

血脂与非酒精性脂肪性肝病因果关系的 两样本孟德尔随机化研究

申鹏悦, 章文强, 黄德琴, 贺麟, 姜侠, 张本

四川大学华西公共卫生学院/华西第四医院, 四川 成都 610041

摘要:目的 采用两样本孟德尔随机化(Mendelian randomization, MR)方法分析血脂(低密度脂蛋白胆固醇、高密度脂蛋白胆固醇、甘油三酯、载脂蛋白 A-I、载脂蛋白 B)与非酒精性脂肪性肝病(non-alcoholic fatty liver disease, NAFLD)的因果关系。方法 使用欧洲人群中公开的大样本全基因组关联研究汇总数据,运用逆方差加权法、MR-Egger 回归、加权中位数估计法以及 MR 多效性残差和异常值检验等方法进行单变量孟德尔随机化分析,并运用多变量逆方差加权法评价血脂对 NAFLD 的独立效应。结果 单变量逆方差加权法显示,高密度脂蛋白胆固醇($OR = 0.774$, 95% CI : 0.708 ~ 0.845, $P < 0.001$)及载脂蛋白 A-I ($OR = 0.878$, 95% CI : 0.796 ~ 0.968, $P = 0.009$)是 NAFLD 的保护因素,甘油三酯($OR = 1.323$, 95% CI : 1.201 ~ 1.458, $P < 0.001$)是 NAFLD 的危险因素;在调整血脂成分、身体质量指数和 2 型糖尿病后,多变量逆方差加权法发现高密度脂蛋白胆固醇及甘油三酯对 NAFLD 的因果效应保持不变。结论 高密度脂蛋白胆固醇及甘油三酯与 NAFLD 之间存在稳健的因果关系。

关键词:孟德尔随机化;血脂;非酒精性脂肪性肝病;因果推断

中图分类号:R575.5 文献标志码:A 文章编号:1003-8507(2024)12-2146-06

DOI:10.20043/j.cnki.MPM.202309532

Causal relationship between lipids and non-alcoholic fatty liver disease: a two-sample Mendelian randomization study

SHEN Peng-yue, ZHANG Wen-qiang, HUANG De-qin, HE Lin, JIANG Xia, ZHANG Ben

West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, Sichuan 610041, China

Abstract: Objective To evaluate the causal relationships between lipids (low-density lipoprotein cholesterol, high-density lipoprotein cholesterol [HDL-C], triglyceride [TG], apolipoprotein A-I [ApoA-I], apolipoprotein B) and non-alcoholic fatty liver disease (NAFLD) through a two-sample Mendelian randomization (MR) study. **Methods** Summary datasets of genome-wide association studies of lipids and NAFLD in European ancestry were leveraged, and single nucleotide polymorphisms associated with lipids were selected as instrumental variables. Several univariable MR analyses were conducted, including inverse-variance weighting method, MR-Egger regression, weighted median estimation, MR pleiotropy residual sum and outlier method. A multivariable inverse-variance weighting method was further used to evaluate the independent causal effect of lipids on NAFLD. **Results** Univariable inverse-variance weighting MR showed that genetically predicted higher HDL-C ($OR = 0.774$, 95% CI : 0.708 - 0.845, $P < 0.001$) and ApoA-I ($OR = 0.878$, 95% CI : 0.796 - 0.968, $P = 0.009$) were associated with a lower risk of NAFLD, while genetically predicted higher TG ($OR = 1.323$, 95% CI : 1.201 - 1.458, $P < 0.001$) was associated with a higher risk of NAFLD. The causal effects of HDL-C and TG on NAFLD remained consistent in multivariable MR when adjusting for other lipids, body mass index, and type-2 diabetes. **Conclusion** Our study supports a putative causal relationship between HDL-C, TG, and NAFLD.

