Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Jiyuan Ren Date

Evaluation of the Role of Nutrition Intake in Coronary Artery Disease Using Mendelian Randomization

By

Jiyuan Ren

Master of Public Health

Epidemiology

_____

Yan Sun

Committee Chair

_____

Evaluation of the Role of Nutrition Intake in Coronary Artery Disease Using Mendelian Randomization

By

Jiyuan Ren
B.Sc.
Virginia Polytechnic Institute and State University
2020

Thesis Committee Chair: Yan Sun PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2023

Abstract

Evaluation of the Role of Nutrition Intake in Coronary Artery Disease Using Mendelian Randomization

By Jiyuan Ren

Background: Previous observational studies showed inconsistent evidence supporting the role of nutrient intake in coronary artery disease (CAD). It is inconclusive if observed associations are caused by nutrient intake or by confounders such as other lifestyle factors. Mendelian Randomization (MR) uses genetic variants as instrumental variables (IVs) to estimate the association between the exposure and the outcome with the control of confounders. This study aims to use two-sample MR to investigate the potential causal relationship between several common nutrients and CAD utilizing data from the UK Biobank and summary statistics of Genome-wide association studies (GWAS) of CAD.

Methods*:* The study population consisted of 201,245 white participants from the UK Biobank. Genome-wide association studies (GWAS) of the nutrient intake were performed with adjustments for age, sex, and top 10 principal components to identify independent significant SNPs (linkage disequilibrium $r^2 < 0.1$) as IVs. CAD GWAS results were derived from CARDIoGRAMplusC4D 1000 Genomes-based study. MR analysis was conducted using inverse-variance weighted MR method and MR-Egger regression.

Results*:* Animal fat, fat, magnesium, trans fatty acids, and Vitamin C were significantly associated with incident CAD after adjusting for CAD risk factors. We identified IVs from the GWAS of 13 nutrients to conduct two-sample MR analyses. Alcohol, fat, lactose, and Vitamin C were significantly associated with CAD risk. A one SD increase in alcohol intake increased the risk of CAD by 1.24-fold (95% CI: 1.02-1.50, P = 0.029). Conversely, the corresponding odds ratio of CAD for one SD increase intake of fat, free sugar, lactose and vitamin C were 0.622 (95%CI: 0.400 - 0.968, P = 0.035), 0.647 (95%CI: 0.431 - 0.970, P = 0.035), 0.805 (95%CI: 0.676 - 0.959, P = 0.015) and 0.372 (95%CI: 0.160 - 0.865, P = 0.022), respectively.

Conclusion*:* Significant GWAS findings showed the important role of genetic factors in nutrients intake. The MR results supported that nutrients intake may have potential causal effect on CAD.

Evaluation of the Role of Nutrition Intake in Coronary Artery Disease Using Mendelian Randomization

By

Jiyuan Ren
B.Sc.
Virginia Polytechnic Institute and State University
2020

Thesis Committee Chair: Yan Sun PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2023

Table of Contents

Evaluation of the Role of Nutrition Intake in Coronary Artery Disease Using Mendelian Randomization

Introduction

Coronary artery disease (CAD) affects the arteries of the heart and reduces the blood supply to the heart. CAD is caused by the buildup of atherosclerotic plaque in the coronary artery which obstructs blood flow and then resulting in insufficient oxygen supply to cardiac muscle cells.[1] In acute cases, CAD could leads to congestive heart failure, cardiac arrhythmias, and even death.[2] As one of the most common cardiovascular diseases, CAD is estimated to affect over 20.5 million adults in United States.[3,4] The prevalence of CAD is approximately 7.1% of the population aged >20. Each year there would be 720,000 new CAD cases[4] and the mortality rate of coronary artery disease is approximately 1.8% in 2020 which meanings approximately 382,820 adults die from CAD each year in United States.[4] During 2018 to 2019, treating CAD costs an estimated $117 billion in direct medical costs and $122.9 billion for indirect cost such as lost productivity due to illness or death in United States.[4] Some CAD risk factors are preventable or manageable to reduce the primary and secondary CAD events. Previous large observational studies have identified CAD risk factors including age, sex, cigarette smoking, physical inactivity, high blood pressure, diabetes, obesity, poor nutrition, and unhealthy cholesterol levels.[5] Among these risk factors, a healthy diet plays an important role in preventing diabetes, obesity, high blood pressure and unhealthy cholesterol levels which lead to reduced CAD risk. Better understanding the causal relationship between nutrient intake, particularly individual nutrient components, and CAD could provide the basis for a precise nutrition strategy for CAD prevention.

Previous observational studies have investigated the association between individual nutrients and CAD risk. In a retrospective study of 10,899 patients, researchers found that vitamin D deficiency was associated with a higher risk of coronary artery disease (OR 1.16, 95% CI:

1.01-1.33).[6] In a prospective study of 127,536 participants, the researchers found that a higher intake of polyunsaturated fatty acids (PUFAs) and carbohydrates from whole grains was linked to a decreased risk of CAD.[7] Conversely, carbohydrates from refined starches or added sugars had a positive correlation with risk of CAD.[7] In an a meta-analysis of randomized controlled trials, researchers found that vitamin D supplementation was associated with a higher risk of major cardiovascular events major cardiovascular events.[8] However, the results of these association between different nutrients and risk of CAD are inconsistent across studies. In a meta-analysis of 32 observational studies, the researchers did not observe a significant association between total saturated fatty acids intake, biomarker, or dietary intervention and CAD incidence. But they identified positive association between trans fatty acids intake, biomarker or dietary intervention and CAD incidence.[9] Nutrients is a complex field of study, the large number of confounding factors like lifestyle and other environmental factors making it difficult to controlling for these confounders and accurately compare results using observational data. Therefore, epidemiologic methods such as Mendelian Randomization (MR) could provide supporting evidence for potential causal effect between individual nutrients and CAD.

