ARTICLE - Coagulation & its Disorders

Two-sample Mendelian randomization analysis reveals causal relationships between blood lipids and venous thromboembolism

Wei Huang,[1,2*] Yan Zou,[2*] Kun Zhang,[2*] Shi Yao,[3] Shi-Hao Tang,[2] Hao Wu,[2] Peng-Fei Wang,[1] Han-Zhong Xue,[1] Tie-Lin Yang,[2] Kun Zhang[1] and Yan Guo[1,2]

[1]Department of Trauma Surgery, Honghui Hospital, College of Medicine, Xi'an Jiaotong University; [2]Key Laboratory of Biomedical Information Engineering of Ministry of Education, Biomedical Informatics & Genomics Center, School of Life Science and Technology, Xi'an Jiaotong University and [3]National and Local Joint Engineering Research Center of Biodiagnosis and Biotherapy, The Second Affiliated Hospital, Xi'an Jiaotong University, Xi'an, Shaanxi, China

WH, YZ and KZ contributed equally as co-first authors.

Correspondence: Y. Guo
guoyan253@xjtu.edu.cn

K. Zhang
hhyyzk@126.com

Received: October 30, 2023.
Accepted: April 22, 2024.
Early view: May 2, 2024.

https://doi.org/10.3324/haematol.2023.284566

Supplementary Material

Two-sample Mendelian randomization analysis reveals causal relationships between blood lipids and venous thromboembolism

Contents

Supplementary Methods

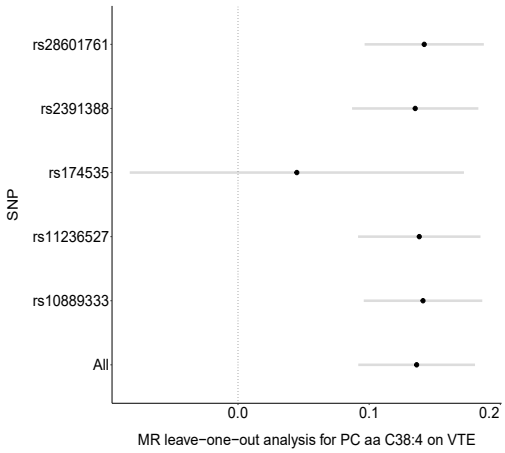
Genetic Instruments Selection and Data Harmonization

Using three-step approaches, we obtained the effective instrumental variables (IVs) of each exposure. Genetic IVs must conform to three hypotheses: (1) have a strong robust correlation with exposure; (2) were independent of confounding associated with exposures and outcomes; (3) only affect the outcome through exposures, but not through other ways[3]. According to these hypotheses, we first selected independent SNPs by clump algorithm module of plink1.9 software[4]. The 1000 Genomes European data was used as reference the panel for linkage disequilibrium (LD) estimation (r^2 threshold = 0.001, window size = 1000kb, *P* value threshold = 5×10^{-8}). Next, we performed a heterogeneity test using the RadialMR package[5] which identified outlier pleiotropic SNPs via modified Q statistics. The threshold for outlier definition is *P* value < 0.05. Finally, we used F-statistics to evaluate the IVs strength for each exposure, while an F-statistic < 10 was considered to be weak intensity[6]. After IVs selection, we harmonized the effect alleles and adjusted β values in the outcome data to make it consistent with the exposure data[1].