Keywords: Mendelian randomization; Lipids; Non-alcoholic fatty liver disease; Causal inference

作者简介:申鹏悦(1998—),女,硕士在读,研究方向:流行病与卫生统计学

通信作者:张本, E-mail: wcsphwcfh@vip.163.com; 姜侠, E-mail: xiajiang@scu.edu.cn

非酒精性脂肪性肝病(non-alcoholic fatty liver disease, NAFLD)是目前全球肝脏疾病负担的主要原因之一。据估计,我国将成为 NAFLD 患病率总体和相对增幅最大的国家,NAFLD 病例数将从 2016 年的 2.463 3 亿例增加至 2030 年的 3.145 8 亿例^[1],将对

我国居民健康产生重要影响,一方面反映在医疗成本的升高,如对肝移植需求增加,另一方面出现严重纤维化的 NAFLD 患者发生不良后果的风险显著增加,包括肝脏特异性发病率、死亡率和总体死亡率^[2]。临床上 NAFLD 患者常合并血脂异常^[3-4],观察性研究结果也提示血脂异常与 NAFLD 发生存在关联,但结论存在争议^[5-6]。传统流行病学研究固有的残余混杂和反向因果问题可能会导致先前研究的结果不一致。由于基因在减数分裂时随机分配类似于随机干预,孟德尔随机化(Mendelian Randomization, MR)研究可以克服传统流行病学研究的固有局限^[7]。各脂质表型的 SNPs 通常修饰不止一种脂质表型,其单变量 MR 结果反映的每种暴露对结果的影响包括了直接影响及与通过与其他暴露相互作用产生的影响,而多变量 MR 由于可以估计每种暴露对结果的直接影响而通常用于脂质的 MR 分析^[8-9]。此前对于脂质与非酒精性脂肪性肝病的 MR 研究结果不全面,使用的全基因组关联研究(genome-wide association study, GWAS)汇总数据较少且并未考虑到脂质间及其他代谢性因素多效性的影响^[10]。因此,本研究使

用欧洲人群中公开的大样本全基因组关联研究汇总数据,采用两样本单变量及多变量孟德尔随机化方法分析低密度脂蛋白胆固醇(low-density lipoproteins cholesterol, LDL-C)、高密度脂蛋白胆固醇(high-density lipoproteins cholesterol, HDL-C)、甘油三酯(triglyceride, TG)、载脂蛋白 A-I (apolipoprotein A-I, ApoA-I) 和载脂蛋白 B (apolipoprotein B, ApoB) 等五项常用血脂指标与 NAFLD 的独立因果关系。

1 资料与方法

1.1 资料来源 本研究中 LDL-C, HDL-C, TG, ApoA-I, ApoB 的 GWAS 汇总数据来源于英国生物银行^[11], NAFLD 的 GWAS 汇总数据来源于一项对英国生物银行、爱沙尼亚生物银行以及其他 4 项队列进行的荟萃分析^[12], 总共有 84 334 例 NAFLD 病例及 770 180 名对照。另外考虑到身体质量指数(body mass index, BMI) 及二型糖尿病(type 2 diabetes, T2D) 是可能的混杂因素, 本研究选取了目前最大的 BMI^[13] 和 T2D^[14] 的 GWAS 汇总数据。所有表型的遗传数据均来自欧洲人群。详见表 1。

表 1 血脂与 NAFLD 汇总数据简要信息

Table 1 Information on GWAS summary statistics of lipids and NAFLD

变量	SNP 数量	样本量	人种	性别	发表年份	PMID
LDL-C		440 546				
HDL-C		403 943				
TG	12 321 875	441 016	欧洲人群	男女混合	2020	32 203 549
ApoA-I		393 193				
ApoB		439 214				
NAFLD	6 797 908	778 614	欧洲人群	男女混合	2021	34 841 290
BMI	27 381 302	694 649	欧洲人群	男女混合	2018	30 239 722
T2D	23 596 171	898 130	欧洲人群	男女混合	2018	30 297 969