Mendelian randomization uses genetic variants as instrumental variables for the exposure variable which is susceptible to confounders.[10] Two-sample MR depends on three assumptions that genetic variants 1) are associated with the exposure 2) are not associated with confounder and 3) influence the outcome only through the exposure.[11] Two-sample MR, which increasing more statistical power, uses association results estimated in two non-overlapping sets of population instead of a single population.[10] Several two-sample MR studies have been conducted to examine the association between different nutrients and risk of CAD. For example, a study published in 2020 used two-sample MR to investigate the association between alcohol consumption and risk of CAD using data form UK Biobank.[12]

The results provided evidence for that higher intake of alcohol was associated with an increased risk of CAD which was consistent with observational studies.[12] Other studies have investigated the association between serum calcium and risk of CAD in European ancestry.[13] The results suggested that higher serum calcium were associated with a lower risk of CAD.[13] However, no study has yet systematically demonstrated the relationship between common nutrients and CAD using two sample MR study. Therefore, in this study, we will use two sample MR and similar strategies to provide new evidence for the potential causal relationship between several common nutrients and risk of CAD utilizing the data provided by the UK Biobank.

Methods

Study Population and data sources

The dataset used for this study is available in the UK Biobank (UKB) public repository. This research has been conducted using the UKB Resource under Application Number "34031". The UK Biobank is a large-scale, prospective population-based study of health and disease of over 500,000 people located in various parts of the United Kingdom between 2006 and 2010.[14] The data collected includes genetic, physical, lifestyle and health information. For the present study, only white participants were included. To do this, we only kept all participants that coded as 'White', 'Irish' or 'Any other white background' in ethnic background dataset.

The nutrients data was obtained from a dietary questionnaire, based on a 24-hour dietary recall of the previous day in the UKB. These questionnaires were sent out at the assessment visit at the end of the recruitment phase. After the close of recruitment, 4 additional questionnaire rounds were conducted online. However, not all participants responded every round, thus some may have only answered one survey out of the five. In order to maximize the sample size, the first available nutrition data for each subject was applied in this study. In

this study, the nutrients including alcohol, animal fat, calcium, energy, fat, free sugar, lactose, magnesium, omega-3 fatty acids, omega-6 fatty acids, potassium, sodium, starch, total sugar, total weight, trans fatty acids, vitamin A, vitamin B6, vitamin B12, vitamin C, vitamin D, vitamin E and zinc.

We identified individuals with a history of CAD utilizing Hospital Episode Statistics data to locate the earliest record of International Classification of Diseases (ICD)-10 codes I20 to I25 and Office of Population Censuses and Surveys-4 (OPCS-4) codes K40-K46, K49, K50, or K75.[15]

Other demographic data and risk factors including age, sex, frequent alcohol use (defined as alcohol consumption ≥ 3 times per week), smoking status, body mass index (BMI), high-density lipoprotein (HDL), low-density lipoprotein (HDL), systolic blood pressure, diastolic blood pressure, cholesterol, diabetes, hypertension, and stroke history were all provided by UKB which were obtained at enrollment of participants.

GWAS of the Nutrition data

Genome-wide association studies of the association between the genotype and nutrition intake were performed using liner model. PLINK2 were used as the primary tool for the GWAS. [16,17] We adjusted for age, sex and Top 10 principal components provided by the UKB. For quality control, we filtered the SNPS based on minor allele frequency (MAF) ≥ 0.01. Genome-wide significance threshold $P \leq 5 \times 10^{-8}$ was applied to account for multiple testing. Subsequently, the SNPs from the nutrition GWAS results were aligned with those of the CAD GWAS, with only SNPs contained in both datasets being retained ensuring maximum number of SNPs can be used for Two-sample MR. SNP2GENE function, which is provided by the Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA) was applied to annotate and interpret the GWAS results.[18] Only independent

significant SNPs which are independent each other at linkage disequilibrium $r^2 < 0.1$ were kept as the genetic instrumental variables (GIVs). Only nutrients with a minimum of two SNPs were retained for the subsequent two-sample Mendelian randomization analysis. These nutrients included alcohol, animal fat, calcium, energy, fat, free sugar, lactose, magnesium, omega-3 fatty acids, omega-6 fatty acids, sodium, total sugars, and total weight of food.

CAD GWAS data source

GWAS results for coronary artery disease are derived from CARDIoGRAMplusC4D 1000 Genomes-based GWAS. This study examined the associations of both common and low-frequency variants with CAD involving 60,801 CAD cases and 123,504 controls.[19] The majority of participants, 77%, are European ancestry, while 13% and 6% were of South Asian and East Asian descent, respectively, with smaller samples of Hispanic and African American individuals.[19] The imputation process was based on the 38 million variants contained in the 1000 Genomes Project phase 1 v3 reference panel.[19]

Statistical Analyses

Statistical analyses were performed using R (version 4.2.1). We first stratify the population by the CAD status. T test and chi-squared test were used to compare the difference of baseline characteristics between CAD case and non-CAD control groups. For individual-level observational data from the UKB, we performed linear regression to analyze the association of 23 nutrients and CAD with an unadjusted model and an adjusted model controlling for age, sex, frequent alcohol use (defined as alcohol consumption ≥ 3 times per week), smoking, BMI, HDL, LDL, systolic blood pressure, diastolic blood pressure, cholesterol, diabetes, hypertension, and stroke history.

R package ‘TwoSampleMR’ were used to performing two sample MR using the two set of genetic association estimates data.[20,21] The genetic association estimates data of 23 nutritional

variables and coronary artery disease (CAD) were harmonized independently, and palindromic SNPs were then removed. Inverse-variance weighted MR method were used to estimate the association between 23 nutrients data and CAD outcome separately. MR-Egger test was also used to test whether there was directional pleiotropy problem between SNPs and the outcome.[22] MR-Egger model only works when number of SNPs > 2. Therefore, for nutrients with only two SNPs which including animal fat, calcium, energy, magnesium, and sodium, only result of inverse-variance weighted model were used.

Results

Descriptive statistics

This study included 201,245 participants with a mean age of 58.2 year, with 44.9% males. All participants were of white European ancestry. Other baseline characteristics and nutrition data are summarized in Table 1. We also compared the baseline characteristics between participants with and without CAD. This encompassed both pre- and post-enrollment CAD participants, incorporating a total of 18,301 cases and 182,944 controls. As Table 1 shown, all relative risk factors for CAD except for SBP were significantly higher in CAD group. A comparison of the baseline characteristics between participants included in this study and those excluded due to missing nutrition data can be found in Supplemental Table 1. As the Supplemental Table 1 shown, participants with nutrition data had a slightly higher average age (58.22 years) compared to those without nutrition data (56.71 years). Mean HDL, LDL, and cholesterol levels were higher in the group with nutrition data. Systolic and diastolic blood pressure were also slightly different between the groups.

