Nordestgaard, 2016). In a meta-analysis of Mendelian randomization studies, which has strong evidence for association between genetic determinants of lipids and cardiovascular disease risk, suggested genetic effects on TG are concordant with the expected risk of cardiovascular disease (Nordestgaard & Tybjærg-Hansen, 2011). Findings from other Mendelian randomization researches have been shown consistent results with a causal association between elevated levels of TG and increased risk of ischemic heart disease (IHD), coronary artery disease (CAD), and other cardiovascular diseases (Budoff, 2016; Rosenson, Davidson, Hirsh, Kathiresan, & Gaudet, 2014).

In addition, most of these studies have been conducted in European ancestry, further work was required to investigate whether these findings apply to participants in East Asian ancestry. Unlike previous studies which were performed in European ancestry, our study suggests different evidence for association between serum TG and risk of AMI. In Korean participants, the Mendelian randomization analysis showed no conclusive evidence that TG is causally associated with AMI.

We also found differences in the lists of SNPs for blood lipids already published in Western (Nordestgaard & Tybjærg-Hansen, 2011; Voight et al., 2012). However,

some genes, such as *ABO, LDLR*, which are well known for LDL-C, were also found in KCPS-II data (Table 7) (Y. J. Kim et al., 2011). There are several genes that have interactions between statin efficacy such as *APOE*, *HMGCR*, *ABCB1*, *SLCO1B1*, *CETP*, *PCSK9*, *KIF6*, *SORT1*/*CELSR2*/*PRSC1*, and *ABCA1* (Leusink, Onland-Moret, De Bakker, De Boer, & Maitland-Van Der Zee, 2016). Within the genes, *PCSK9*, *APOE*, *HMGCR* are known to influence LDL-C response to statin therapy (Ruiz-Iruela et al., 2019). We investigated why these genes were not detected in KCPS-II data. In the case of *PCSK9*, we found genetic variants associated with LDL-C trait were excluded through genome-wide significance ($P<5*10^{-8}$). Also, SNPs associated with *HMGCR* were removed thorough the process of excluding highly interrelated SNPs with linkage disequilibrium (LD) ($r^2 < 0.001$) from the GWAS.

2. Strength and limitations of this study

Our study has several advantages. First, since our research is based on a case-cohort study design from a large prospective cohort study, a broad individual dataset has many advantages rather than using a summary dataset. Subgroup analysis and effect moderation cannot be tested with summary data from large GWAS consortia (Lawlor, 2016). Also, compared to the nested case-control design, the case-cohort

design has several benefits. It is more efficient in terms of cost and sub-cohort can be used as a comparison group to study different disease outcomes.

Second, our paper provides a comparison of conventional epidemiological findings and Mendelian randomization analysis findings. This comparison could be possible as we performed an investigation in a single population sample (Burgess et al., 2019). Multivariable Cox regression, one-sample MR and two-sample MR approaches have different strengths and limitations with regard to main bias sources. (Appendix 1). The method of multivariable analysis alleviates the effect of bias, such as reverse causation and confounding (Richmond et al., 2019). However, residual or unmeasured confounding, selection bias, and measurement errors can also be estimates of biased effect estimates (Lawlor, 2016). In one-sample MR analysis, it can minimize the potential for bias due to measurement errors and confounders. Moreover, it is able to evaluate individual-level confounding factors thoroughly. Nevertheless, in one-sample MR, there's a possibility of "collider bias" due to study sampling and weak instrument biases towards the confounded regression analysis result. In the case of two-sample MR analysis, due to using two samples, sample size and statistical power are better improved than one-sample MR. However, the difficulty of performing subgroup analysis with summary data from large GWAS Consortium remains a limitation.

The study's main strength is the first attempt of one-sample Mendelian

randomization analysis, which explores the association between all three lipid traits and acute myocardial infarction in the Korean population. Up to now, there was little evidence of the causal relationship between blood lipids and the risk of acute myocardial infarction in Korea, while many studies have been announced in Western.

Furthermore, potential limitations can be derived from our study. First, there is an inevitable limitation in one-sample Mendelian randomization concerned with weak instrument bias. However, by using a calculated weighted genetic risk score (WGRS), we tried to increase the statistical power of this study.

Second, limitation relates to measurement errors in this research. Since the data in this study is secondary data, serum lipids and other variables were measured only once, so it may not be accurate and may offer little value. Moreover, there′s no description of any duration of using lipid-lowering drugs, because it was a self-reported data. In the future, replication research with larger samples and accurate results data will be needed.