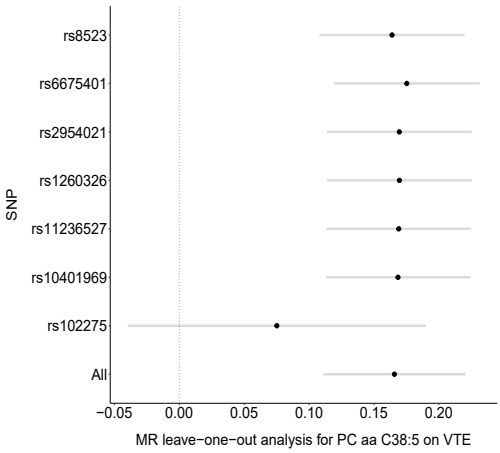
MR Analyses

The IVW method with multiplicative random effects model can be applied to the summary data estimates in the presence of observed heterogeneity[7], which was deemed as the main MR method in our study. The MR-RAPS method is robust to both systematic and idiosyncratic pleiotropy, especially for MR estimation with many weak instruments[8]. It is recommended in cases where exposure and outcomes are both complex traits. MR-Egger method allows all genetic variants to be pleiotropic but requires to be satisfied with the Instrument Strength Independent of Direct Effect (InSIDE) assumption. It assumes that the pleiotropic effect is the same in all variables. This means that pleiotropy leads to bias, but not to additional heterogeneity[7]. The enhancement of the pleiotropy robustness of the MR-Egger method leads to the violation of no measurement error (NOME) in the SNP exposure effects assumption, which can be evaluated by the regression dilution I^2 (GX)[9]. When I^2 (GX) is close to 1, the attenuation due to NOME violation will be negligible. If I^2 (GX) < 0.9, the Simulation Extrapolation (SIMEX) method should be employed to correct this regression dilution bias. Since invalid instrumental variables do not directly affect the median estimate, the weighted median method is able to accurately calculate causal association effects when less than 50% of the genetic variation violates the MR hypothesis[10]. For the weighted mode method, the NOME assumption is not necessary. It relaxed the IV assumption, showing less bias and a lower type I error rate[11].

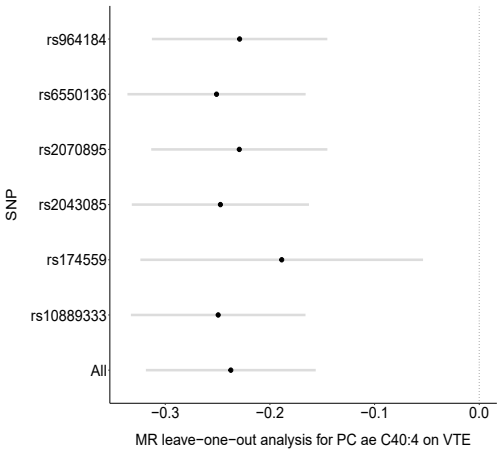
Supplementary Figures



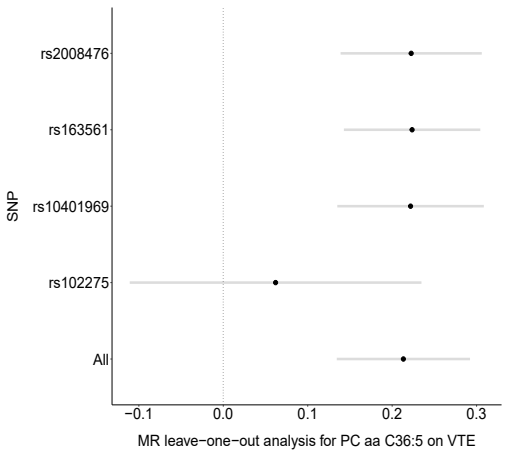
MR leave−one−out analysis for PC aa C38:4 on VTE



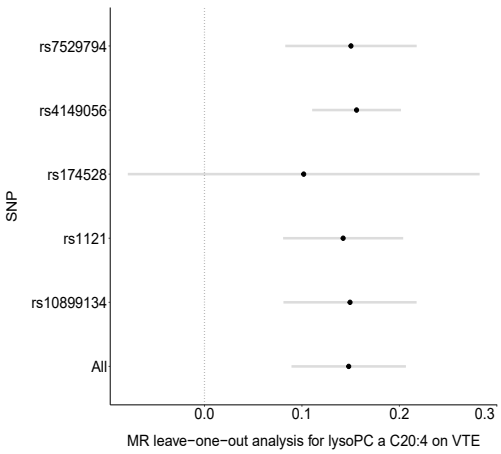
MR leave−one−out analysis for PC aa C38:5 on VTE



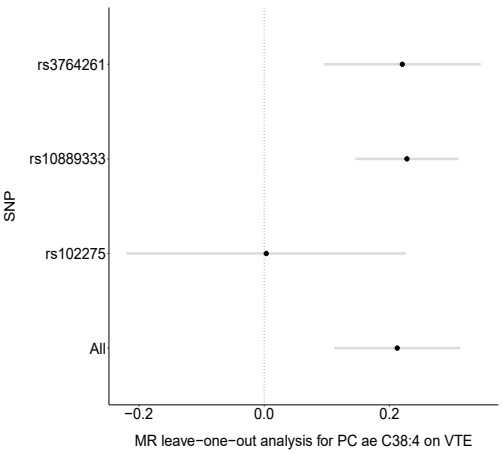
MR leave−one−out analysis for PC ae C40:4 on VTE