Associations between nutrient intake and incidence CAD

For individual-level observational data, 17 out of 23 nutrients were significantly associated with an increased risk of CAD in unadjusted logistic regression models. The corresponding

odds ratio of CAD for one SD increased intake ranged from 1.02 for total weight of food (95% CI: 1.003-1.041, P = 0.02) to 1.11 for Vitamin B6 (95% CI: 1.087-1.126, P<0.001). After adjusted for common risk factors of CAD, only animal fat, fat, magnesium, trans fatty acids and Vitamin C were significantly associated with risk of CAD. However, the direction of these correlations is opposite to that of the unadjusted model with 1 SD increase in intake of animal fat was associated with an odds ratio of 0.97 for CAD (95%CI: 0.954 - 0.995, P = 0.0136). The corresponding odds ratio of CAD for 1 SD increase intake of other nutrients were 0.97 for fat (95%CI: 0.953- 0.993, P = 0.009), 0.97 for magnesium (95%CI: 0.953 - 0.994, P = 0.013), 0.97 for trans fatty acids (95%CI: 0.952 - 0.993, P = 0.008) and 0.98 for Vitamin C (95%CI: 0.956 - 0.997, P=0.028). Detailed results were shown in Figure 1, 2 and Supplemental Table 2.

GWAS of nutrient intake

The results of the nutrition GWAS were summarized in Supplemental Table 3. The table details the results of the GWAS that was conducted to identify genetic variants associated with nutrients intake. The number of SNPs for the GWAS results of different nutrients intake were highly variable. Inflation factors for all GWAS results were lower than the 1.10 threshold. After annotation using FUMA, we did not identify enough independent significant SNPs (number of SNPs ≥ 2) for potassium, trans fat acids, Vitamin A, Vitamin B6, Vitamin B12, Vitamin D, Vitamin E, and zinc to perform two-sample MR.

Two-sample MR

Palindromic SNPs were removed for some nutrients in this step. Only 13 of 23 nutrients have enough SNPs for performing two sample MR. The results of two-sample MR were summarized in Figure 3 and Supplemental Table 4 Among these 13 nutrients, only alcohol, fat, lactose, and Vitamin C were significantly associated with the risk of CAD. Using the

Inverse-Variance Weighted (IVW) method, A one SD increase in alcohol intake was associated with a 1.24-fold increased risk of CAD (95%CI: 1.02-1.50, P = 0.029). Conversely, the corresponding odds ratio of CAD for one SD increase intake of other nutrients was 0.622 for fat (95%CI: 0.400 - 0.968, P = 0.035), 0.647 for free sugar (95%CI: 0.431 - 0.970, P = 0.035), 0.805 for lactose (95%CI: 0.676 - 0.959, P = 0.015) and 0.372 for Vitamin C (95%CI: 0.160 - 0.865, P = 0.022). The MR-Egger method, which helps address potential pleiotropy, requires a minimum of three SNPs for each nutrient, and several nutrients in our study did not meet this requirement. As a result, we were unable to apply the MR-Egger method for those nutrients.

Discussion

This study systematically investigated the relationship between common nutrients and CAD incidence using observational association analysis and the two-sample MR method. The two-sample MR results supported that alcohol intake increases the CAD risk. The results are consistent with large-scale observational studies[23] and recent two-sample MR studies[24], which also supported that alcohol intake increases CAD risk.

Vitamin C was also found to have an inverse association with CAD incidence with a relatively very low odds ratio. The mechanism of Vitamin C in CAD is closely associated with its antioxidative property. It can reduce the serum uric acid (UA) level which is an inflammation factor that is associated with CAD risk.[25] A previous case-control study have found that serum level of vitamin C is significantly higher in normal group compared to CAD cases.[25] A previous MR study also confirmed the protect effect cardioembolic stroke.[26] However, another Mendelian randomization study failed to find an association between Vitamin C and CAD.[27] The results of these studies are conflicting and further research is needed to investigate the role of Vitamin C in CAD. In addition, further studies should also

investigate the underlying mechanisms of the possible protective effect of Vitamin C in CAD.

In this study, we also found that fat intake is negatively associated with the risk of CAD. High fat diet was believed to be a risk factor for CAD. A high-fat diet can cause increased levels of LDL cholesterol[28] and high levels of LDL cholesterol is a common risk factor for CAD.[1] However, we observed an opposite MR association which was in the same direction of association in the observational analysis in the UK Biobank. One possible explanation is that the fat data contains both beneficial components, such as polyunsaturated fatty acids, and detrimental components, such as saturated and trans fats. Previous studies have shown that polyunsaturated fatty acids are protective factors for CAD. In contrast, higher intake of saturated fats and trans fats is associated with higher risk of CAD.[7,29] Combining these fats with different effects could be one of the reasons for the unanticipated result. Due to the limitation of the dataset, we were unable to assess the composition of the fat data in detail. Therefore, further research is necessary to explore the relationship between different fat types and CAD risk.

Another unexpected finding was the inverse correlation between the consumption of free sugar and CAD. This finding was unexpected because it is contrary to the widely held belief that consuming too much free sugar can lead to increased risk of CAD. High consumption of carbohydrates from refined starches or added sugars is a common risk factor for CAD.[7] This suggests that there may be other unmeasured factors that can influence the risk of CAD.

The MR results also demonstrated an inverse association between lactose intake and CAD. Lactose is a sugar found in milk and milk products suggesting that individuals who consume higher levels of dairy products are less likely to develop CAD. This finding is consistent with some previous studies that have suggested that dairy product intake may have a protective effect against calcified atherosclerotic plaques in the coronary arteries (CAC), a predictor for

CAD.[30] This study found an inverse association of cheese intake with CAC but not yogurt and milk consumption. While we were unable to tell the specific source of the lactose in our dataset, our findings provide further evidence of the protective effect of dairy products against CAD.

This study has several limitations. Firstly, this study mainly focuses on the European ancestry to ensure more stable GWAS results. As such, results may not be generalizable to other ethnicities and ancestries. Secondly, Nutrition data in UKB are based on the 24-hour recall of food consumption by the participants. This method is subject to recall bias and may not be accurate which can potentially lead to inaccurate results. Thirdly, the strengths of the genetic instruments for different nutrients vary considerably. For example, we identified thousands of SNPs associated with alcohol intake. However, we failed to find sufficient SNPs for performing two sample MR for several other nutrients. Another limitation is the pleiotropy assumption in Mendelian Randomization. Pleiotropy occurs when a genetic variant influences outcomes directly, which may lead to biased MR results.[22] In our analysis, we were unable to apply the MR-Egger method for several nutrients due to an insufficient number of SNPs, which would have helped address potential pleiotropy. Additionally, the dietary assessments in UKB are limited to a few nutrients and did not provide specific source of the nutrients, which may not capture the full range of dietary intake. Therefore, further studies with more complete dietary assessment and mixed population should be conducted to confirm the results.

