

Deciphering correlation and causation in risk factors for heart disease with Mendelian randomization

Aryamann Singh¹, Parsa Akbari²

¹ Thomas Jefferson High School for Science and Technology, Alexandria, Virginia

² Downing College, Cambridge, United Kingdom, CB2 1DQ

SUMMARY

In this analysis, we studied the use of Mendelian randomization to identify the risk factors of coronary artery disease (CAD), a major cause of cardiovascular disease. Identifying risk factors of CAD are critical to understanding and managing the disease. Our analysis combined results from 28 genetic analyses from 12 unique studies. For each genetic variation, we obtained the variant ID, chromosome, base-pair position, reference and alternative alleles of the genetic variation, and estimated effect and *p*-value of the genetic variation on the outcome. We hypothesized that traits which are correlated with CAD outcomes will be causally associated with CAD risk in a genetic Mendelian randomization analysis. Our analysis showed that several traits such as blood pressure readings (systolic, OR 0.51 (95% CI: 0.34-0.69), *p*-value = 5.4×10^{-9}) and (diastolic, OR 0.56 (95% CI: 0.41-0.71), *p*-value = 7.6×10^{-14}), LDL cholesterol levels (OR 0.54 (95% CI: 0.47-0.60), *p*-value = 4.4×10^{-56}), and BMI (OR 0.41 (95% I: 0.35-0.48), *p*-value = 6.30×10^{-33}) were significant risk factors for CAD. C-reactive protein (OR -0.09 (95% CI: -0.18–0.00), *p*-value = 0.05) was a protective risk factor of CAD due to its negative odds ratio. In contrast, eosinophil count (OR -0.007 (95% CI: -0.06-0.04), *p*-value = 0.79) had no statistically significant association. Blood cells had weak associations with CAD, and uric acid's role as a causal or reversible risk factor of CAD was inconclusive, requiring further study.

INTRODUCTION

Coronary artery disease (CAD), a condition caused by the buildup of plaque in the wall of the arteries that supply blood to the heart, is the leading cause of death for men and women in the United States (1). Nearly 610,000 Americans and 17.8 million people worldwide die annually from heart disease (2,3). Projections indicate that by 2035, nearly 45% of the adult American population will have some type of cardiovascular disease – like CAD (4).

Cardiovascular disease is the leading cause of death in the United States, however cardiovascular disease is preventable (4). Prevention methods are primarily focused on controlling risk factors by maintaining a healthy diet lacking saturated fats and trans fats, getting physical exercise, and avoiding smoking (5). Controlling other risk factors such as tobacco use, obesity, smoking, and raised blood pressure can largely prevent cardiovascular disease (3). This emphasizes the need for identification of further causal risk factors of cardiovascular

disease. However, identification of risk factors is traditionally done with observational studies, which are limited due to effects between confounding variables and environmental factors (6). For this reason, many observational studies are not able to distinguish between correlative and causative relationships. Mendelian randomization offers a solution to the difficulty of causal inference in observational studies by leveraging genetic information to inform statistical tests of causality (6,7).

Mendelian randomization (MR) is a statistical method that can identify factors which increase the risk or protect against a disease (6,7). When observing patients, it is difficult to differentiate between a correlative or a causative relationship between protective or risk factors and disease outcomes. MR can address this by using an instrumental variable analysis, which relies on genetic variation (7). Genetic variation is inherited at birth and cannot be affected by disease status or the environment (6). MR utilizes genetic variations as instrumental variables to identify causative effects. In this analysis, we studied the use of MR to identify the risk factors of CAD. Our hypothesis was that MR can identify known risk factors for CAD, and also that there would be causal effects from suggested risk factors such as eosinophil, lymphocyte, basophil, and platelet cell count on CAD.

The analytical advantage of MR is the ability to distinguish correlation from causation between risk factors and disease outcomes (6,8). The MR protocol begins by identifying key genetic variations associated with a risk factor and the effect of these genetic variations on a disease outcome risk (6). In our analysis, we collected the latest results from a genome-wide association study of 16 human phenotypes with purported mechanistic associations with CAD risk. For example, our MR study showed that an increase in low density lipoprotein (LDL) cholesterol levels will also increase the risk of CAD. The MR analysis identified that genetic variations which increase LDL cholesterol levels will also on average increase the risk of CAD. This MR analysis relies first on a study of LDL cholesterol levels that identified key genetic variations influencing the LDL cholesterol levels phenotype. We further query and extract the effect of these genetic variations on CAD risk utilizing a large study of the genetics of CAD patients. Using this combined dataset, the MR protocol is applied to derive a causal effect between LDL cholesterol levels and CAD risk.