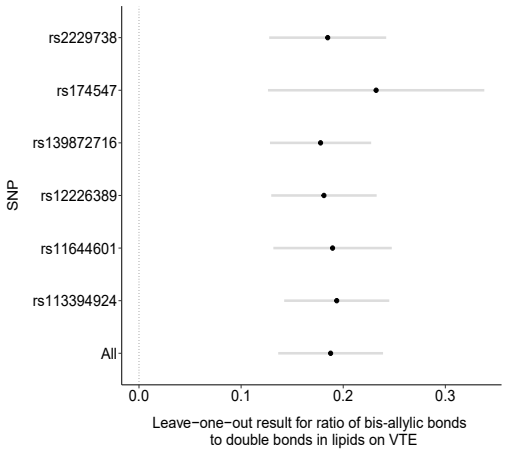
MR leave−one−out analysis for PC aa C36:5 on VTE



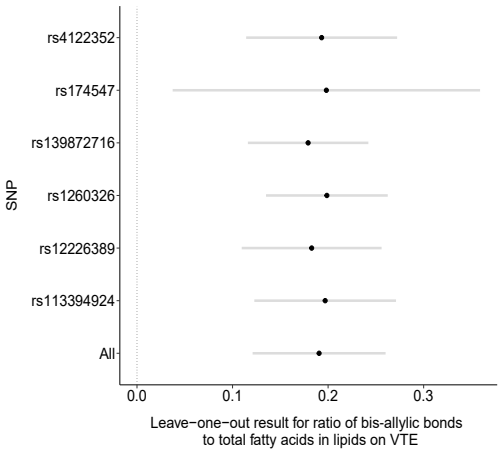
MR leave−one−out analysis for lysoPC a C20:4 on VTE



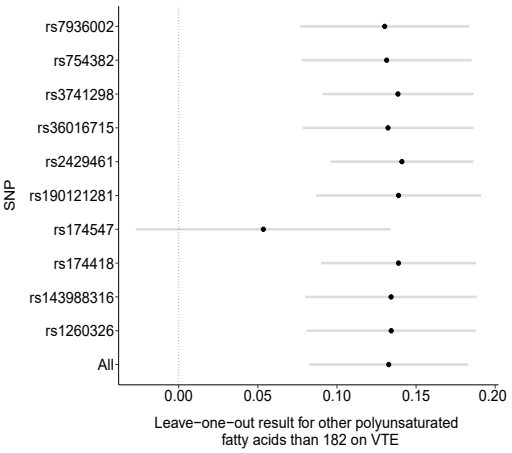
MR leave−one−out analysis for PC ae C38:4 on VTE



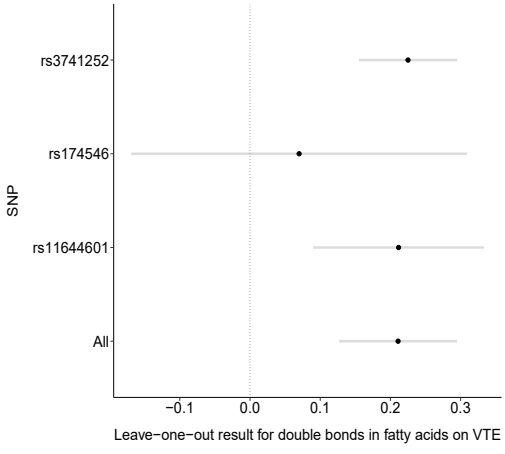
Leave−one−out result for ratio of bis−allylic bonds
to double bonds in lipids on VTE



