ABSTRACT

Background and Objectives: Associations between blood lipids and risk of ischemic heart disease (IHD) have been reported in observational studies. However, due to confounding and reverse causation, observational studies are influenced by bias, thus their results show inconsistency in the effects of lipid levels on IHD. In this study, we evaluate whether lipid levels have an effect on the risk of IHD in a Korean population.

Methods: A 2-sample Mendelian randomization (MR) study, using the genetic variants associated with lipid levels as the instrumental variables was performed. Genetic variants significantly associated with lipid concentrations were obtained from the Korean Genome and Epidemiology Study (n=35,000), and the same variants on IHD were obtained from the Korean Cancer Prevention Study-II (n=13,855). Inverse variance weighting (IVW), weighted median, and MR-Egger approaches were used to assess the causal association between lipid levels and IHD. Radial MR methods were applied to remove outliers subject to pleiotropic bias.

Results: Causal association between low-density lipoprotein-cholesterol (LDL-C) and IHD was observed in the IVW method (odds ratio, 1.013; 95% confidence interval, 1.007–1.109). However, high-density lipoprotein-cholesterol (HDL-C) and triglyceride (TG) did not show causal association with IHD. In the Radial MR analysis of the relationship between HDL-C, TG and IHD, outliers were detected. Interestingly, after removing the outliers, a causal association between TG and IHD was found.

Conclusions: High levels LDL-C and TG were causally associated with increased IHD risk in a Korean population, these results are potentially useful as evidence of a significant causal relationship.

Keywords: Mendelian randomization analysis; Lipid; Ischemic heart disease

INTRODUCTION

Serum lipid levels play a significant role in the development of atherosclerotic cardiovascular diseases.1 When there is an increased concentration of low-density lipoprotein-cholesterol (LDL-C) or a decreased level of high-density lipoprotein-cholesterol (HDL-C), there is a progressive development of atherosclerotic plaque, which can cause cardiovascular disease.2
In observational epidemiological studies, increased LDL-C levels and increased triglyceride levels and decreased HDL-C levels are associated with an increased risk for coronary artery disease (CAD). However, it may be difficult to disentangle causation from association, especially in the presence of confounding and reverse causation. Effect estimates from observational studies may be prone to measurement error or biases, which can lead to the attenuation of estimates.

Mendelian randomization (MR) provides an alternative way of dealing with the limitations of observational studies. Recent MR studies on lipid level biomarkers showed that elevated LDL-C and triglyceride (TG) are causal risk factors for coronary heart disease (CHD). Unlike previous observational studies, genetically raised HDL-C levels might not be protective against myocardial infarction. The available evidence for the causal relevance of lipid biomarker from randomized trials or MR studies is inconsistent. In addition, although multivariable MR analyses has suggested a causal effect of increased TG on CAD risk, it is still controversial whether elevated serum TG levels is a risk factor for ischemic heart disease (IHD).

Therefore, in this study, we examined whether lipid levels are causally related to IHD risk using a 2-sample MR approach with a Korean population.

**METHODS**

**Two-sample Mendelian randomization analysis**

MR analyses is a method commonly used in epidemiology to estimates the causal relationship between exposure and disease based on natural experiments assigned to high and low levels of exposure depending on individual genetic variation. To investigate the causal relationship between lipid concentrations and IHD, we performed 2-sample inverse variance weighted (IVW) MR analyses using the Korean genome wide association study (GWAS)'s results data from their Korean Genome and Epidemiology Study (KoGES) as well as the results data from the Korean Cancer Prevention Study-II (KCPS-II) Biobank.