Interesting results were obtained from the study. Several traits such as blood pressure readings, LDL cholesterol levels, tobacco smoking, body mass index (BMI), uric acid, C-reactive protein levels (a biomarker of inflammation), and type 2 diabetes were significant risk factors for CAD. Blood cell counts such as eosinophils, basophils, neutrophils, monocytes, lymphocytes, platelets, and red blood cells had

weak associations with CAD. We were unable to identify uric acid as a causal or reversible risk factor of CAD, so further study is required. These results show that the MR methodology to identify causal risk factors recapitulates known traits that are causative of CAD, and that blood cell counts and uric acid levels may be causative risk factors for CAD. Identification of further risk factors for CAD will enable better patient diagnostics and potential therapeutics to manage CAD risk.

RESULTS

We performed a confirmatory study showing that Mendelian randomization methodology was able to identify known causal risk factors for CAD (Figure 1). The causal association was therefore more resistant to confounding variables because the Mendelian randomization protocol relies on genetic variations as instrumental variables, which are randomized at birth and cannot be affected by reverse causality. The results demonstrated the change in CAD risk for one standard deviation (SD) in the exposure. For example, one standard deviation of increased systolic blood pressure results in 0.51 (95% CI: 0.34-0.69, p -value: 5.4×10^{-9}) odds ratio higher CAD risk.

Traits which Increase the Risk of CAD

Diastolic blood pressure and cholesterol are strongly associated with increased CAD risk

Diastolic blood pressure is a measure of arterial pressure when the heart is resting in between beats (9). High blood pressure increases risk of CAD as this adds force to the arterial wall, leading to damage and contributing to plaque formation (10). Our association analysis identified a highly significant association between diastolic blood pressure and

CAD risk, with a causal effect of 0.56 higher odds ratio of CAD for one standard deviation increased diastolic blood pressure (95% CI: 0.41-0.71, p -value: 7.44×10^{-14}). Furthermore, we identified a strong association between higher cholesterol (LDL) and risk of CAD, with a causal effect of 0.54 higher odds ratio of CAD for one standard deviation increased LDL cholesterol level (95% CI: 0.47-0.6, p -value: 4.44×10^{-56}). The buildup of such cholesterol in the walls of the arteries leads to the increased chances of CAD.

Uric acid promotes the risk of coronary artery disease

Uric acid, a chemical waste product present in blood, is created when the body breaks down purines (11). High levels of uric acid can lead to systemic inflammation and preclinical atherosclerosis, which is associated with future cardiovascular disease risk (12). Our study identified a highly significant association between uric acid and CAD risk, with a causal effect of 0.13 higher odds ratio of CAD for one standard deviation increased uric acid levels (95% CI: 0.06-0.2, p -value: 4.35×10^{-4}) (Figure 1).

Current tobacco smoking increases risk of coronary artery disease

Tobacco smoke contains harmful chemicals, such as nicotine, hydrogen cyanide, arsenic, and ammonia, which can inflame and swell cells that line the interior surface of blood vessels, leading to their contraction (13). Such infection can lead to cardiovascular complications, primarily CAD. This condition occurs when the arteries that carry the blood to the heart muscle are narrowed by the buildup of plaque, and with the tobacco chemicals causing the blood to thicken and form clots inside the veins and arteries, the blockage can further lead to a heart attack (13). Tobacco smoking increases the

