Association between human blood metabolome and the risk of coronary heart disease: Mendelian randomization study

Jia Zhu¹, Xiaojun Xia¹, Haodong Jiang¹, Congying Wang¹, and Yunpeng Jin¹*

1. Department of Cardiology, The Fourth Affiliated Hospital, Zhejiang University School of Medicine, N1 Shangcheng Road, Yiwu, 322000 Zhejiang, People's Republic of China; 22318337@zju.edu.cn (J.Z.); 22218014@zju.edu.cn (X.-J.X); 22318338@zju.edu.cn(H.-D.J.); congyingwang@zju.edu.cn(C.-Y.W.)

* Corresponding author. Department of Cardiology, The Fourth Affiliated Hospital, Zhejiang University School of Medicine, N1 Shangcheng Road, Yiwu 322000, Zhejiang, People's Republic of China. Tel: +86 18867171112. E-mail: 8013013@zju.edu.cn.

Abstract:
In this study, we employed Mendelian Randomization (MR) to elucidate the causal relationships between specific blood metabolites and Coronary Heart Disease (CHD). By analyzing data from Genome-Wide Association Studies (GWAS) and the FinnGen database, we conducted a two-sample MR analysis focusing on 40 metabolites and 6 metabolite ratios linked to CHD risk. Our findings highlight a group of metabolites significantly influencing CHD risk, either augmenting or mitigating it. Rigorous sensitivity checks, including MR-Egger and MR-PRESSO, negated the influence of horizontal pleiotropy and reinforced the robustness of our results. Furthermore, reverse MR analysis unveiled a bidirectional influence between certain metabolites and CHD, suggesting a complex mutual interaction. This study not only unravels intricate connections between metabolites and CHD, but also paves the way for potential biomarkers instrumental in CHD prevention and therapy. However, it acknowledges certain limitations, such as the modest sample size and a primary focus on European genetic data, underscoring the need for further investigations in more diverse population cohorts.

Introduction:
Coronary Heart Disease (CHD) is a complex chronic inflammatory disorder, characterized predominantly by the narrowing and remodeling of coronary arteries. Despite comprehensive research into the pathogenesis of CHD¹, ², leading to numerous therapeutic approaches for coronary artery disease³, recent epidemiological studies indicate that CHD remains a leading cause of mortality⁴. Therefore, understanding new mechanisms in the development and progression of CHD is crucial for its prevention and treatment⁵-⁶.

Metabolites in the blood, reflective of physiological state changes, are increasingly recognized for their role in the early detection and prognosis of CHD⁷-⁹. However, due to confounding factors and reverse causation, the causal relationship between human metabolites and CHD remains to be fully elucidated.

Employing Mendelian Randomization (MR), a sophisticated epidemiological approach, which offers a more convenient means of eliminating confounders and verifying causal relationships compared to Randomized Controlled Trials (RCTs)¹⁰-¹², this study utilizes two-sample MR analysis. Utilizing the latest Genome-Wide Association Studies (GWAS) and the FinnGen database, this analysis investigates the causal relationship between exposures (metabolites) and outcomes (CHD), providing potential targets for subsequent CHD interventions.

Materials and methods:

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Study design:
In this study, a two-sample Mendelian Randomization (MR) framework is employed to rigorously evaluate the causal links between specific metabolites and Coronary Heart Disease (CHD), as delineated in Figure 1. Within the ambit of MR analysis, the instrumental variables (IVs), constituted by Single Nucleotide Polymorphisms (SNPs), are mandated to fulfill three stringent criteria: 1. The selected SNPs, serving as IVs, must exhibit a robust association with the exposure, evidenced by an F-statistic > 10. 2. These SNPs should demonstrate no association with the outcome variable. 3. The SNPs must remain uncorrelated with any potential confounding variables, thereby ensuring the integrity and validity of the causal inferences drawn.

Figure 1. Overview of this MR study

Genome-wide association study (GWAS) data sources for blood metabolites and CHD:
Genetic variations related to metabolite levels were derived from an extensive Genome-Wide Association Study (GWAS) within the Canadian Longitudinal Study on Aging (CLSA) cohort, analyzing 1091 types of metabolites and 309 metabolite ratios. The GWAS data for CHD originated from the Finnish database (N_case = 43518, N_control = 333759).