Third, MR assumptions were not perfectly satisfied. One assumption that the genetic IV is robustly associated with the exposure of interest, is satisfied because IVs were identified from the GWAS on lipids. However, the other two assumptions include that the genetic IV must not associated with potential confounders, and can

only affect the outcome via the modifiable exposure, were not fully satisfied. In our study, we tested the association of WGRS with some potential AMI risk factors, and found that each WGRS of lipid was associated with other lipid traits or other confounders. This result means that the potential bias exists in the estimate of the effect of lipids on AMI risk, and the true causal effect is potentially underestimated. Also, using a large number of SNPs as instrument variables increase the chance to induce bias caused by pleiotropy in MR results. However, the implementation of additional MR analyses with different sensitivity to these pleiotropic effects, which provide similar risk estimates, has brought robustness to our results.

Potential existence of pleiotropy is due to the mechanism of lipids, three lipid traits affect each other considerably, as we can see the correlation between WGRS of LDL-C and serum HDL-C level (Table 11). Also, the biological role of many genetic variants used to measure these characteristics by MR and the underlying mechanical pathways of observed effects is not well known. Additional research of two-sample Mendelian randomization with sensitivity analysis approaches to clarify the potential pleiotropy for these mediators may be required (Burgess et al., 2019).

Despite of limitation, this study gave an explanation of role of blood lipids on AMI risk in the Korean population. To explain the specific role of TG in AMI risk, an additional study on subclasses of lipids, measured by nuclear magnetic

resonance spectroscopy samples, should be considered in the Korean population (Kamstrup, Tybjærg-hansen, Steffensen, & Nordestgaard, 2009; Y. J. Kim et al., 2011).

VI. CONCLUSION

In conclusion, our study provides a significant causal association between LDL-C and the risk of AMI using a population-based cohort. Our research using Mendelian randomization analysis is not restricted to the possible limitations of conventional epidemiology studies such as reverse causation and confounding effect.

The present study is the first Mendelian randomization study regarding AMI risk and blood lipids for the Asian population. The highly consistent results of MR in the Korean population suggest that lower LDL-C are still likely to have a net benefit for preventing overall AMI risk in the Korean population as well as in Western. According to our results, we suggest exploring the mechanism of TG on AMI risk and a possible reason for non-significant results in TG should be in the further research area.

REFERENCES

Allara E, Morani G, Carter P, Gkatzionis A, Zuber V, Foley CN, Rees JMB, Mason AM, Bell S, Gill D, Lindström S, Butterworth AS, Di Angelantonio E, Peters J, Burgess S; INVENT consortium. Genetic Determinants of Lipids and Cardiovascular Disease Outcomes: A Wide-Angled Mendelian Randomization Investigation. Circ Genom Precis Med. 2019 Dec;12(12):e002711.

Budoff M. Triglycerides and triglyceride-rich lipoproteins in the causal pathway of cardiovascular disease. Am J Cardiol. 2016;118(1):138–45.

Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations. Wellcome Open Res. 2019;4:186.

Das S, Abecasis GR, Browning BL. Genotype Imputation from Large Reference Panels. Annu Rev Genomics Hum Genet. 2018;19(1):73–96.

Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkilä K, Hyppönen E, Isaacs A, Jackson AU, Johansson A, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikäinen LP, Magnusson PK, Mangino M, Mihailov E, Montasser ME, Müller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W,

Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney AS, Döring A, Elliott P, Epstein SE, Eyjolfsson GI, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P, Kastelein JJ, Khaw KT, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimäki T, Lin SY, Lindström J, Loos RJ, Mach F, McArdle WL, Meisinger C, Mitchell BD, Müller G, Nagaraja R, Narisu N, Nieminen TV, Nsubuga RN, Olafsson I, Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruokonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stančáková A, Stirrups K, Swift AJ, Tiret L, Uitterlinden AG, van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen YD, Collins FS, Cooper RS, Danesh J, Dedoussis G, de Faire U, Feranil AB, Ferrières J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllensten U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Järvelin MR, Jula A, Kähönen M, Kaprio J, Kesäniemi A, Kivimaki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, März W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njølstad I, Pedersen NL, Power C, Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PE, Sheu WH, Shuldiner AR, Siegbahn A, Spector TD,

Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BH, Altshuler D, Ordovas JM, Boerwinkle E, Palmer CN, Thorsteinsdottir U, Chasman DI, Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Mohlke KL, Ingelsson E, Abecasis GR, Daly MJ, Neale BM, Kathiresan S. Common variants associated with plasma triglycerides and risk for coronary artery disease. Nat Genet. 2013 Nov;45(11):1345-52.