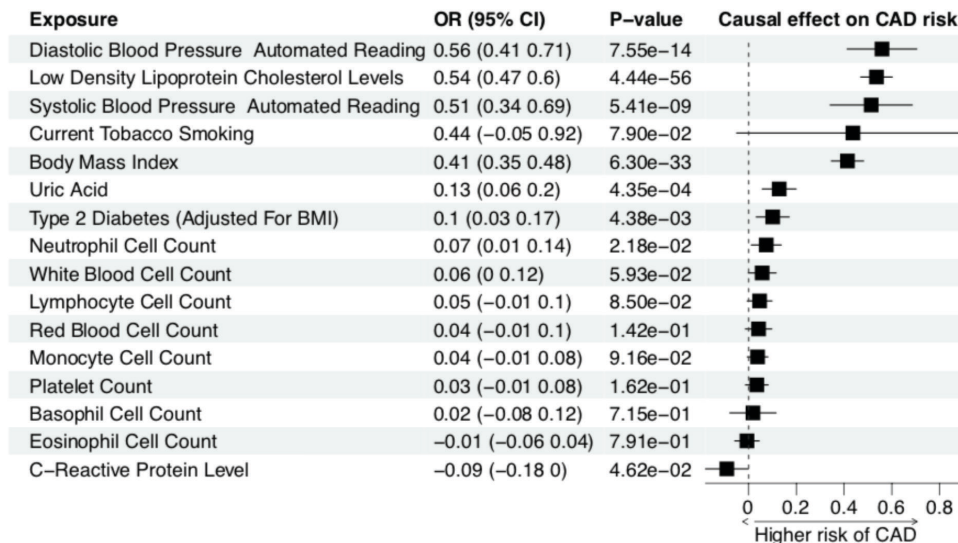


Figure 1: Forest plot representing the change in odds ratio of coronary artery disease for one standard deviation unit increase in exposures. A forest plot of 12 distinct exposures and their individual association tests with coronary artery disease risk as studied by Mendelian randomization. Most exposures increased the risks of coronary artery disease: blood pressure readings systolic (OR 0.51, 95% CI: 0.34-0.69, p -value = 5.4×10^{-9}) and diastolic (OR 0.56, 95% CI: 0.41-0.71, p -value = 7.6×10^{-14}), LDL cholesterol levels (OR 0.54, 95% CI: 0.47-0.60, p -value = 4.4×10^{-56}), BMI (OR 0.41, 95% CI: 0.35-0.48, p -value = 6.30×10^{-33}), uric acid (OR 0.13, 95% CI: 0.06-0.2, p -value = 4.35×10^{-4}), and type 2 diabetes (OR 0.1, 95% CI: 0.03-0.17, p -value = 4.38×10^{-3}). Only C-reactive protein (OR -0.09 (95% CI: -0.18-0.00), p -value = 0.05) levels protected against coronary artery disease. Blood cell counts had low associations with the disease but increased its risks, with an exception to eosinophil since it had a 0 change in odds ratio. Genetic variations for risk factor were studied in an independent analysis with varying sample size as listed in Table 1.

risks of CAD, with a causal effect of 0.44 higher odds ratio of CAD for one standard deviation increased current tobacco smoking levels (95% CI: -0.05-0.92, p -value: 7.9×10^{-2}) (**Figure 1**).

BMI and type 2 diabetes are associated with coronary artery disease

BMI is the measure of body fat calculated from height and weight. A high BMI falls within the obesity range and is associated with several complications, such as hypertension, dyslipidemia, diabetes mellitus, metabolic syndrome, and cardiovascular diseases (14). Based on our results, BMI was significantly associated with higher risks of CAD, with a causal effect of 0.41 higher odds ratio of CAD for one standard deviation increased BMI levels (95% CI: 0.35-0.48, p -value: 6.3×10^{-33}). Weight loss and decreasing BMI, however, is associated with the alleviation of preexisting cardiovascular risk factors (15). Additionally, type 2 diabetes, the inability of cells to respond to insulin properly, can create high blood glucose and damage blood vessels and nerves, making the heart muscle stiffer (16). Such damage can result in heart diseases such as CAD. Type 2 diabetes increased the risks of CAD, with a causal effect of 0.1 higher odds ratio of CAD for one standard deviation increased type 2 diabetes levels (95% CI: 0.03-0.17, p -value: 4.38×10^{-3}) (**Figure 1**).

Traits which are Inconclusive or Decrease the Risk of CAD

CRP protects against the risk of coronary artery disease

C-Reactive Protein (CRP) is identified as a protein produced by the liver when white blood cells fight against inflammatory diseases and infection (17). Higher CRP levels indicate present inflammation in the body, which results in white blood cells fighting against bacterial agents created by the invading inflammatory disease, like CAD (18). Our study identified an association between CRP and CAD risk, with a causal effect of -0.09 higher odds ratio of CAD for one standard deviation increased CRP levels (95% CI: -0.18-0, p -value: 4.62×10^{-2}) (**Figure 1**). The negative odds ratio value signifies CRP as a protective factor of CAD.

Blood cell counts have weak relationships with coronary artery disease

The blood cell counts in **Figure 1** (neutrophil, white blood, lymphocyte, red blood, monocyte, platelet, basophil, and eosinophil) are all weakly associated with CAD. Limited conclusions were made on the eosinophil, basophil, platelet, monocyte, and red blood cell counts due to the proximity of their respective odds ratios to 0. On the other hand, neutrophil, lymphocyte, and white blood cell counts may play roles in the development of CAD. The neutrophil-to-lymphocyte cell count ratio can act as a marker of inflammation that is directly associated with CAD (19). Moreover, high counts of white blood cells can indicate inflammation and developing CAD (20). While these three blood cell counts indicate increased risks of CAD, limited conclusions were made due to low associations (**Figure 1**).

DISCUSSION

The causal risk factors, systolic and diastolic blood pressure readings, are observed to strongly increase the risk of CAD. This is well supported by a number of observational

studies showing high blood pressure to be a major risk factor for coronary heart disease, indicating that low systolic and diastolic blood pressure readings are associated with a low risk for developing coronary heart disease (**Figure 1**) (10). Moreover, LDL cholesterol levels have been extensively shown to also increase the risk of CAD. The majority of the body's cholesterol is LDL, and unhealthy lifestyles can produce excess amounts, leading to harmful effects such as the buildup of cholesterol in the walls of the arteries (21). Studies have investigated the relationships between LDL and high density lipoprotein (HDL) cholesterol levels and mortality among people 85 years and older (22). High LDL cholesterol and low HDL cholesterol concentrations were both linked with an increased mortality risk of infection, concluding that high LDL cholesterol levels can lead to the increased risks of CAD (22).