Selection of instrumental variables:
In this study, we identified a set of Single Nucleotide Polymorphisms (SNPs) associated with blood metabolites, each surpassing the genome-wide significance threshold (P-value < $1 \times 10^{-5}$), thereby qualifying them as instrumental variables (IVs). This identification was achieved through a comprehensive and sequential filtration process:
Firstly, employing PLINK software (version v1.90), we executed a clumping procedure to refine the SNPs selection. This process adhered to a strict linkage disequilibrium (LD) r^2 threshold of less than 0.001, within a 10,000 kb window, ensuring the minimization of confounding genetic linkage. Secondly, to enhance the robustness of our IVs, we systematically excluded those with weaker statistical relevance, particularly focusing on those with an F-statistic < 10. This stringent selection criterion was pivotal in reinforcing the reliability of our instrumental variables.

Finally, to address and mitigate potential confounders specifically pertaining to Coronary Heart Disease (CHD), an exhaustive search was conducted using the Phenoscanner database (http://www.phenoscanner.medschl.cam.ac.uk/). Any IVs found to be associated with these confounders were rigorously excluded from our analysis, thereby refining our assessment and
ensuring the integrity of our causal inferences

**Statistical analysis**

The association between metabolites and Coronary Heart Disease (CHD) underwent a rigorous evaluation through the application of the 'Mendelian-Randomization' package (Version 0.4.3)\(^2\). Our analysis leveraged advanced statistical methods including inverse variance weighting (IVW), weighted median\(^2\), and mode-based estimation, facilitating a multifaceted assessment. To enhance the reliability of our results, we meticulously computed Cochran's Q statistic\(^2\) accompanied by corresponding p-values, offering a comprehensive quantification of heterogeneity across the employed instrumental variables (IVs).

Concurrently, we utilized the MR-EAGGER methodology\(^2\), renowned for its efficacy in identifying pleiotropy via significant intercepts, to effectively address horizontal pleiotropy within our dataset. We also implemented a leave-one-out analysis to further validate the robustness of our findings. Complementary to these quantitative methods, our study integrated detailed scatter plots and funnel charts, meticulously constructed to provide a visual interpretation of our data, thereby underscoring the validity of our analytical approach\(^2\).

**Reverse MR analysis**

In our analysis, we adopted a more stringent threshold ($p < 5 \times 10^{-12}$) for selecting Single Nucleotide Polymorphisms (SNPs) associated with Coronary Heart Disease (CHD) from the FinnGen database. Subsequently, the filtered Instrumental Variables (IVs) underwent a similar selection process as initially employed for the metabolite-associated SNPs.

**Results**

**Levels of blood metabolites associate with CHD risk**

To investigate the influence of blood metabolites and changes in their ratios on coronary heart disease (CHD), we conducted a two-sample Mendelian Randomization (MR) analysis using the Inverse Variance Weighted (IVW) method as our primary approach. Our study identified 46 metabolites and 7 metabolite ratios impacting CHD risk. After adjusting for the False Discovery Rate (FDR), we excluded results with $\text{PFDR} > 0.05$. This exclusion applied to the Adenosine 5'-monophosphate (AMP) to palmitate (16:0) ratio ($P_{\text{FDR}}=0.0568$), along with levels of 3beta-hydroxy-5-cholestenoate ($P_{\text{FDR}}=0.0513$), Cystathionine ($P_{\text{FDR}}=0.0523$), Decanoylcarnitine (C10) ($P_{\text{FDR}}=0.5$), Methylsuccinoylcarnitine ($P_{\text{FDR}}=0.527$), Ornithine ($P_{\text{FDR}}=0.527$), and 2-hydroxyhippurate (salicylurate) ($P_{\text{FDR}}=0.526$). Ultimately, our analysis confirmed that 40 metabolites and 6 metabolite ratios are significant influencers of CHD risk. Of these, 24 metabolites and 3 metabolite ratios appear to increase CHD risk, while 16 metabolites and 3 ratios seem to offer a protective effect. (Figure 2 and Supplementary Figure 1) The outcomes presented herein have undergone comprehensive global tests through both MR-Egger and MR-PRESSO methodologies, effectively ruling out the presence of horizontal pleiotropy. Additionally, scatter plots and funnel plots have been constructed to demonstrate the stability and robustness of the results.
Figure 2. Forest plots showed the causal associations metabolites and metabolites ratio and CHD Reverse MR analysis assessing the effect of CHD on LEVELS Of blood metabolities.