The exposure variables of interest were the genetic variants that are associated with 3 serum lipid fraction in the Korean population. From KoGES, genetic information related to lipid levels was obtained. The KoGES study included 223,000 people and collected data between 2004 and 2013. The Korean chip project was genome-based chronic disease project that collected information on 830,000 specific genetic variations in the Korean population. Among the KoGES population, 35,000 participants who took part in the Korean-chip data study were recruited. After quality control excluded some participants (Hardy-Weinberg equilibrium <10^-4, minor allele frequency<1%), a total of 33,598 subjects were selected for further analysis using the K-chip (http://nih.go.kr/menu.es?mid=a50303010100). Before conducting GWAS analyses, we performed imputation analysis using the 1000 Genome (1000G) Phase 3 dataset of East Asian populations (Japanese in Tokyo and Chinese in Beijing) as a reference panel. Then, linear regression with adjustment for age and sex was performed to identify the associations of genetic variants with serum lipid level. Statistical analysis was carried out using PLINK version 1.07 (https://zzz.bwh.harvard.edu/plink/). Based on genome-wide significance (p<10^-8), we found 2,029 single nucleotide polymorphisms (SNPs) associated with HDL-C, 1,207 SNPs with LDL-C, and 3,757 SNPs with TG. Then, a clumping test (https://mrcieu.github.io/TwoSampleMR), which estimates the linkage disequilibrium between the SNPs selected using 1000G data, was performed to select instrumental
variables. Finally, 29, 20, and 20 independent SNPs of HDL-C, LDL-C, and TG ($r^2>0.001$) were chosen as the instrumental variables.

The outcome of interest was IHD when looking at the KCPS-II Biobank. The number of participants in the cohort was 156,701 (94,840 men and 61,861 women), all these received medical examinations at one of 18 health promotion centers in Seoul and Gyeonggi-do Province between 2004 and 2013. The published Biobank profile papers provide detailed information. We define the IHD events we are interested in as the first nonfatal occurrence of IHD. We identified these events from insurance claims submitted to the National Health Insurance System (NHIS) which is the single provider of universal healthcare coverage in Korea. We ascertained nonfatal IHD events according to the International Classification of Diseases 10th revision (I20–I25) from the health insurance claims data of the NHIS. We conducted a validation study in collaboration with the NHIS. Participants provided written permission for the use of their personal information to give our study 673 IHD events between 1994 and 2007, each case was confirmed with individual hospital medical records and 73% of the myocardial infarction cases were validated.

Genotyping for outcome was carried out with K-chip (http://nih.go.kr/menu.es?mid=a50303010100). The quality control procedure for SNPs was carried out using the same procedure as selection for the aforementioned test. GWAS analyses were conducted using the PLINK 1.07 program (http://zzz.bwh.harvard.edu/plink/).

Assumption of 2-sample Mendelian randomization

Two-sample MR analyses were performed to investigate the existence of a causal relationship between lipid levels and IHD.

MR necessitates that genetic instruments are related with a modifiable exposure of interest (assumption 1), and any relationship between the instruments and outcome is mediated by the exposure (assumption 2). The instrumental variable only affects the outcome through its effect on risk factors (assumption 3). An SNP that violates these assumptions is referred to as an invalid instrumental variable and its consideration in MR analyses may bias the results. In the absence of reliable information on the ratio and direction of functional pathways and multifaceted genetic variants, the consistency of the estimated direction of the effect was checked using additional MR methods, including MR-Egger weighted medians. The MR-Egger method provides a less biased effect estimate by assuming significant heterogeneity of directional pleiotropy and measurement error, this is unlike the IVW method which does not assume the intercept term of the model. The results are related and require a larger sample size for the same basal variance in exposure. The IVW and MR-Egger methods further assume that the pleiotropic effects of genetic variants are independent of the assumption of direct effect, and independent of the association with exposure known as instrument strength. In the case of violation, the weighted median method can provide consistent effect estimates even when up to 50% of the genetic variants do not fit into instrumental strengths that depend on the assumption of direct effect.

Sensitivity analyses

The MR-Radial method was used to identify heterogeneity (p value=0.05/number of SNPs). The directional pleiotropy method occurs when the net effect of the horizontal pleiotropy method over all SNPs is not zero, which would introduce bias into the IVW method estimation. When heterogeneity was detected, we removed outliers and repeated both the analysis and the IVW radial regression.
We used the TwoSampleMR package\textsuperscript{20} and the RadialMR package\textsuperscript{21} to perform IVW, MR-Egger regression, and weighted median analyses in R (version 3.4.3; R Foundation, Vienna, Austria). We ran the analysis using code provided at https://mrcieu.github.io/TwoSampleMR/ and https://github.com/WSpiller/RadialMR/.

**Ethics statement**
The Severance Medical Ethics Committee approved this study (No. 4-2011-0277).

**RESULTS**

**Causal effect of lipids on ischemic heart disease**

We screened out 69 lipid-specific instrumental variables (29 for HDL-C, 20 for LDL-C and 20 for TG) when evaluating their associations with IHD (Supplementary Tables 1-3 and Figure 1).