Conclusion

In this study, genome-wide association studies (GWAS) were conducted to examine the genetic determinants of nutrient intake. Significant GWAS findings showed the important role of genetic factors in nutrients intake. Our MR results demonstrated an association

between multiple nutrients and the risk of CAD using the two-sample MR approach. Our findings provide novel evidence for potential causal relationships between several common nutrients and the risk of CAD. A better understanding of the potential causal effect of individual nutrients on CAD could lead to more effective strategies for preventing CAD events.

References

1. Malakar AK, Choudhury D, Halder B, Paul P, Uddin A, Chakraborty S. A review on coronary artery disease, its risk factors, and therapeutics. *Journal of Cellular Physiology*. 2019-10-01 2019;234(10):16812-16823. doi:10.1002/jcp.28350
2. Cassar A, Holmes DR, Rihal CS, Gersh BJ. Chronic Coronary Artery Disease: Diagnosis and Management. *Mayo Clinic Proceedings*. 2009-12-01 2009;84(12):1130-1146. doi:10.4065/mcp.2009.0391
3. Duggan JP, Peters AS, Trachiotis GD, Antevil JL. Epidemiology of Coronary Artery Disease. *Surg Clin North Am*. Jun 2022;102(3):499-516. doi:10.1016/j.suc.2022.01.007
4. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart Disease and Stroke Statistics—2023 Update: A Report From the American Heart Association. *Circulation*. 2023-01-25 2023;doi:10.1161/cir.0000000000001123
5. Torpy JM. Coronary Heart Disease Risk Factors. *JAMA*. 2009-12-02 2009;302(21):2388. doi:10.1001/jama.302.21.2388
6. Vacek JL, Vanga SR, Good M, Lai SM, Lakkireddy D, Howard PA. Vitamin D deficiency and supplementation and relation to cardiovascular health. *Am J Cardiol*. Feb 01 2012;109(3):359-63. doi:10.1016/j.amjcard.2011.09.020
7. Li Y, Hruby A, Bernstein AM, et al. Saturated Fats Compared With Unsaturated Fats and Sources of Carbohydrates in Relation to Risk of Coronary Heart Disease: A Prospective Cohort Study. *J Am Coll Cardiol*. Oct 06 2015;66(14):1538-1548. doi:10.1016/j.jacc.2015.07.055
8. Mao PJ, Zhang C, Tang L, et al. Effect of calcium or vitamin D supplementation on vascular outcomes: a meta-analysis of randomized controlled trials. *Int J Cardiol*. Oct 30 2013;169(2):106-11. doi:10.1016/j.ijcard.2013.08.055
9. Chowdhury R, Warnakula S, Kunutsor S, et al. Association of dietary, circulating, and supplement fatty acids with coronary risk: a systematic review and meta-analysis. *Ann Intern Med*. Mar 18 2014;160(6):398-406. doi:10.7326/M13-1788
10. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int J Epidemiol*. Dec 01 2016;45(6):1717-1726. doi:10.1093/ije/dyx028
11. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ*. Jul 12 2018;362:k601. doi:10.1136/bmj.k601
12. Larsson SC, Burgess S, Mason AM, Michaëlsson K. Alcohol Consumption and Cardiovascular Disease. *Circulation: Genomic and Precision Medicine*. 2020-06-01 2020;13(3)doi:10.1161/circgen.119.002814
13. Xu L, Lin SL, Schooling CM. A Mendelian randomization study of the effect of calcium on coronary artery disease, myocardial infarction and their risk factors. *Scientific Reports*. 2017-02-14 2017;7(1):42691. doi:10.1038/srep42691
14. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine*. 2015-03-31 2015;12(3):e1001779. doi:10.1371/journal.pmed.1001779
15. Liu C, Dhindsa D, Almuwaqqat Z, et al. Association Between High-Density Lipoprotein Cholesterol Levels and Adverse Cardiovascular Outcomes in High-risk Populations. *JAMA Cardiology*. 2022-07-01 2022;7(7):672. doi:10.1001/jamacardio.2022.0912
16. *PLINK*. Version 2.0. www.cog-genomics.org/plink/2.0/
17. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK:

rising to the challenge of larger and richer datasets. *GigaScience*. 2015-12-01 2015;4(1)doi:10.1186/s13742-015-0047-8

18. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*. 2017-11-28 2017;8(1)doi:10.1038/s41467-017-01261-5

19. A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. *Nature Genetics*. 2015-10-01 2015;47(10):1121-1130. doi:10.1038/ng.3396

20. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife*. 2018-05-30 2018;7doi:10.7554/elife.34408

21. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLOS Genetics*. 2017-11-17 2017;13(11):e1007081. doi:10.1371/journal.pgen.1007081

22. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *European Journal of Epidemiology*. 2017-05-01 2017;32(5):377-389. doi:10.1007/s10654-017-0255-x

23. Song RJ, Nguyen X-MT, Quaden R, et al. Alcohol Consumption and Risk of Coronary Artery Disease (from the Million Veteran Program). *The American Journal of Cardiology*. 2018-05-01 2018;121(10):1162-1168. doi:10.1016/j.amjcard.2018.01.042

24. Biddinger KJ, Emdin CA, Haas ME, et al. Association of Habitual Alcohol Intake With Risk of Cardiovascular Disease. *JAMA Network Open*. 2022-03-25 2022;5(3):e223849. doi:10.1001/jamanetworkopen.2022.3849

25. Torkzaban A, Naeini AA, Hassanzadeh A, Namdari M. The Relationship between Serum Vitamin C and Uric Acid Levels, Antioxidant Status and Coronary Artery Disease: a Case-Control Study. *Clinical Nutrition Research*. 2020-01-01 2020;9(4):307. doi:10.7762/cnr.2020.9.4.307

26. Chen L, Sun X, Wang Z, et al. The impact of plasma vitamin C levels on the risk of cardiovascular diseases and Alzheimer's disease: A Mendelian randomization study. *Clinical Nutrition*. 2021-10-01 2021;40(10):5327-5334. doi:10.1016/j.clnu.2021.08.020