Additionally, current tobacco smoking as an exposure increases the risk of CAD (**Figure 1**). Studies have found that smoking is a major risk factor for CAD as one particular experiment aimed to determine the relationship between smoking status and the risk of developing CAD (23). Findings specified that for smokers under the age of 50, the risk of developing CAD is ten times greater than for nonsmokers of the same age (23). In addition, smoking was found to double the risk of mortality from ischemic heart disease (also known as coronary artery disease), compared with a lifetime of not smoking (23). These studies supported the results obtained from the MR experiment (**Figure 1**).

Moreover, BMI and type 2 diabetes are observed to increase the risk of CAD. One particular study performed a retrospective study on a large group of patients who underwent cardiac catheterization and were experiencing chest pain, finding that a BMI over 30 – which indicates obesity – is a risk factor for early development of CAD (14). Likewise, type 2 diabetes, an exposure that is also promoted by obesity, is associated with cardiovascular complications (24). Specifically, the Centers for Disease Control (CDC) finds that high blood sugar can damage the heart's blood vessels and nerves (25). Another such study specifies that diabetes mellitus is associated with higher risks of cardiovascular diseases, such as CAD (25).

Regarding the results of uric acid, several studies have debated the role of uric acid as a causal risk factor for CAD. Even though the plot indicates that uric acid promotes the risk of CAD, one study's results reveal that increases in uric acid may protect against the progression of CAD (26). Several other studies proved the opposite, indicating that increased levels of serum uric acid lead to the development of CAD, but emphasized the need of clinical trials to confirm the results (27). Uric acid may act as a pathogen and enhance the harmful effects of cardiovascular risk factors on the vascular tissue and myocardium (28). However, a review determined that even though there is a considerable amount of evidence that associates uric acid with CAD, there is no evidence that uric acid acts a causal or reversible relationship to vascular diseases (29). This debate makes it unable to determine uric acid's role as a risk factor.

CRP level is the other exposure that seems to protect against the risk of CAD (**Figure 1**). One particular study's findings indicated that CRP may increase in cardiovascular disease to respond to infectious agents generating inflammatory reactions in the coronary vessels (18). In terms

of CAD, another study plotted the risk ratios for CAD by CRP levels, obtaining a log-linear shape that indicates that CRP levels have associations with the risk of coronary heart disease (17). The MR experiment in Figure 1 denoted the protection of CAD from CRP levels, which is supported by the studies examined, since CRP levels do, in fact, increase to protect the immune system from bacterial agents originating from cardiovascular diseases.

We investigated several blood cell counts – neutrophil, white blood cell, lymphocyte, red blood cell, monocyte, platelet, basophil, and eosinophil – as risk factors for CAD. Due to the proximity of their change in odds ratio to 0, the blood cell counts have weak associations with CAD (**Figure 1**). White blood cell counts have shown to increase the risk of CAD, with evidence showing that white blood cell counts can indicate the progression of the disease (30). A study that attempted to determine the role of white blood cell counts as predictors for CAD discovered that increased levels of most types of white blood cell counts were associated with the increased risk of CAD (20). Our MR results successfully identified the role of white blood cells as exposures that increase the chances of CAD, even with a low association. While eosinophils are a type of white blood cell and have similar functions as disease-fighting cells, eosinophil blood cell count has no role as a risk factor due to the nearly 0 change in odds ratio (**Figure 1**). Moreover, red blood cell count and platelet cell count have similar roles as white blood cell count in acting as a causal risk factor. However, due to such weak association with CAD in our study, strong conclusions cannot be made. Multiple studies support this decision as the need for more research studies is required for further assessment (31).

Moreover, our findings further support research that claims that blood pressure readings, LDL cholesterol levels, tobacco smoke, BMI, and type 2 diabetes increase risks of CAD. Our MR analysis also indicated that CRP levels protect against CAD, which is supported by several other outside studies, proving that our research was able to contribute to the need for causal risk factor identification of CAD. In the case of uric acid, we found that it promotes CAD, which differs from what other such studies have proved. Some studies claim that uric acid protects against CAD, while others agree with our results that it progresses the outcome. Our study, in fact, provides new insight by suggesting an association between uric acid and increased CAD. Blood cell counts provided unique results in our study. The counts of white blood cells, neutrophil cells, and lymphocyte cells are seen to increase risks of CAD, but with weak association. One particular study (30) indicates a causal relationship between white blood cells and CAD, but we believe that further research is required before making definitive conclusions, especially since results showed that white blood cells increase risks of CAD while a few specific types of white blood cells (eosinophils, monocytes, and basophils) have no such or weak relationships with CAD.