The Cochran's Q test showed that there is little evidence of instrumental heterogeneity for LDL-C ($p$ value=0.378) or TG ($p$ value=0.242). However, instrumental heterogeneity was observed for HDL-C at the marginal significance level of 0.05 ($p$ value=0.032) (Figure 2).

Motivated by this, we employed the IVW method for LDL-C and TG, and applied the MR-Egger method for HDL-C when estimating their causal effects on IHD. Among the 3 lipids, we identified that only LDL-C was associated with IHD (Table 1). Specifically, the estimated odds ratio (OR) per 1 mg/dL increase of LDL-C on IHD was 1.013 (95% confidence interval [CI], 1.007–1.019; $p$ value=9.44E-06) and was statistically significant. The OR per 1 mg/dL increase of HDL-C and TG for IHD was 0.995 (95% CI, 0.98–1.01; $p$ value=0.552) and 1.002 (95% CI, 0.999–1.004; $p$ value=0.065) (Table 1), respectively. In the case of LDL-C, MR-Egger regression analysis showed the presence of pleiotropy ($p$ value=0.008). However, the possibility of pleiotropy was not found to be significant in further intercept analysis.

**Sensitivity analyses to validate estimated causal effects**

We subsequently performed sensitivity analyses to validate the causal association observed above. Using the weighted median method, the estimated OR per 1 mg/dL increase of HDL-C

---

**Figure 1.** Scatter plots to visualize causal effect of lipid levels on IHD risk. (A) HDL-C; (B) LDL-C; (C) TG.

SNP = single nucleotide polymorphism; IHD = ischemic heart disease; HDL-C = high-density lipoprotein-cholesterol; LDL-C = low-density lipoprotein-cholesterol; TG = triglyceride; MR = Mendelian randomization; IVW = inverse variance weighting.
and TG on IHD was 1.001 (95% CI, 0.982–1.02; p value=0.879) and 1.002 (95% CI, 0.999–1.004; p value=0.085); the LDL-C has a significant association (OR, 1.011; 95% CI, 1.003–1.02, p value=2.30E-03) (Table 1). However, the IVW Radial MR results delineated that the 2 instrumental variables of HDL-C and TG that carry large effect sizes on IHD could potentially be outliers (i.e., HDL-C on rs2876971 in gene \textit{UBE2L3} and rs769446 in gene \textit{APOE}; TG on rs438811 in gene \textit{APOC1} and rs9436224 in gene \textit{DOCK7}) (Figure 2). Outliers did not exert any substantial influence on the estimated causal effect of HDL-C on IHD. However, after removal, the estimate of TG’s effect on IHD was 0.002 (standard error=0.001, p value=0.006) (Table 2).

**Figure 2.** Radial plots of SNP-IHD associations (beta*sqrt(Wj)) versus SNP-lipid level associations (sqrt(Wj)), with the IVW slope shown as a solid sky-blue line. Outliers are highlighted in orange. (A) HDL-C; (B) LDL-C; (C) TG.

**Table 1.** Mendelian randomization for the effects of lipid levels on ischemic heart disease risk

| Variables | Method                  | OR (95% CI)   | p value | Number of SNPs |
|-----------|-------------------------|---------------|---------|----------------|
| HDL-C     | Inverse variance weighted | 0.995 (0.98–1.010) | 0.552  | 29             |
|           | Weighted median         | 1.001 (0.982–1.020) | 0.879  | 29             |
|           | MR-Egger regression     | 1.007 (0.980–1.034) | 0.667  | 29             |
| LDL-C     | Inverse variance weighted | 1.013 (1.007–1.019) | 9.44E-6 | 20             |
|           | Weighted median         | 1.011 (1.003–1.020) | 2.30E-3| 20             |
|           | MR-Egger regression     | 1.014 (1.005–1.024) | 7.99E-3| 20             |
| TG        | Inverse variance weighted | 1.002 (0.999–1.004) | 0.065  | 20             |
|           | Weighted median         | 1.002 (0.999–1.004) | 0.085  | 20             |
|           | MR-Egger regression     | 1.002 (0.997–1.006) | 0.329  | 20             |

OR = odds ratio; CI = confidence interval; SNP = single nucleotide polymorphism; HDL-C = high-density lipoprotein-cholesterol; LDL-C = low-density lipoprotein-cholesterol; TG = triglyceride; MR = Mendelian randomization.