HIRA. 최근 5년간 "허혈성 심질환" 진료경향 분석. HIRA 11권 2호. 2017;59–67.

Holmes MV, Asselbergs FW, Palmer TM, Drenos F, Lanktree MB, Nelson CP, Dale CE, Padmanabhan S, Finan C, Swerdlow DI, Tragante V, van Iperen EP, Sivapalaratnam S, Shah S, Elbers CC, Shah T, Engmann J, Giambartolomei C, White J, Zabaneh D, Sofat R, McLachlan S; UCLEB consortium, Doevendans PA, Balmforth AJ, Hall AS, North KE, Almoguera B, Hoogeveen RC, Cushman M, Fornage M, Patel SR, Redline S, Siscovick DS, Tsai MY, Karczewski KJ, Hofker MH, Verschuren WM, Bots ML, van der Schouw YT, Melander O, Dominiczak AF, Morris R, Ben-Shlomo Y, Price J, Kumari M, Baumert J, Peters A, Thorand B, Koenig W, Gaunt TR, Humphries SE, Clarke R, Watkins H, Farrall M, Wilson JG, Rich SS, de Bakker PI, Lange LA, Davey Smith G, Reiner AP, Talmud PJ, Kivimäki M, Lawlor DA, Dudbridge F, Samani NJ, Keating BJ, Hingorani AD, Casas JP. Mendelian randomization of blood lipids for coronary heart disease. Eur Heart J. 2015 Mar 1;36(9):539-50.

Hong JS, Kang HC, Lee SH, Kim J. Long-term trend in the incidence of acute myocardial infarction in Korea: 1997-2007. Korean Circ J. 2009;39(11):467–

76.

Jee YH, Emberson J, Jung KJ, Lee SJ, Lee S, Back JH, Hong S, Kimm H, Sherliker P, Jee SH, Lewington S. Cohort Profile: The Korean Cancer Prevention Study-II (KCPS-II) Biobank. Int J Epidemiol. 2018 Apr 1;47(2):385-386f.

Jee YH, Lee SJ, Jung KJ, Jee SH. Alcohol intake and serum glucose levels from the perspective of a mendelian randomization design: The KCPS-II Biobank. PLoS One. 2016;11(9):1–9.

Jørgensen AB, Frikke-Schmidt R, West AS, Grande P, Nordestgaard BG, Tybjæ rg-Hansen A. Genetically elevated non-fasting triglycerides and calculated remnant cholesterol as causal risk factors for myocardial infarction. Eur Heart J. 2013;34(24):1826–33.

Kamstrup PR, Tybjæ rg-hansen A, Steffensen R, Nordestgaard BG. Genetically Elevated Lipoprotein(a) and Increased Risk of Myocardial Infarction. J Am Med Assoc. 2009;301(22):2331–9.

Kawashiri M aki, Tada H, Nomura A, Yamagishi M. Mendelian randomization: Its impact on cardiovascular disease. J Cardiol [Internet]. 2018;72(4):307–13.

Khera A V., Kathiresan S. Genetics of coronary artery disease: Discovery, biology and clinical translation. Nat Rev Genet [Internet]. 2017;18(6):331–44.

Kim Y, Ahn Y, Cho MC, Kim CJ, Kim YJ, Jeong MH. Current status of acute myocardial infarction in Korea. Korean J Intern Med. 2019;34(1):1–10.

Kim YJ, Go MJ, Hu C, Hong CB, Kim YK, Lee JY, et al. Large-scale genome-

wide association studies in east Asians identify new genetic loci influencing metabolic traits. Nat Genet. 2011;43(10):990–7.

Kimm H, Yun JE, Lee SH, Jang Y, Jee SH. Validity of the diagnosis of acute myocardial infarction in Korean National Medical Health Insurance claims data: The Korean Heart Study. Korean Circ J. 2012;42(1):10–5.

Lawlor DA. Commentary: Two-sample Mendelian randomization: Opportunities and challenges. Int J Epidemiol. 2016;45(3):908–15.

Lee SJ, Jee YH, Jung KJ, Hong S, Shin ES, Jee SH. Bilirubin and Stroke Risk Using a Mendelian Randomization Design. Stroke. 2017;48(5):1154–60.

Leusink M, Onland-Moret NC, De Bakker PIW, De Boer A, Maitland-Van Der Zee AH. Seventeen years of statin pharmacogenetics: A systematic review. Pharmacogenomics. 2016;17(2):163–80.

Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet. 2010;11(7):499–511.