27. Zhu J, Ling Y, Tse LA, Kinra S, Li Y. Circulating vitamin C and the risk of cardiovascular diseases: A Mendelian randomization study. *Nutr Metab Cardiovasc Dis*. Jul 22 2021;31(8):2398-2406. doi:10.1016/j.numecd.2021.04.023

28. Burén J, Ericsson M, Damasceno N, Sjödin A. A Ketogenic Low-Carbohydrate High-Fat Diet Increases LDL Cholesterol in Healthy, Young, Normal-Weight Women: A Randomized Controlled Feeding Trial. *Nutrients*. 2021-03-02 2021;13(3):814. doi:10.3390/nu13030814

29. Tayyem RF, Al-Shudifat AE, Hammad S, Agraib LM, Azab M, Bawadi H. Fat intake and the risk of coronary heart disease among Jordanians. *Nutr Hosp*. Apr 16 2020;37(2):313-320. doi:10.20960/nh.02761

30. Neisius U, Zhou G, Ward RE, Ellison RC, Gaziano JM, Djoussé L. Dairy product consumption and calcified atherosclerotic plaques in the coronary arteries: The NHLBI Family Heart Study. *Clin Nutr ESPEN*. Jun 2022;49:517-521. doi:10.1016/j.clnesp.2022.02.112

Tables and Figures

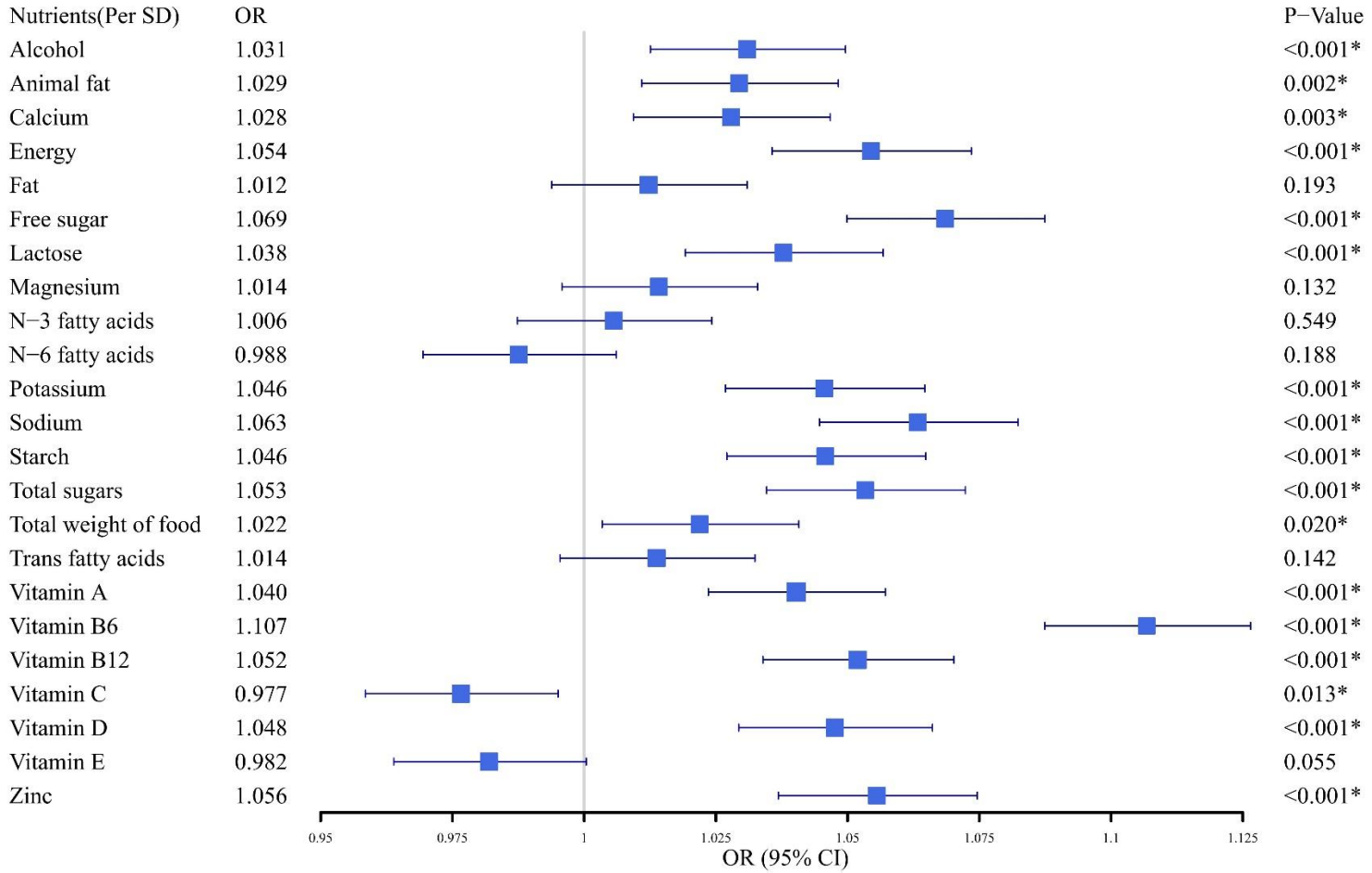
Table 1. Baseline characteristics : Categorized by CAD status.