**Table 2.** Inverse weighted variance radial regression results for lipid levels on ischemic heart disease risk

| Variables | Method                  | Beta (SE)   | p value | Number of SNPs |
|-----------|-------------------------|-------------|---------|----------------|
| HDL-C     | Inverse variance weighted | −0.004 (0.007) | 0.533  | 27             |
| TG        | Inverse variance weighted | 0.002 (0.001) | 0.006  | 18             |

SE = standard error; SNP = single nucleotide polymorphism; HDL-C = high-density lipoprotein-cholesterol; TG = triglyceride.

*After removing outliers, remaining SNPs were analyzed.
DISCUSSION

This study assessed the causal relationship of LDL-C, HDL-C and TG on IHD using a 2-sample MR method. All these lipids are traditionally well known as risk factors for IHD. The 2-sample used for the MR approach included KoGES for exposure data and KCPS-II for outcome data.

The MR results showed that LDL-C and TG were causally related to IHD, but HDL-C was not causally related to IHD, which is consistent with results of previous studies.

Recent MR studies of IHD biomarkers show that LDL-C has a causal relationship with IHD, while the causality of association between HDL-C and IHD is not clear. These previous results are consistent with the results of this study. In the case of TG, however, it is still controversial whether elevated serum TG levels are a risk factor for IHD. In our study, the relationship between TG and IHD was heterogeneous. In other words, the 2 SNPs were determined to be outliers in the radial plot. The outliers associated with TG in this study were rs438811 in gene APOC1 and rs9436224 in gene DOCK7. The IVW method results analyzed with these 2 outliers showed no significant association with an OR,1.002 and a p value=0.065. However, after removing the 2 outliers using IVW radial regression, results showed a significant causal association with an OR,1.002 and a p value=0.006.

This study demonstrates the potential of TG as a causal factor for IHD. In a recent study, it was found that treatments that lower LDL-C or TG levels should lead to reductions in CHD risk proportional to their reduction in ApoB. However, cardiovascular outcome studies with TG-lowering agents still produce inconsistent results, meaning that there is no convincing evidence available that suggests lowering TG by any approach can reduce mortality. Therefore, causality studies for TG and IHD should be repeated in the future.

The evidence that LDL-C and TG are causally related to IHD is based on genetically determined LDL-C and TG. Genetically determined LDL-C and TG are characterized by being free of confounding variables, this problem has always affected the ability of observational studies to uncover true relevance. This does not mean that only genetically increased LDL-C and TG are the cause of IHD. This means that only genetic factors are used in the methodology for causality, and so not only the genetic increases LDL-C or TG but also the increases due to other life style risk factors can cause IHD.

There are several health implications for biomarkers’ causality in certain diseases. First, causal biomarkers can reduce the incidence of related diseases if biomarkers can be reduced through drug or non-drug therapies. Second, causal biomarkers can be candidates for drug development. Therefore, HDL-C, which did not show causality, cannot be a candidate for drug development but can be used as an IHD predictor for finding high risk groups.

To our knowledge, this is the first study to examine the causality between lipid levels and IHD risk in a Korean population using 2-sample MR. In addition, all other previously published MR studies have been conducted primarily in populations of European descent. Thus, all instrument variables have been extracted from GWASs of European descent. In this study, we used a restricted set of SNPs for lipid levels within a Korean population (KoGES data) using individual-level data to avoid potential population substructure. We also verified the IHD events using the NHIS database. Koreans have good access to medical services as almost all citizens are enrolled in national health insurance.
This study has several limitations. First, this study includes possible measurement errors: according to previous research, clinical data from the health promotion centers were one-time measurements of serum lipids and other variables. Cholesterol levels are likely to fluctuate over time, failure to consider these changes may have been a factor that obscures relevance. Second, due to using KCPS-II data, which is a non-community-based sample, any potential risk estimator cannot be validated. Third, selected lipid instruments may not be useful in other ethnic groups. Therefore, our study findings may not be generalized to other ethnicities. Fourth, this study did not examine all IHD events. The IHD events in this study were used only in the 'non-fatal IHD events' validated in hospital records, and therefore fatal coronary events cannot be explained. Finally, MR is a popular method for estimating causal effects of risk factors in observational studies, but it is difficult to completely exclude the effect of relevant unobserved potential confounders.