Moon S, Kim YJ, Han S, Hwang MY, Shin DM, Park MY, Lu Y, Yoon K, Jang HM, Kim YK, Park TJ, Song DS, Park JK, Lee JE, Kim BJ. The Korea Biobank Array: Design and Identification of Coding Variants Associated with Blood Biochemical Traits. Sci Rep. 2019 Feb 4;9(1):1382.

Nitsch D, Molokhia M, Smeeth L, Destavola BL, Whittaker JC, Leon DA. Special Article Limits to Causal Inference based on Mendelian Randomization : A Comparison with Randomized Controlled Trials. 2006;163(5):397–403.

Nordestgaard BG. Triglyceride-Rich Lipoproteins and Atherosclerotic Cardiovascular Disease: New Insights from Epidemiology, Genetics, and Biology. Circ Res. 2016;118(4):547–63.

Nordestgaard BG, Tybjærg-Hansen A. Genetic determinants of LDL, lipoprotein(a), triglyceride-rich lipoproteins and HDL: Concordance and discordance with cardiovascular disease risk. Curr Opin Lipidol. 2011;22(2):113–22.

Nordestgaard BG, Varbo A. Triglycerides and cardiovascular disease. Lancet. 2014;384(9943):626–35.

Richmond RC, Anderson EL, Dashti HS, Jones SE, Lane JM, Strand LB, Brumpton B, Rutter MK, Wood AR, Straif K, Relton CL, Munafò M, Frayling TM, Martin RM, Saxena R, Weedon MN, Lawlor DA, Smith GD. Investigating causal relations between sleep traits and risk of breast cancer in women: mendelian randomisation study. BMJ. 2019 Jun 26;365:l2327.

Rosenson RS, Davidson MH, Hirsh BJ, Kathiresan S, Gaudet D. Genetics and causality of triglyceride-rich lipoproteins in atherosclerotic cardiovascular disease. J Am Coll Cardiol. 2014;64(23):2525–40.

Ruiz-Iruela C, Candás-Estébanez B, Pintó-Sala X, Baena-Díez N, Caixàs-Pedragós A, Güell-Miró R, Navarro-Badal R, Calmarza P, Puzo-Foncilla JL, Alía-Ramos P, Padró-Miquel A. Genetic contribution to lipid target achievement with statin therapy: a prospective study. Pharmacogenomics J. 2020 Jun;20(3):494-504.

Sarwar N, Sandhu MS, Ricketts SL, Butterworth AS, Di Angelantonio E, Boekholdt SM, Ouwehand W, Watkins H, Samani NJ, Saleheen D, Lawlor D, Reilly MP, Hingorani AD, Talmud PJ, Danesh J. Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies. Lancet. 2010 May 8;375(9726):1634-9.

Sharp SJ, Poulaliou M, Thompson SG, White IR, Wood AM. A review of

published analyses of case-cohort studies and recommendations for future reporting. PLoS One. 2014;9(6).

Sim DS, Jeong MH. Differences in the Korea acute myocardial infarction registry compared with western registries. Korean Circ J. 2017;47(6):811–22.

Stephen Burgess SGT. Mendelian randomization methods for using genetic variants in causal estimation. 2015. 13–24 p.

Tada H, Nohara A, Kawashiri MA. Serum triglycerides and atherosclerotic cardiovascular disease: Insights from clinical and genetic studies. Nutrients. 2018;10(11).

Tchetgen EJT, Walter S, Vansteelandt S, Martinussen T, Glymour M. Instrumental variable estimation in a survival context. Epidemiology. 2015;26(3):402–10.

Teumer A. Common Methods for Performing Mendelian Randomization. Front Cardiovasc Med. 2018;5(May):1–7.

Thomas DC, Conti D V. Commentary: The concept of "Mendelian randomization." Int J Epidemiol. 2004;33(1):21–5.

Ueshima H, Sekikawa A, Miura K, Turin TC, Takashima N, Kita Y, et al. Cardiovascular disease and risk factors in Asia: A selected review. Circulation. 2008;118(25):2702–9.