	Overall (n=201,245)	Without CAD (n=182,944)	With CAD (n=18,301)	P-Value[a]
Age at Data Collection (mean (SD))	58.22 (7.93)	57.77 (7.94)	62.75 (6.31)	<0.001
Male (%)	90416 (44.9)	77968 (42.6)	12448 (68.0)	<0.001
Ever smoker (%)	88296 (44.0)	89712 (49.2)	8678 (47.6)	<0.001
Frequent alcohol use (%)[b]	98390 (49.0)	78020 (42.7)	10276 (56.4)	<0.001
HDL, mg/dL (mean (SD))	1.48 (0.39)	1.50 (0.39)	1.31 (0.35)	<0.001
LDL, mg/dL (mean (SD))	3.57 (0.85)	3.59 (0.84)	3.29 (0.96)	<0.001
DBP, mmHg (mean (SD))	81.79 (10.58)	81.78 (10.54)	81.93 (10.94)	0.059
SBP, mmHg (mean (SD))	138.97 (19.37)	138.49 (19.30)	143.68 (19.44)	<0.001
Cholesterol, mg/dL (mean (SD))	5.72 (1.12)	5.76 (1.10)	5.28 (1.28)	<0.001
BMI, kg/m^2 (mean (SD))	26.94 (4.64)	26.78 (4.59)	28.57 (4.84)	<0.001
Diabetes (%)	7961 (4.0)	5914 (3.2)	2047 (11.2)	<0.001
Hypertension (%)	48780 (24.3)	40477 (22.1)	8303 (45.5)	<0.001
Stroke History (%)	2215 (1.1)	1601 (0.9)	614 (3.4)	<0.001
Total weight of food, g (mean (SD))	3,250.62 (880.47)	3,247.98 (875.67)	3,277.01 (926.70)	<0.001
Animal fat, g (mean (SD))	40.98 (23.05)	40.99 (23.02)	40.93 (23.30)	0.76
Calcium, mg (mean (SD))	994.30 (391.97)	993.71 (391.41)	1,000.19 (397.49)	0.033
Magnesium, mg (mean (SD))	335.36 (106.81)	335.26 (106.65)	336.66 (112.26)	0.093
Zinc, mg (mean (SD))	9.78 (3.86)	9.77 (3.85)	9.99 (4.13)	<0.001
Energy, kj (mean (SD))	8,676.94 (2,813.53)	8,664.37 (2,800.68)	8,802.51 (2,936.02)	<0.001
Fat, g (mean (SD))	73.27 (33.03)	73.30 (32.98)	72.94 (33.54)	0.155
Free sugar, g (mean (SD))	60.77 (40.40)	60.51 (40.16)	63.32 (42.68)	<0.001
Lactose, g (mean (SD))	14.15 (8.59)	14.12 (8.57)	14.42 (8.77)	<0.001
N-3 fatty acids, g (mean (SD))	1.99 (1.20)	1.99 (1.20)	2.00 (1.24)	0.445
N-6 fatty acids, g (mean (SD))	10.98 (6.07)	10.99 (6.07)	10.88 (6.05)	0.017
Potassium, mg (mean (SD))	3,705.59 (1,216.59)	3,699.89 (1,207.99)	3,762.62 (1,298.13)	<0.001

Protein, g (mean (SD))	81.02 (28.25)	80.85 (28.05)	82.68 (30.20)	<0.001
Sodium, mg (mean (SD))	1,964.54 (914.17)	1,958.94 (910.56)	2,020.59 (947.70)	<0.001
Starch, g (mean (SD))	129.29 (56.47)	128.99 (56.33)	132.27 (57.84)	<0.001
Total sugars, g (mean (SD))	126.03 (54.72)	125.75 (54.40)	128.81 (57.76)	<0.001
Trans fatty acids, g (mean (SD))	1.20 (0.80)	1.20 (0.80)	1.18 (0.80)	0.002
Vitamin A retinol equivalents, ug (mean (SD))	980.03 (1,272.33)	974.96 (1,252.69)	1,030.73 (1,453.23)	<0.001
Vitamin B6, mg (mean (SD))	2.08 (0.79)	2.07 (0.78)	2.17 (0.85)	<0.001
Vitamin B12, ug (mean (SD))	6.20 (3.99)	6.18 (3.95)	6.43 (4.35)	<0.001
Vitamin C, mg (mean (SD))	129.65 (88.07)	129.78 (87.87)	128.38 (90.08)	0.04
Vitamin D, ug (mean (SD))	3.62 (3.50)	3.60 (3.48)	3.84 (3.63)	<0.001
Vitamin E, mg (mean (SD))	11.06 (5.35)	11.07 (5.35)	10.96 (5.39)	0.005

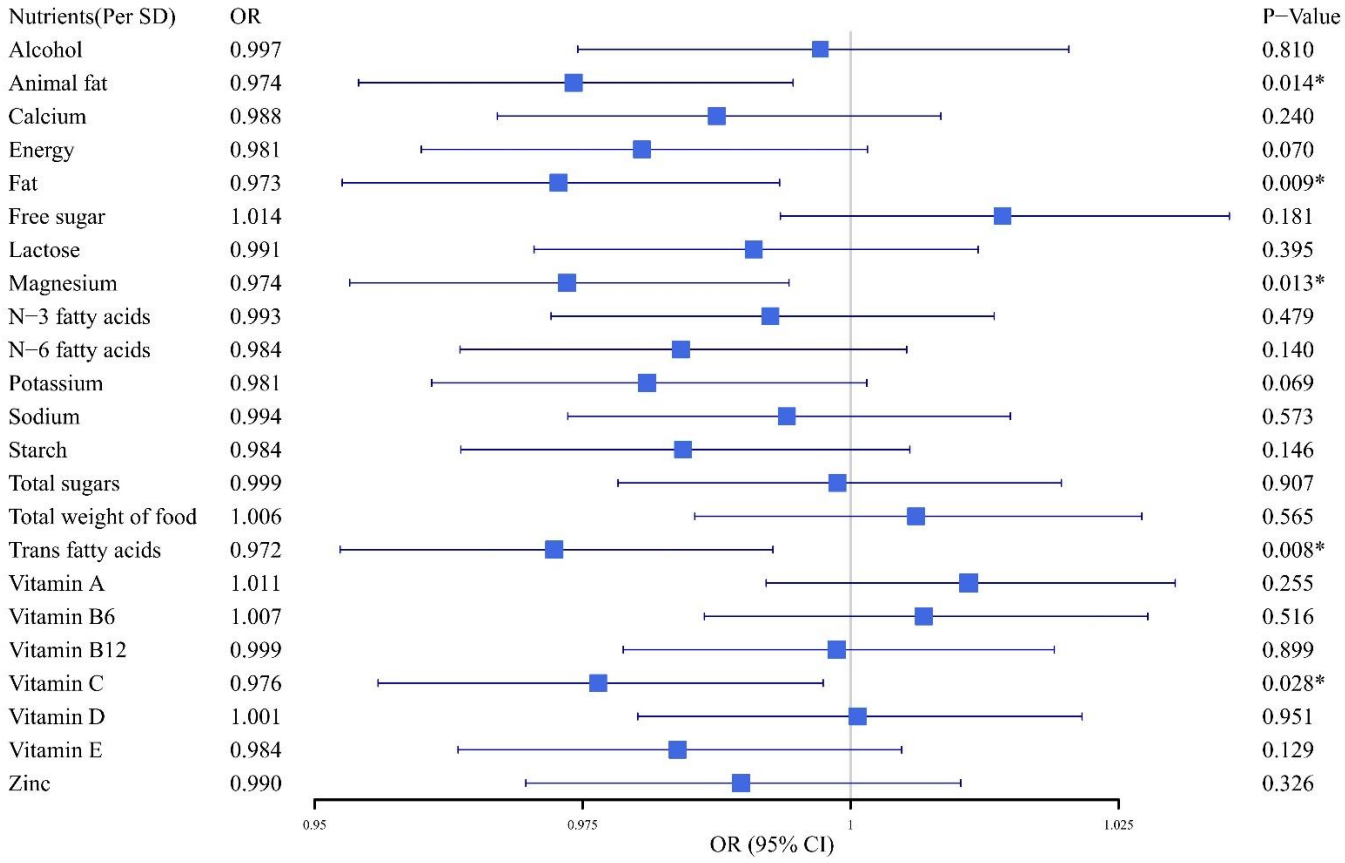
a. A chi-squared test was used for categorical variables and a two-sample t-test was used for continuous variables.
b. Frequent alcohol use is defined as alcohol consumption ≥3 times per week.