Additionally, MR analysis relies on instrumental variables derived from genome-wide association analysis, which is reliant on study sample sizes, in particular disease case counts. Therefore, our analysis is limited to disease outcomes which are well studied and at higher incidence. The results may also not be generalizable to all ethnic groups as the data primarily involves participants of European ancestry. Furthermore, some studies focus specifically on coding variation in the genome, in our analysis the type 2 diabetes

dataset by Mahajan *et al.*, 2018 with 81,412 cases is a three-fold increase in effective sample size compared to the previous largest study but is limited in its focus on only studying coding variants in the genome (32). This limitation will reduce the power to detect causal associations as genetic variations in intergenic regions which could contribute to the Mendelian randomization analysis.

Overall, our results support the hypothesis that in a MR study, traits which are correlated with CAD outcomes will be causally associated with CAD risk. To further validate our findings a prospective study conducted to measure risk factor levels over time would provide further evidence for the influence of traits on CAD. Furthermore, studying other cardiovascular disease outcomes including myocardial infarction, ischemic and hemorrhagic stroke, and heart failure may show that individual risk factors associated with CAD risk more broadly influence other disease outcomes.

MATERIALS AND METHODS

Dataset

Our analysis combined results from 28 genetic analyses from 12 unique studies, with data being collected from supplementary tables and repositories associated with the studies listed in **Table 1**.

Mahajan *et al.*, 2018 collected coding variant data on 81,412 type 2 diabetes cases and 370,832 control samples which were collected from the UK Biobank and GERA (Resource for Genetic Epidemiology on Adult Health and Aging) cohorts (32). Systolic blood pressure was recorded using an Omron device in over 317,754 individuals in the UK biobank cohort (33), recorded as the average of two measurements taken in the same visit. The blood pressure measurements were studied by an open data initiative by the Neale lab at the Broad Institute of MIT and Harvard; variants were filtered based on quality control metrics and expected Hardy-Weinberg equilibrium (HWE) (34). Van der Harst *et al.*, 2017 executed a genome-wide association study in 34,541 CAD cases and 261,984 control samples from the UK Biobank resource in order to expand the number of genome-wide significant loci. They were able to identify 64 novel genetic risk loci for CAD, broadening our knowledge of the genetic architecture of CAD (35). The data of blood cell counts such as eosinophils, basophils, neutrophils, monocytes, lymphocytes, platelets, and white blood cells were recorded from 563,085 individuals from the UK Biobank, which allowed for the discovery of 5,106 new genetic variants that are independently associated with 29 blood cell phenotypes (36). Ligthart *et al.*, 2018 used data from 88 studies consisting of 204,402 European individuals to perform two genome-wide association studies of circulating amounts of CRP, revealing 58 distinct genetic loci for CRP (37). Yengo *et al.*, 2018 performed genome-wide association studies of height and BMI that utilized 700,000 UK Biobank individuals. Around 941 single nucleotide polymorphisms (SNPs)– variation of a single position in a DNA sequence – were associated with BMI, and the study as a whole demonstrated that a high sample size results in increased prediction accuracy, which allows for additional understanding of complex trait biology (38). Klimentidis *et al.*, 2020, performed a genome-wide association study of variants associated with lower LDL cholesterol and increased type 2 diabetes risk, using 431,167 UK Biobank individuals for LDL cholesterol and

Trait	Author	Sample size	Disease cases	Source
Type 2 diabetes (adjusted for BMI)	Mahajan A. (32)	298,957	48,286	PMID: 29632382
Systolic blood pressure reading	B. Neale (34)	317,754		Neale lab repository
Uric acid	K. Ishigaki (42)	109,029		PMID: 29403010
Coronary artery disease	Van der Harst P. (35)	547,261	122,733	PMID: 29212778
Eosinophil cell count	D. Vuckovic (36)	563,946		PMID: 32888494
Current tobacco smoking	Ben Elsworth (41)	462,434		DOI: 10.1101/2020.08.10.244293
Basophil cell count	D. Vuckovic (36)	563,946		PMID: 32888494
Neutrophil cell count	D. Vuckovic (36)	563,946		PMID: 32888494
Monocyte cell count	D. Vuckovic (36)	563,946		PMID: 32888494
Lymphocyte cell count	D. Vuckovic (36)	563,946		PMID: 32888494
C-Reactive protein level	S. Ligthart (37)	204,402		PMID: 30388399
BMI	L. Yengo (38)	681,275		PMID: 30124842
Platelet count	D. Vuckovic (36)	350,474		PMID: 32888494
Red blood cell count	Astle WJ (40)	172,952		PMID: 27863252
White blood cell count	D. Vuckovic (36)	563,946		PMID: 32888494
LDL cholesterol levels	Klimentidis Y.C. (39)	431,167		PMID: 32493714

Table 1: Sources for the genome-wide association study data that was utilized in our Mendelian randomization analysis. Genetic data for the 16 human phenotypes studied in our Mendelian randomization analysis were collected and integrated from multiple sources.