To investigate the variations in metabolites and their ratios associated with the onset of coronary heart disease (CHD), we again employed a two-sample Mendelian Randomization (MR) analysis. After comprehensive adjustments for the False Discovery Rate (FDR), we found that CHD onset is characterized by increased levels of four specific metabolites: Dihomo-linolenate (20:3n3 or n6) ($P_{FDR}=0.0452$), 2-hydroxyhippurate (salicylate) ($P_{FDR}=0.0035$), Salicylate ($P_{FDR}=0.0018$), and X-15503 ($P_{FDR}=0.0186$).
Discussion

Through an exhaustive analysis of extensive genetic data, this research uncovers bidirectional causal relationships between alterations in blood metabolites and their ratios, and the risk of coronary heart disease (CHD), as elucidated via Mendelian Randomization (MR) analysis. The study discerns 40 metabolites and 6 metabolite ratios significantly correlated with CHD risk (P<0.05). A reverse causation analysis reveals an elevation in four blood metabolites subsequent to CHD onset. Augmented levels of 2-hydroxyhippurate (salicylate), Salicylate, and X-15503 are inversely associated with CHD risk, implying a potential protective role under normative conditions. However, the emergence of CHD seems to instigate a compensatory increase in these metabolites, elucidating a multifaceted interaction. This observation is congruent with prior research indicating stress responses triggered by CHD, leading to elevated metabolite levels. In contrast, an escalation in Dihomo-linolenate (20:3n3 or n6) levels is connected to an increased risk of CHD, with further elevation post-CHD onset, suggesting a bidirectional relationship and hinting at an intricate network of interactions or shared etiological factors.

Employing the latest GWAS summary data, this study eclipses previous research in terms of scale of exposure data (8299 compared to 7828) and a wider range of metabolites, ensuring a more thorough and rigorous analysis. While the selection criteria were exacting, more lenient conditions were adopted for identifying factors influencing CHD (genome-wide significance level of p<1×10^-5, r^2<0.001 within 10000kb). For the reverse causation analysis, owing to the broader conceptualization of CHD relative to CAD and the availability of additional instrumental variables, more stringent criteria were applied (genome-wide significance level of 1×10^-12, r^2<0.001 within 10000kb). The presence of multiple instrumental variables (IVs) for the metabolites necessitated a range of sensitivity analyses, which did not reveal any significant pleiotropy or heterogeneity.

Nonetheless, the study is subject to certain limitations. The relatively modest sample size for the selected metabolites might introduce bias. Furthermore, the exclusive reliance on genetic data from European populations curtails the universality of the findings. Future research is warranted in more diverse populations to enhance the generalizability of these relationships. The insights gleaned from this study pave the way for novel biomarkers and potential targets in CHD prevention and treatment strategies, offering new perspectives in the realm of cardiovascular disease research.

Conclusion

Our research reveals that alterations in 40 metabolites and 6 metabolite ratios can influence the occurrence of coronary heart disease (CHD). These findings suggest a complex network of interactions between them and CHD, offering novel insights for understanding the mechanisms of CHD progression, and paving the way for improved prediction, prevention, and treatment strategies.

Acknowledgements:
This study was possible thanks to publicly available GWAS summary statistics, including those from the GWAS catalog and IEU OpenGWAS project.

**Author contributions:**

J.Z. performed the research, analyzed the data, and wrote the manuscript. X.-J.X. and H.-D.J. drew the figures. C.-Y.W. wrote a part of the manuscript. Y.-P.J. revised, and approved the final manuscript and is the corresponding author. All authors read and approved the final version of the manuscript.

**Funding**

This research received no external funding.

**Data availability**

The summary statistics of SNP-metabolite associations from the CLSA study were available from the NHGRI-EBI GWAS Catalog (https://www.ebi.ac.uk/gwas/) with accession number GCST90199621-90201020. The summary statistics of CHD GWAS were available from the Finngen (https://www.finngen.fi/en)

**References**