Figure 1. Univariate associations between nutrients (per SD increase) and CAD incidence (unadjusted)



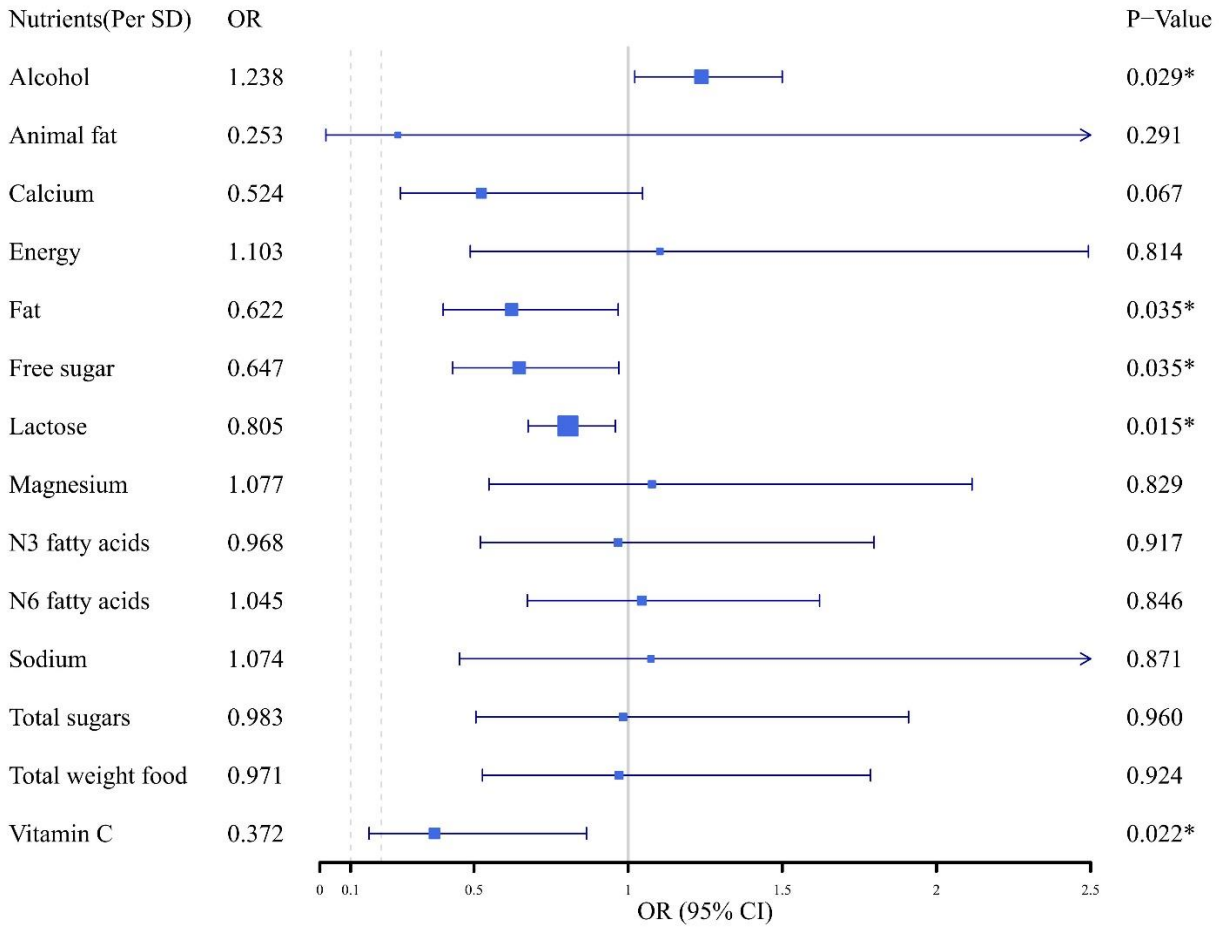
Notes: Horizontal bar indicates 95% CI of the OR.

Figure 2. Univariate associations between nutrients (per SD increase) and the risk of CAD (adjusted)



Notes: Model adjusted for age, sex, frequent alcohol use (defined as alcohol consumption ≥ 3 times per week), smoking status, body mass index (BMI), high-density lipoprotein (HDL), low-density lipoprotein (HDL), systolic blood pressure, diastolic blood pressure, cholesterol, diabetes, hypertension, and stroke history. The horizontal bar indicates 95% CI of the OR. * P-value <0.05

Figure 3. Mendelian Randomization Associations between nutrients (per SD increase) and CAD incidence .