898,130 Diabetes Genetics Replication And Meta-analysis consortium individuals for type 2 diabetes (39). They identified 31 loci associated with lower LDL cholesterol and increased type 2 diabetes and broadened understanding of higher type 2 diabetes risk among individuals with low LDL cholesterol (39). To test 29.5 million genetic variants for association with red blood cell, white blood cell, and platelet properties in 173,480 individuals, Astle *et al.*, 2016 performed a genome-wide association analysis in the UK Biobank, resulting in the understanding of the properties of the allelic framework of traits (40). Elsworth *et al.*, 2020 presented the OpenGWAS database which imports and publishes the summary datasets and metadata from genome-wide association studies in order to make the results from such studies more programmatically accessible. The database contains around 126 billion genetic associations from 14,582 genome-wide association study datasets, representing exposures and disease outcomes, including the current tobacco smoking exposure which was utilized in our study (41). Ishigaki *et al.*, 2018 conducted a genome-wide association study of 58 quantitative traits in 162,255 Japanese individuals from the BioBank Japan Project – a hospital registry that collected clinical information – to broaden the understanding of the genetics of the studied traits. One of the 58 traits included uric acid, which was found to be genetically associated with ischemic stroke (42).

The following acquisition data was compiled to prepare for MR analyses using the R programming language. For each genetic variation of interest, we obtained: a) the rsID, b) chromosome and base-pair position, c) reference and alternative alleles of the genetic variation, d) estimated effect of the genetic variation on the outcome, e) *p*-value for the estimated effect of the genetic variation on the outcome. Summary statistics were analyzed with the MR protocol using the TwoSampleMR package.

Mendelian Randomization Sensitivity Analyses

MR uses a large number of genetic variants in its application, which could potentially lead to pleiotropy – when one genetic variation influences several unrelated traits. When there are direct effects from the genetic variants to the disease outcome other than through the pathway mediated by the exposure, they produce false positives and biased causal results, demeaning the validity of the results (43). To combat pleiotropy, two MR methods can be used: MR Egger and Inverse Variance Weighted. MR Egger helps to provide a less biased causal estimate but lacks statistical power (43). The Inverse Variance Weighted method, on the other hand, infers the strength of the causal effect between an exposure and an outcome, while possessing significant statistical power (43).

Software Pipeline

RStudio (Build 351) was utilized to statistically compute the exposures and outcomes in the study and generate forest plots to visualize the potential relationships between the two. MR programs were created using the TwoSampleMR package for each disease outcome. The TwoSampleMR package was utilized to extract association statistics for the genetic variations associated with the exposure and their effect on the outcome in order to estimate the causal relationship. This process is referred to as a MR experiment – the use of genetic variations as instrumental variables to infer the causal effects of multiple risk factors on a single outcome – and is

more efficient than the typical MR experiment that infers the causal effect of a single risk factor on a single outcome. After extraction, a loop was run on each exposure to receive its instruments and the effects of its instruments on the outcome. Within the loop, each exposure and outcome data was then harmonized to be on the same effect allele so that comparisons were able to be made. Finally, the MR was performed, and a scatter plot and dataset were generated to represent the data. This process was utilized for each outcome examined in the study. Regarding the visualization of the data, scatterplots were generated for each exposure as it plotted the SNP effect on the exposure against the SNP effect on the outcome. In a similar fashion, the datasets were created for each exposure as they provided data regarding the *p*-values of the exposure on outcome for each MR method performed (MR Egger and Inverse Variance Weighted). The values retrieved from the Inverse Variance Weighted method were favored due to its efficiency and ability to remain high powered. The *p*-values of each exposure on outcome from the Inverse Variance Weighted method were analyzed, and if it was found to be less than or equal to 0.05 at a 95% confidence interval, then it would be statistically significant, indicating the rejection of the null hypothesis and concluding that the exposure is a risk factor for the outcome. After analysis of *p*-values, forest plots were generated to represent the significant exposures and outcomes.

Received: May 15, 2022

Accepted: August 24, 2022

Published: February 8, 2023

REFERENCES

1. Yang, Liuqiao, *et al.* “Association Between Thromboelastography and Coronary Heart Disease.” *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, vol. 28, May 2022, p. e935340.
2. Brown, Jonathan C., *et al.* “Risk Factors for Coronary Artery Disease.” *StatPearls*, StatPearls Publishing, 2021.
3. *Cardiovascular Diseases (CVDs)*. [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)). Accessed 16 July 2022.
4. “CDC Prevention Programs.” www.heart.org, <https://www.heart.org/en/get-involved/advocate/federal-priorities/cdc-prevention-programs>. Accessed 19 Nov. 2022.
5. *Coronary Artery Disease: Prevention, Treatment and Research*. 7 July 2021, <https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronary-artery-disease-prevention-treatment-and-research>.
6. Smith, G. D., and S. Ebrahim. “Mendelian Randomization: Can Genetic Epidemiology Contribute to Understanding Environmental Determinants of Disease?” *International Journal of Epidemiology*, vol. 32, no. 1, Feb. 2003, <https://doi.org/10.1093/ije/dyg070>.
7. Burgess, Stephen, Adam Butterworth, *et al.* “Mendelian Randomization Analysis with Multiple Genetic Variants Using Summarized Data.” *Genetic Epidemiology*, vol. 37, no. 7, Nov. 2013, pp. 658–65.
8. Burgess, Stephen, George Davey Smith, *et al.* “Guidelines for Performing Mendelian Randomization Investigations.” *Wellcome Open Research*, vol. 4, 2019,