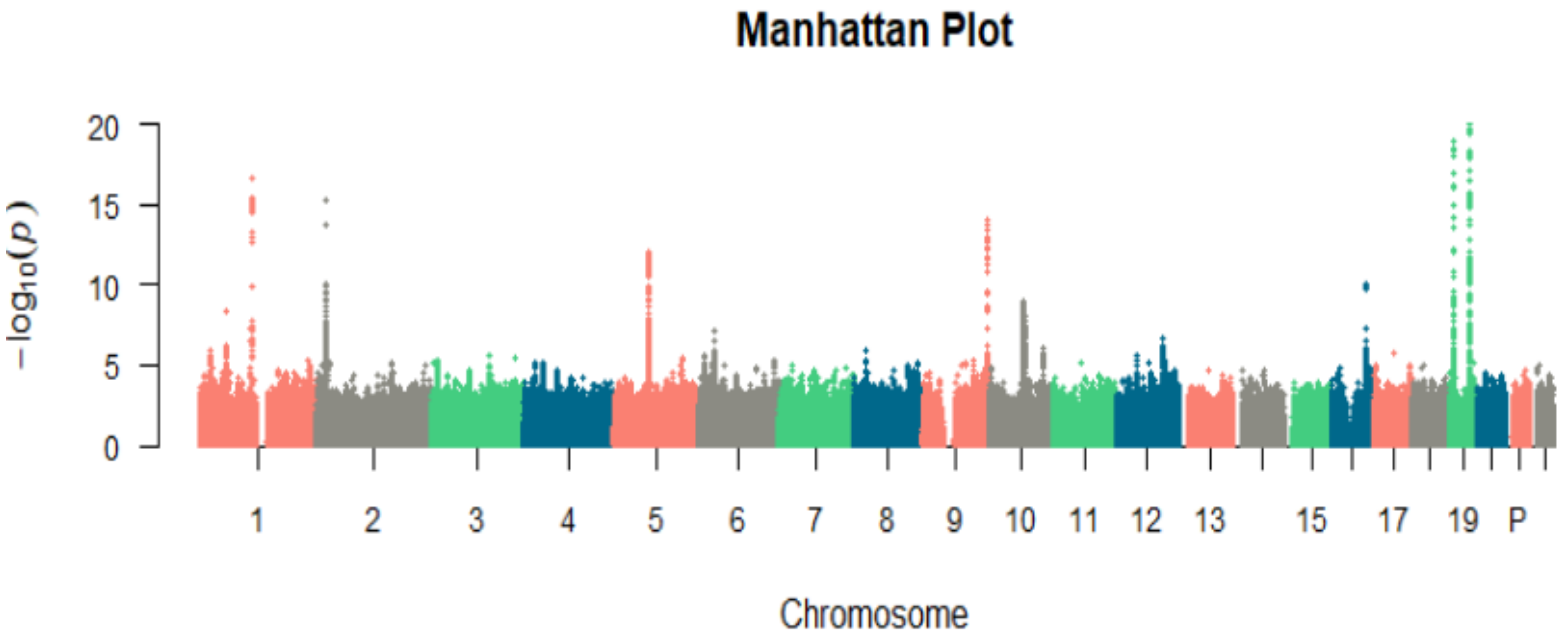
Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, Hindy G, Hólm H, Ding EL, Johnson T, Schunkert H, Samani NJ, Clarke R, Hopewell JC, Thompson JF, Li M, Thorleifsson G, Newton-Cheh C, Musunuru K, Pirruccello JP, Saleheen D, Chen L, Stewart A, Schillert A, Thorsteinsdottir U, Thorgeirsson G, Anand S, Engert JC, Morgan T, Spertus

J, Stoll M, Berger K, Martinelli N, Girelli D, McKeown PP, Patterson CC, Epstein SE, Devaney J, Burnett MS, Mooser V, Ripatti S, Surakka I, Nieminen MS, Sinisalo J, Lokki ML, Perola M, Havulinna A, de Faire U, Gigante B, Ingelsson E, Zeller T, Wild P, de Bakker PI, Klungel OH, Maitland-van der Zee AH, Peters BJ, de Boer A, Grobbee DE, Kamphuisen PW, Deneer VH, Elbers CC, Onland-Moret NC, Hofker MH, Wijmenga C, Verschuren WM, Boer JM, van der Schouw YT, Rasheed A, Frossard P, Demissie S, Willer C, Do R, Ordovas JM, Abecasis GR, Boehnke M, Mohlke KL, Daly MJ, Guiducci C, Burtt NP, Surti A, Gonzalez E, Purcell S, Gabriel S, Marrugat J, Peden J, Erdmann J, Diemert P, Willenborg C, König IR, Fischer M, Hengstenberg C, Ziegler A, Buysschaert I, Lambrechts D, Van de Werf F, Fox KA, El Mokhtari NE, Rubin D, Schrezenmeir J, Schreiber S, Schäfer A, Danesh J, Blankenberg S, Roberts R, McPherson R, Watkins H, Hall AS, Overvad K, Rimm E, Boerwinkle E, Tybjaerg-Hansen A, Cupples LA, Reilly MP, Melander O, Mannucci PM, Ardissino D, Siscovick D, Elosua R, Stefansson K, O'Donnell CJ, Salomaa V, Rader DJ, Peltonen L, Schwartz SM, Altshuler D, Kathiresan S. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. Lancet. 2012 Aug 11;380(9841):572-80.