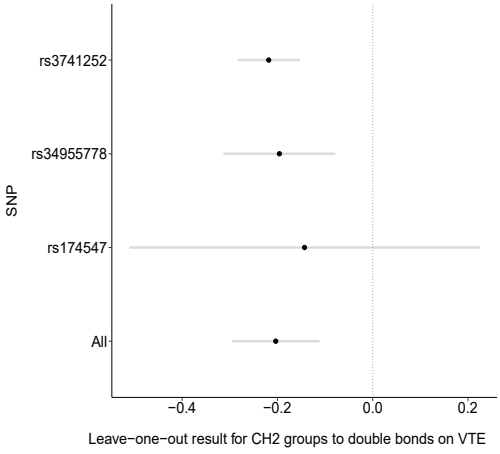
Leave−one−out result for ratio of bis−allylic bonds
to total fatty acids in lipids on VTE



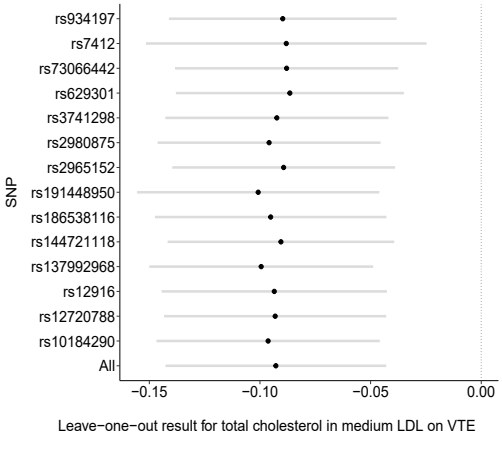
Leave−one−out result for other polyunsaturated
fatty acids than 182 on VTE



Leave−one−out result for double bonds in fatty acids on VTE



Leave−one−out result for CH2 groups to double bonds on VTE



Leave−one−out result for total cholesterol in medium LDL on VTE

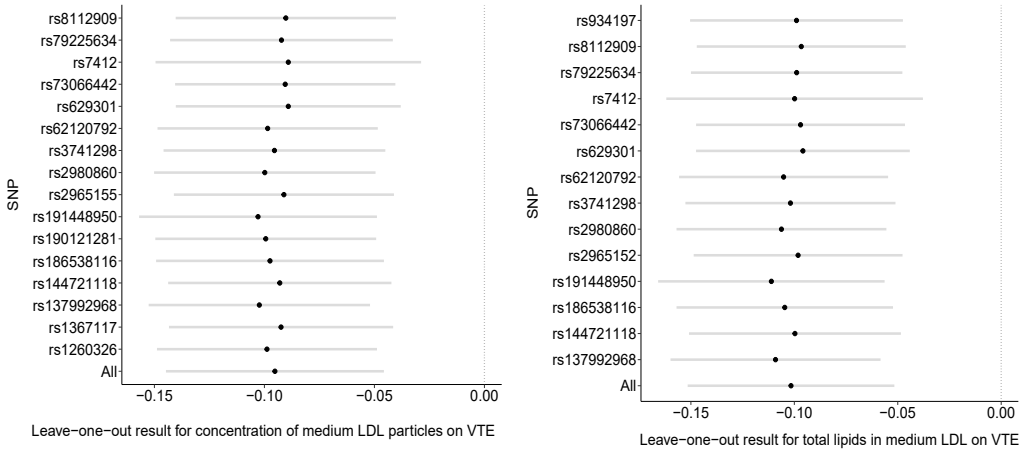


Figure S1. The plot of leave-one-out analysis for 14 significant lipids and lipid-related traits on VTE.