- <https://doi.org/10.12688/wellcomeopenres.15555.2>.
9. *Measuring Blood Pressure*. <https://medlineplus.gov/lab-tests/measuring-blood-pressure/>. Accessed 16 July 2022.
 10. Weber, Thomas, *et al.* "Hypertension and Coronary Artery Disease: Epidemiology, Physiology, Effects of Treatment, and Recommendations : A Joint Scientific Statement from the Austrian Society of Cardiology and the Austrian Society of Hypertension." *Wiener Klinische Wochenschrift*, vol. 128, no. 13-14, July 2016, pp. 467–79.
 11. "High Uric Acid Level." *Cleveland Clinic*, <https://my.clevelandclinic.org/health/symptoms/17808-high-uric-acid-level>. Accessed 15 May 2022.
 12. Chang, Chun-Chin, *et al.* "Association between Serum Uric Acid and Cardiovascular Risk in Nonhypertensive and Nondiabetic Individuals: The Taiwan I-Lan Longitudinal Aging Study." *Scientific Reports*, vol. 8, no. 1, Mar. 2018, pp. 1–6.
 13. Gallucci, G., *et al.* "Cardiovascular Risk of Smoking and Benefits of Smoking Cessation." *Journal of Thoracic Disease*, vol. 12, no. 7, July 2020, <https://doi.org/10.21037/jtd.2020.02.47>.
 14. Alkhawam, Hassan, *et al.* "Coronary Artery Disease in Patients with Body Mass Index ≥ 30 kg/m²: A Retrospective Chart Analysis." *Journal of Community Hospital Internal Medicine Perspectives*, vol. 6, no. 3, 2016, p. 31483, <https://doi.org/10.3402/jchimp.v6.31483>.
 15. Gregory, Anne B., *et al.* "The Relationship between Body Mass Index and the Severity of Coronary Artery Disease in Patients Referred for Coronary Angiography." *Cardiology Research and Practice*, vol. 2017, Apr. 2017, p. 5481671.
 16. *Type 2 Diabetes*. <https://medlineplus.gov/ency/article/000313.htm>. Accessed 17 July 2022.
 17. "C-Reactive Protein Concentration and Risk of Coronary Heart Disease, Stroke, and Mortality: An Individual Participant Meta-Analysis." *The Lancet*, vol. 375, no. 9709, Jan. 2010, pp. 132–40.
 18. Lagrand, W. K., *et al.* "C-Reactive Protein as a Cardiovascular Risk Factor: More than an Epiphenomenon?" *Circulation*, vol. 100, no. 1, July 1999, pp. 96–102.
 19. Verdoia, Monica, *et al.* "Neutrophil to Lymphocyte Ratio and the Extent of Coronary Artery Disease: Results From a Large Cohort Study." *Angiology*, vol. 67, no. 1, Jan. 2016, pp. 75–82.
 20. Madjid, Mohammad, and Omid Fatemi. "Components of the Complete Blood Count as Risk Predictors for Coronary Heart Disease: In-Depth Review and Update." *Texas Heart Institute Journal / from the Texas Heart Institute of St. Luke's Episcopal Hospital, Texas Children's Hospital*, vol. 40, no. 1, 2013, pp. 17–29.
 21. "LDL: The 'Bad' Cholesterol." *Blood, Heart and Circulation*, Dec. 2017, <https://medlineplus.gov/ldlthebadcholesterol.html>.
 22. Weverling-Rijnsburger, Annelies W. E., *et al.* "High-Density vs Low-Density Lipoprotein Cholesterol as the Risk Factor for Coronary Artery Disease and Stroke in Old Age." *Archives of Internal Medicine*, vol. 163, no. 13, July 2003, pp. 1549–54.
 23. Hajar, Rachel. "Risk Factors for Coronary Artery Disease: Historical Perspectives." *Heart Views: The Official Journal of the Gulf Heart Association*, vol. 18, no. 3, July 2017, pp. 109–14.
 24. Malone, John I., and Barbara C. Hansen. "Does Obesity Cause Type 2 Diabetes Mellitus (T2DM)? Or Is It the Opposite?" *Pediatric Diabetes*, vol. 20, no. 1, Feb. 2019, pp. 5–9.
 25. Chiha, Maguy, *et al.* "Diabetes and Coronary Heart Disease: A Risk Factor for the Global Epidemic." *International Journal of Hypertension*, vol. 2012, Oct. 2012, p. 697240.
 26. Bagheri, Babak. "Uric Acid and Coronary Artery Disease, Two Sides of a Single Coin: A Determinant of Antioxidant System or a Factor in Metabolic Syndrome." *JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH*, 2016, <https://doi.org/10.7860/jcdr/2016/16335.7281>.
 27. Rahimi-Sakak, Fatemeh, *et al.* "Serum Uric Acid and Risk of Cardiovascular Mortality: A Systematic Review and Dose-Response Meta-Analysis of Cohort Studies of over a Million Participants." *BMC Cardiovascular Disorders*, vol. 19, no. 1, Oct. 2019, p. 218.
 28. Ndrepepa, Gjin. "Uric Acid and Cardiovascular Disease." *Clinica Chimica Acta; International Journal of Clinical Chemistry*, vol. 484, Sept. 2018, pp. 150–63.
 29. Alderman, Michael H. "Uric Acid and Cardiovascular Risk." *Current Opinion in Pharmacology*, vol. 2, no. 2, Apr. 2002, pp. 126–30.
 30. Twig, Gilad, *et al.* "White Blood Cell Count and the Risk for Coronary Artery Disease in Young Adults." *PLoS One*, vol. 7, no. 10, Oct. 2012, p. e47183.
 31. Puddu, Paolo Emilio, *et al.* "Red Blood Cell Count in Short-Term Prediction of Cardiovascular Disease Incidence in the Gubbio Population Study." *Acta Cardiologica*, vol. 57, no. 3, June 2002, pp. 177–85.
 32. Mahajan, Anubha, *et al.* "Refining the Accuracy of Validated Target Identification through Coding Variant Fine-Mapping in Type 2 Diabetes." *Nature Genetics*, vol. 50, no. 4, Apr. 2018, pp. 559–71.
 33. Bycroft, Clare, *et al.* "The UK Biobank Resource with Deep Phenotyping and Genomic Data." *Nature*, vol. 562, no. 7726, Oct. 2018, pp. 203–09.
 34. *Imputed-v2-Gwas at Master · Nealelab/UK_Biobank_GWAS*. Github, https://github.com/Nealelab/UK_Biobank_GWAS. Accessed 9 July 2022.
 35. van der Harst, Pim, and Niek Verweij. "Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease." *Circulation Research*, vol. 122, no. 3, Feb. 2018, pp. 433–43.
 36. Vuckovic, Dragana, *et al.* "The Polygenic and Monogenic Basis of Blood Traits and Diseases." *Cell*, vol. 182, no. 5, Sept. 2020, pp. 1214–31.e11.
 37. Ligthart, Symen, *et al.* "Genome Analyses of >200,000 Individuals Identify 58 Loci for Chronic Inflammation and Highlight Pathways That Link Inflammation and Complex Disorders." *American Journal of Human Genetics*, vol. 103, no. 5, Nov. 2018, pp. 691–706.
 38. Yengo, Loic, *et al.* "Meta-Analysis of Genome-Wide Association Studies for Height and Body Mass Index in ~700000 Individuals of European Ancestry." *Human Molecular Genetics*, vol. 27, no. 20, Oct. 2018, pp. 3641–49.