Notes: Horizontal bar indicates 95% CI of the OR (inverse-variance weighted MR method). *P-value <0.05

Supplemental Table 1. Comparison of baseline characteristics among subjects based on availability of nutrition data

	Participants with nutrition data	Participants without nutrition data	P-Value
N	201245	301291	
Age (mean (SD))	58.22 (7.93)	56.71 (8.22)	<0.001
HDL, mg/dl (mean (SD))	1.48 (0.39)	1.43 (0.38)	<0.001
LDL, mg/dl (mean (SD))	3.57 (0.85)	3.55 (0.88)	<0.001
DBP, mmHg (mean (SD))	81.79 (10.58)	82.50 (10.78)	<0.001
SBP, mmHg (mean (SD))	138.97 (19.37)	140.28 (19.90)	<0.001
Cholesterol, mg/dl (mean (SD))	5.72 (1.12)	5.67 (1.16)	<0.001
Frequent alcohol use (%)[a]	98390 (49.0)	118632 (39.6)	<0.001
Current/former smoker (%)	88296 (44.0)	137753 (46.1)	<0.001
BMI, kg/m² (mean (SD))	26.94 (4.64)	27.76 (4.88)	<0.001
Male (%)	90416 (44.9)	138718 (46.0)	<0.001
Diabetes (%)	7961 (4.0)	18441 (6.2)	<0.001
Hypertension (%)	48780 (24.3)	86982 (29.1)	<0.001
Stroke history (%)	2215 (1.1)	5453 (1.8)	<0.001
Cad incidence (%)	12037 (6.0)	23406 (7.8)	<0.001
All cad cases (%)	18301 (9.1)	37639 (12.5)	<0.001

a. Frequent alcohol use defined as alcohol consumption ≥3 times per week.

Supplemental Table 2. Univariate associations between nutrients (per SD increase) and the CAD incidence

	Model 1[a]		Model 2[a]	
Nutrients (Per SD)	**OR (95% CI)**	**P value**	**OR (95% CI)**	**P value**
Alcohol	1.03(1.01-1.05)	**<0.001**	0.997(0.975-1.02)	0.810
Animal fat	1.03(1.01-1.05)	**0.002**	0.974(0.954-0.995)	**0.014**
Calcium	1.03(1.01-1.05)	**0.003**	0.988(0.967-1.008)	0.240
Energy	1.05(1.04-1.07)	**<0.001**	0.981(0.96-1.002)	0.070
Fat	1.01(0.994-1.03)	0.193	0.973(0.953-0.993)	**0.010**
Free sugar	1.07(1.05-1.09)	**<0.001**	1.01(0.993-1.035)	0.182
Lactose	1.04(1.02-1.06)	**<0.001**	0.991(0.97-1.012)	0.395
Magnesium	1.01(0.996-1.03)	0.132	0.974(0.953-0.994)	**0.013**
N3 fatty acids	1.01(0.987-1.023)	0.549	0.993(0.972-1.013)	0.479
N6 fatty acids	0.988(0.969-1.01)	0.188	0.984(0.964-1.005)	0.140
Potassium	1.05(1.03-1.06)	**<0.001**	0.981(0.961-1.002)	0.069
Sodium	1.06(1.04-1.08)	**<0.001**	0.994(0.974-1.015)	0.573
Starch	1.05(1.03-1.06)	**<0.001**	0.984(0.964-1.006)	0.146
Total sugars	1.05(1.03-1.07)	**<0.001**	0.999(0.978-1.02)	0.907
Total weight of food	1.02(1.003-1.04)	**0.020**	1.006(0.985-1.027)	0.565
Trans fatty acids	1.01(0.995-1.03)	0.142	0.972(0.952-0.993)	**0.008**
Vitamin A	1.04(1.02-1.06)	**<0.001**	1.011(0.992-1.03)	0.255
Vitamin B6	1.11(1.09-1.13)	**<0.001**	1.007(0.986-1.028)	0.516
Vitamin B12	1.05(1.03-1.07)	**<0.001**	0.999(0.979-1.019)	0.899
Vitamin B	0.977(0.959-0.995)	**0.013**	0.976(0.956-0.997)	**0.028**
Vitamin D	1.05(1.03-1.07)	**<0.001**	1.001(0.98-1.022)	0.951
Vitamin E	0.982(0.964-1.0004)	0.055	0.984(0.963-1.005)	0.129
Zinc	1.06(1.04-1.07)	**<0.001**	0.99(0.97-1.01)	0.326

a. Model 1 unadjusted, Model 2 adjusted for for age, sex, frequent alcohol use (defined as alcohol consumption ≥ 3 times per week), smoking status, BMI, HDL, LDL, systolic blood pressure, diastolic blood pressure, cholesterol, diabetes, hypertension, and stroke history.

Supplemental Table 3. Summary of GWAS of nutrient intake.

Nutrients	SNPs with maf≥0.01	Independent significant SNPs [a]	SNPs used in two sample MR[b]	Inflation factor
Alcohol	4393	87	28	1.053
Animal fat	15	2	2	1.044
Calcium	198	2	2	1.042
Energy	78	2	2	1.052
Fat	184	5	4	1.051
Free sugar	500	9	8	1.051
Lactose	1490	48	30	1.041
Magnesium	209	4	2	1.068
N3 fatty acids	104	4	3	1.034
N6 fatty acids	173	7	6	1.038
Potassium	48	1	N/A	1.058
Sodium	32	2	2	1.033
Starch	53	1	N/A	1.037
Total sugars	145	4	3	1.056
Total weight of food	758	4	3	1.078
Trans fatty acids	20	1	N/A	1.034
Vitamin A	7	1	N/A	1.011
Vitamin B6	0	0	N/A	1.069
Vitamin B12	0	0	N/A	1.028
Vitamin C	153	2	2	1.080
Vitamin D	0	0	N/A	1.028
Vitamin E	18	1	N/A	1.073
Zinc	0	0	N/A	1.033

a. Independent significant SNPs are SNPs that independent each other at $r^2 \geq 0.1$.
b. Several SNPs were removed because of palindromic and not inferable.

Supplemental Table 4. Mendelian Randomization **Associations between nutrients (per SD increase) and CAD incidence .**

Nutrients	OR (95% CI)	P VALUE
ALCOHOL	1.24 (1.02-1.50)	0.029*
Animal fat	0.253 (0.0198-3.24)	0.291
Calcium	0.524 (0.262-1.05)	0.067*
Energy	1.10 (0.488-2.49)	0.814
Fat	0.622 (0.400-0.968)	0.035*
Free sugar	0.647 (0.431-0.970)	0.035*
Lactose	0.805 (0.676-0.959)	0.015*
Magnesium	1.08 (0.549-2.12)	0.829
N3 fatty acids	0.968 (0.521-1.798)	0.917
N6 fatty acids	1.05 (0.673-1.621)	0.846
Sodium	1.07 (0.453-2.54)	0.871
Total sugars	0.983 (0.506-1.91)	0.960
Total weight food	0.971 (0.528-1.79)	0.924
Vitamin c	0.372 (0.160-0.865)	0.022*

Notes: *P-value <0.05