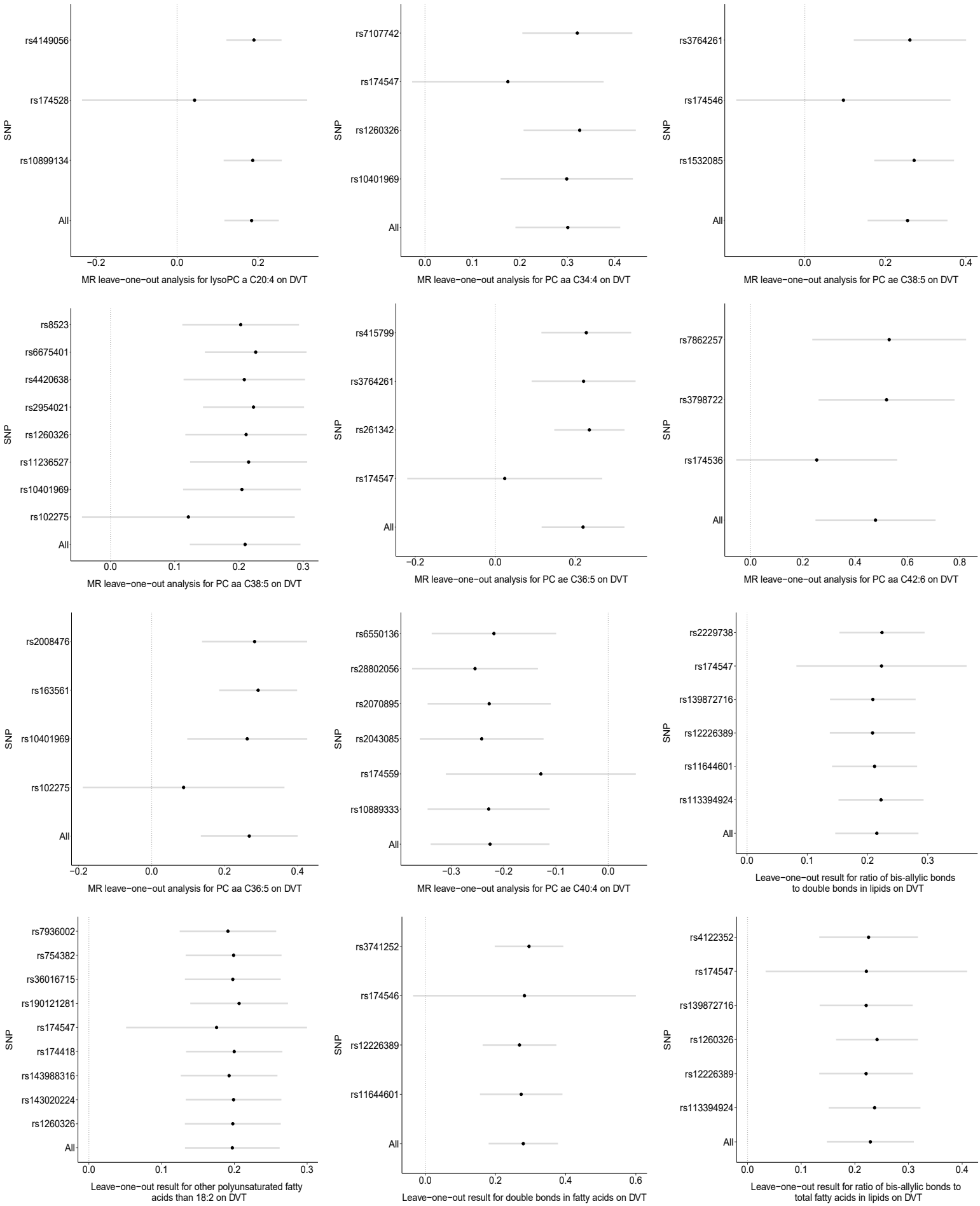


Figure S2. The plot of leave-one-out analysis for 12 significant lipids and lipid-related traits on DVT.