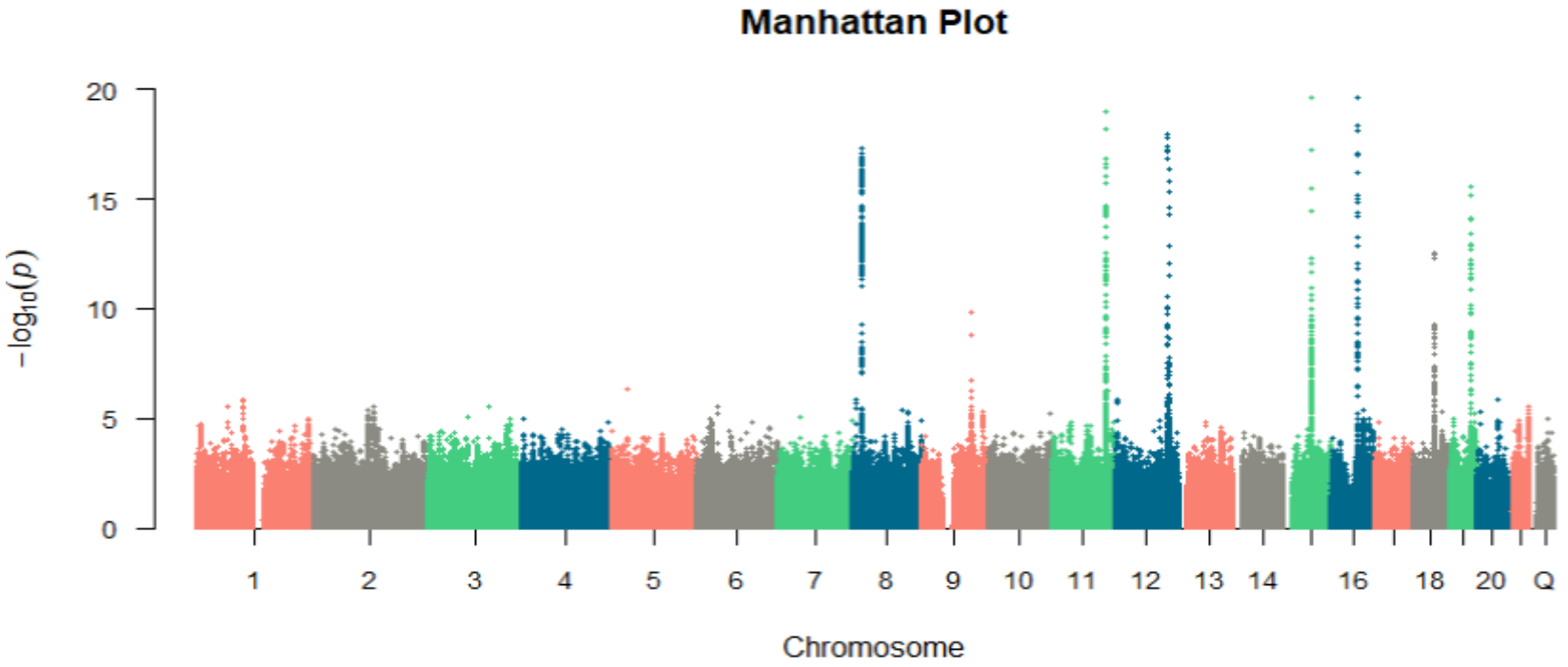
APPENDIX

Appendix 1. Comparison of analysis methods applied in this study

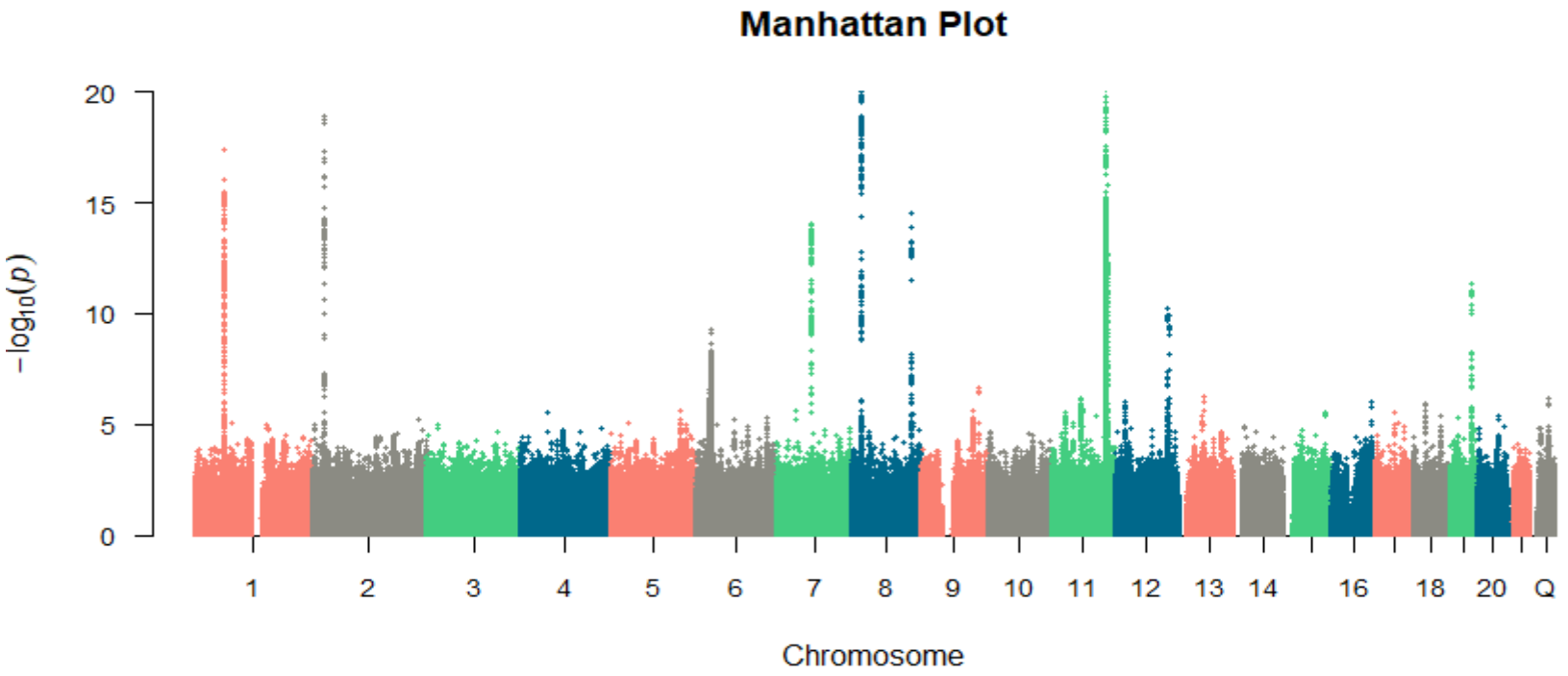
	Multivariable Cox regression of incident cases	One-sample Mendelian randomization	Two-sample Mendelian randomization
Strengths	Minimize the possibility of reverse causality Present the estimation of unadjusted and adjusted hazard ratio with CIs Enable to evaluate confounders thoroughly	Possible to investigate of subgroup analyses Incidence (hazard) rate of disease can be measured if implemented in a Cox regression framework Able to thoroughly evaluate individual-level confounding factors	Horizontal pleiotropy can be explored through use of different genetic instruments and MR-Egger Using two non-overlapping samples avoids 'Winner's curse' sample size and statistical power are improved adaptability and enhanced power to perform a variety of sensitivity analyses
Limitations	Be vulnerable to unmeasured or residual confounding Low statistical power selection bias with study sampling differential diagnosis	Weak instrument biases towards the confounded regression analysis result Low statistical power Winner's curse in which genetic variants identified in the same dataset as applied in MR analysis may bias estimates upwards Horizontal pleiotropy Potential for 'collider bias' due to study sampling	Weak instrument biases towards the null It is difficult to perform subgroup analyses with summary data from large GWAS consortia When using summary data from publicly available GWAS results, it is often impossible to confirm whether the confounders of the risk factor-outcome association is related to the genetic instrument