39. Klimentidis, Yann C., *et al.* “Phenotypic and Genetic Characterization of Lower LDL Cholesterol and Increased Type 2 Diabetes Risk in the UK Biobank.” *Diabetes*, vol. 69, no. 10, Oct. 2020, pp. 2194–205.
40. Astle, William J., *et al.* “The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease.” *Cell*, vol. 167, no. 5, Nov. 2016, pp. 1415–29. e19.
41. Elsworth, Ben, *et al.* “The MRC IEU OpenGWAS Data Infrastructure.” *bioRxiv*, 10 Aug. 2020, p. 2020.08.10.244293, <https://doi.org/10.1101/2020.08.10.244293>.
42. Kanai, Masahiro, *et al.* “Genetic Analysis of Quantitative Traits in the Japanese Population Links Cell Types to Complex Human Diseases.” *Nature Genetics*, vol. 50, no. 3, Feb. 2018, pp. 390–400.
43. Lin, Zhaotong, *et al.* “Combining the Strengths of Inverse-Variance Weighting and Egger Regression in Mendelian Randomization Using a Mixture of Regressions Model.” *PLoS Genetics*, vol. 17, no. 11, Nov. 2021, p. e1009922.

Copyright: © 2023 Singh and Akbari. All JEI articles are distributed under the attribution non-commercial, no derivative license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>). This means that anyone is free to share, copy and distribute an unaltered article for non-commercial purposes provided the original author and source is credited.