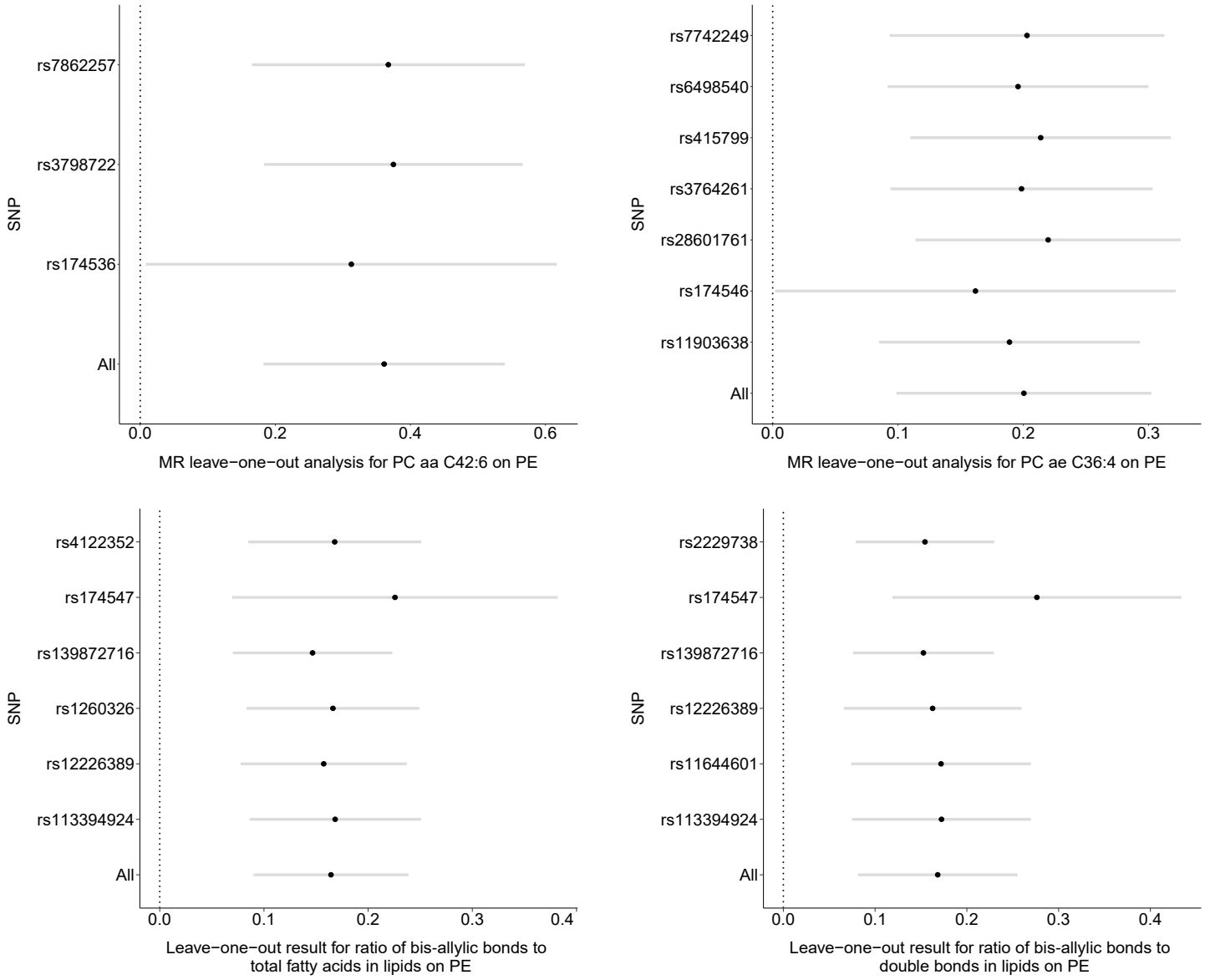


Figure S3. The plot of leave-one-out analysis for 4 significant lipids and lipid-related traits on PE.

Supplementary References

1. Hartwig FP, Davies NM, Hemani G, Smith GD. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int J Epidemiol*. 2016;45:1717-1726. doi: 10.1093/ije/dyx028
2. Hemani G, Zhengn J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7. doi: ARTN e3440810.7554/eLife.34408
3. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7. doi: 10.7554/eLife.34408
4. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559-575. doi: 10.1086/519795
5. Bowden J, Spiller W, Del Greco MF, Sheehan N, Thompson J, Minelli C, Smith GD. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression (vol 47, pg 1264, 2018). *Int J Epidemiol*. 2018;47:2100-2100. doi: 10.1093/ije/dyy265
6. Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol*. 2011;40:755-764. doi: 10.1093/ije/dyr036
7. Bowden J, Del Greco MF, Minelli C, Smith GD, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med*. 2017;36:1783-1802. doi: 10.1002/sim.7221
8. Zhao QY, Wang JS, Hemani G, Bowden J, Small DS. Statistical Inference in Two-Sample Summary-Data Mendelian Randomization Using Robust Adjusted Profile Score. *Ann Stat*. 2020;48:1742-1769. doi: 10.1214/19-Aos1866
9. Bowden J, Del Greco MF, Minelli C, Smith GD, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of thestatistic. *Int J Epidemiol*. 2016;45:1961-1974. doi: 10.1093/ije/dyw220
10. Bowden J, Smith GD, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol*. 2016;40:304-314. doi: 10.1002/gepi.21965
11. Hartwig FP, Smith GD, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*. 2017;46:1985-1998. doi: 10.1093/ije/dyx102