In conclusion, the study provides evidence that LDL-C and TG are causal biomarkers for the increased risk of developing IHD. Further studies that replicate this causal association among other ethnicities with larger samples and accurate outcome data are needed.

ACKNOWLEDGEMENTS

The authors thank the staff of the Korean National Insurance Service. This study was conducted with bioresources from National Biobank of Korea, the Center for Disease Control and Prevention, Republic of Korea (KBN-2016-013).

SUPPLEMENTARY MATERIALS

Supplementary Table 1
Associations of individual genetic instruments for high density lipoprotein-cholesterol levels with ischemic heart disease risk

Click here to view

Supplementary Table 2
Associations of individual genetic instruments for low density lipoprotein-cholesterol levels with ischemic heart disease risk

Click here to view

Supplementary Table 3
Associations of individual genetic instruments for triglyceride levels with ischemic heart disease risk

Click here to view
REFERENCES

1. Tarchalski J, Guzik P, Wysocki H. Correlation between the extent of coronary atherosclerosis and lipid profile. *Mol Cell Biochem* 2003;246:25-30.

2. Ference BA, Ginsberg HN, Graham I, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *Eur Heart J* 2017;38:2459-72.

3. Di Angelantonio E, Sarwar N, Perry P, et al. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA* 2009;302:1993-2000.

4. McBride P. Triglycerides and risk for coronary artery disease. *Curr Atheroscler Rep* 2008;10:386-90.

5. Sheehan NA, Didelez V, Burton PR, Tobin MD. Mendelian randomisation and causal inference in observational epidemiology. *PLoS Med* 2008;5:e117.

6. Linsel-Nitschke P, Götz A, Erdmann J, et al. Lifelong reduction of LDL-cholesterol related to a common variant in the LDL-receptor gene decreases the risk of coronary artery disease—A Mendelian Randomisation study. *PLoS One* 2008;3:e2986.

7. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a Mendelian randomisation study. *Lancet* 2012;380:572-80.

8. Varbo A, Bena M, Tybjærg-Hansen A, Jørgensen AB, Frikke-Schmidt R, Nordestgaard BG. Remnant cholesterol as a causal risk factor for ischemic heart disease. *J Am Coll Cardiol* 2013;61:427-36.

9. Burgess S, Freitag DF, Khan H, Gorman DN, Thompson SG. Using multivariable Mendelian randomization to disentangle the causal effects of lipid fractions. *PLoS One* 2014;9:e108891.

10. Shah S, Casas JP, Drenos F, et al. Causal relevance of blood lipid fractions in the development of carotid atherosclerosis: Mendelian randomization analysis. *Circ Cardiovasc Genet* 2013;6:63-72.

11. Holmes MV, Asselbergs FW, Palmer TM, et al. Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J* 2015;36:539-50.

12. Elousa R, Sayols-Baixeras S. The genetics of ischemic heart disease: from current knowledge to clinical implications. *Rev Esp Cardiol (Engl Ed)* 2017;70:794-62.

13. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. *Int J Epidemiol* 2016;45:908-15.

14. Kim Y, Han BGKoGES group. Cohort profile: The Korean Genome and Epidemiology Study (KoGES) Consortium. *Int J Epidemiol* 2017;46:1350.

15. Lee YH, Emberson J, Jung KJ, et al. Cohort profile: The Korean Cancer Prevention Study-II (KCPS-II) Biobank. *Int J Epidemiol* 2018;47:385-386f.

16. 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 (1). *Korean Circ J* 2012;42:10-5.

17. Didelez Y, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat Med* 2007;26:309-14.

18. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 2016;40:304-44.

19. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44:512-25.

https://e-kcj.org

https://doi.org/10.4070/kcj.2020.0131
20. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *elife* 2018;7:e34408.

21. Bowden J, Spiller W, Del Greco M F, et al. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. *Int J Epidemiol* 2018;47:1264-78.

22. Navar AM. The evolving story of triglycerides and coronary heart disease risk. *JAMA* 2019;321:347-9.

23. Reiner Z. Hypertriglyceridaemia and risk of coronary artery disease. *Nat Rev Cardiol* 2017;14:401-11.

24. Bennett DA, Holmes MV. Mendelian randomisation in cardiovascular research: an introduction for clinicians. *Heart* 2017;103:1400-7.