Appendix 2. Manhattan plots: Age, sex adjusted regression of LDL-C



Appendix 3. Manhattan plots: Age, sex adjusted regression of HDL-C



Appendix 4. Manhattan plots: Age, sex adjusted regression of TG

Korean Abstract

멘델리안 무작위분석법을 이용한 혈중 지질과 급성심근경색의 인과성 연구

김 근 희

연세대학교 보건대학원

배경 및 연구 목적

우리나라는 식생활과 생활 양식의 서구화 및 인구 고령화로 인해 급성심근경색 (Acute myocardial infarction, AMI)의 치명률, 발병률 및 유병률은 매년 꾸준히 증가하고 있다. 혈중 지질 및 지단백질 바이오 마커는 심근경색 및 뇌졸중과 같은 심혈관 질환의 주요 위험 요인이다. 그 중 저밀도 지질단백질 (Low-density lipoprotein cholesterol, LDL-C)은 심혈관 질환에 부정적인 영향을 미치는 것으로 보고되며, 고밀도 지단백질(High-density lipoprotein cholesterol, HDL-C)은 심혈관 질환에 긍정적인 영향을 미치는 것으로 보고되었다. 반면에 중성지방 (Triglyceride, TG)은 생물학적 매커니즘과 중성지방에 연관된 바이오 마커의 다면발현적인 영향으로 인해 급성심근경색의 연관성에 대하여 논란의 여지가 있다. 또한 현재까지 혈중 지질 지표 종류가 심근경색에 미치는 영향에 대한 연구는 대부분 서양인을 대상으로 한 연구였으며, 아시아인을 대상으로 혈중 지질 종류에 대한 급성심근경색의 위험을 MR으로 재현한 연구는 미흡한 실정이다. 따라서 본 연구에서는 유전자 도구 변수를 바탕으로 최신 멘델리안 무작위 분석법 (Mendelian randomization, MR)을 통해 혈중 지질 농도와 심근경색의 인과적 관계를 확인하고자 하였다.

방법

1994년부터 2013년까지 수집된 한국 암 예방 연구-Ⅱ (KCPS-Ⅱ; Korean Cancer Prevention Research-Ⅱ) 코호트 자료에서 자발적 동의 하에 유전정보, 혈중 지질, 질환 정보를 포함한 연구자료를 제공한 16,553명 중 case-cohort design을 위해 선정된 서브 코호트에 속한 대상자와 급성심근경색 발생 대상자를 포함하여 최종 5,197명이 본 연구에 포함되었다. 전장유전체상관성분석 (Genome-wide association study, GWAS)을 시행하여 노출 변수인 혈중 지질과 높은 연관성을 보이는 단일염기다형성(single nucleotide polymorphism, SNP)에 가중치를 부여하여 통합한 유전위험점수(weighted genetic risk score, WGRS)가 멘델리안 무작위 분석을 위한 도구변수로 사용되었다. 혈중 지질에 대한 심근경색 사이의 연관성을 다변량 콕스 회귀분석법과 one-sample 멘델리안 무작위 분석법 (Mendelian Randomization, MR)을 통해 확인하였다.

연구결과

GWAS 분석을 통해 선정된 각 혈중 지질들과 관련된 상호 독립적인 단일염기다형성(Single nucleotide polymorphisms, SNP) 은 유전위험점수로 가중치 계산하여 유전 도구변수로 선정되었으며, 분석에 사용된 유전위험점수는 각 혈중 지질들과의 강한 관련성을 보였다 (LDL-C F_{WGRS}=64.37; HDL-C F_{WGRS}=234.85; TG F_{WGRS}=301.81). 관찰 연구인 다변량 콕스 회귀분석에서 LDL-C와 TG는 confounding 변수들을 통제 했을 때, 급성심근경색 위험과 통계적으로 유의한 양의 관련성을 보였으며 HDL-C는 통계적으로 유의한 음의 관련성을 보였다. 반면, HDL-C, TG에 연관된 도구 변수와 급성심근경색 위험과의 연관성은 MR분석에서 유의하지 않은 결과를 보였으며($HR_{HDL-WGRS}$=0.998; 95% confidence interval, 0.911-1.093; $HR_{TG-WGRS}$=1.493; 95% confidence interval, 0.310-7.188), LDL-C는 MR 분석을 통해서도 일관된 유의한 양의 연관성 결과를 보여주어 두 요소간의 인과성을 보여주었다

($HR_{LDL-WGRS}$=1.046; 95% confidence interval, 1.019−1.078).

결론

혈중 지질과 AMI의 인과성 분석 결과, LDL-C는 AMI와 인과적인 관련성이 발견되었으며, 반면에 HDL-C, TG는 AMI에 유의한 인과적인 영향이 나타나지 않았다. 본 연구는 혈중 지질 농도와 급성심근경색의 위험의 인과적인 연관성이 한국인 코호트에서 재현되었다는 점에서 그 의의를 지닌다. 추후 한국인의 혈중 지질 농도와 급성심근경색에 유의한 단일염기다형성 (SNP; Single nucleotide polymorphisms) 발견 및 앞으로 대규모 자료를 통하여 two-sample MR 및 혈중 지질의 생물학적 매커니즘에 대한 연구가 계속 진행되어야 할 것이다.

핵심어: 멘델리안 무작위분석법, 전장유전체상관성분석, 혈중 지질, 급성심근경색, 유전위험점수