注:SNP 为单核苷酸多态性。

1.2 研究设计 本研究以血脂为暴露因素, NAFLD 为结局变量, 首先采用单变量孟德尔随机化(univariable MR, UVMR)方法分析血脂与 NAFLD 的因果关联。由于血脂间存在高度的相关性, 本研究采用多变量孟德尔随机化(multivariable MR, MVMR)方法分析血脂的独立效应, 以调整其他血脂的干扰。考虑到多重共线性问题, 本研究没有在 MVMR 分析中同时调整 LDL-C 与 ApoB, HDL-C 与 ApoA-I^[15]; 然后在 MVMR 模型的基础上继续调整了 BMI 及 T2D 这两个混杂因素^[16-17], 以评估血脂对 NAFLD 的独立效应。

1.3 工具变量选择 单变量 MR 分析, 首先参照千人全基因组面板, 选择各个血脂表型中具有显著意义($P < 5.0 \times 10^{-8}$)的 SNPs, 并使用 PLINK clumping 方

法^[18](遗传距离 $< 1\ 000\ \text{kb}$, 遗传连锁参数 $r^2 < 0.001$)筛选与血脂显著相关的独立 SNPs, 然后在此基础上, 使用 steiger 过滤方法^[19]筛选出相比结局与暴露更相关的 SNPs 作为工具变量。MVMR 分析中的模型 1 首先合并了上述方法筛选出的 LDL-C, HDL-C 及 TG 表型的工具变量, 并对这些工具变量再次使用 PLINK clumping 方法进行了处理:遗传距离为 $< 1\ 000\ \text{kb}$, 遗传连锁参数 $r^2 < 0.001$ 。同样的, 模型 2 的表型包括 LDL-C, TG 以及 ApoA-I; 模型 3 的表型包括 HDL-C, TG 以及 ApoB; 模型 4 的表型包括了 TG, ApoA-I 以及 ApoB。通过计算 F 统计量检验是否存在弱工具变量偏倚^[7], $F > 10$ 则说明不存在弱工具变量偏倚, 其计算公式为 $F = \frac{N-K-1}{K} \times \frac{R^2}{1-R^2}$, $R^2 = 2 \times EAF \times (1-EAF) \times \beta^2$, 其中 N 为各血脂表型

的人群样本量, K 为 SNPs 的个数, R^2 为 SNPs 所解释血脂表型的变异比例。

1.4 统计分析 MR 分析以逆方差加权法 (inverse-variance weighted, IVW)^[20] 作为主要分析, 当所选 SNPs 没有水平多效性时 (MR-Egger 截距 $P > 0.05$), IVW 方法被认为是最可信的^[21]。除此之外, 本研究还采用 MR-Egger 回归^[22], 加权中位数估计法 (weighted median estimator, WME)^[23] 以及 MR 多效性残差和异常值检验 (MR pleiotropy residual sum and outlier, MR-PRESSO) 进行了补充分析。区别于 IVW 方法, MR-Egger 方法在回归时考虑了截距项, 可用于纠正工具变量存在水平多效性带来的偏差。WME 可以在存在 50% 的无效工具变量时也能提供较为一致的因果估计。MR-PRESSO 能够识别水平多效性异常的 SNPs, 并提供异常值校正后估计效应。基于 IVW 方法的 MVMR 分析^[24] 用于识别各表型独立的效应。本研究还进行了多项敏感性分析: 去除回文结构的 SNPs, 保证效应等位基因能够正确识别; 根据 GWAS catalog (<https://www.ebi.ac.uk/gwas/>) 去除了与潜在混杂表型相关的 SNPs; 采用留一法 (leave-one-out)^[25] 依次去除 SNP 进行 IVW 分析, 若余下的 SNPs 的 MR 结果仍与总结果保持一致, 说明结果稳健。

本研究应用 Cochran Q 检验判断 SNP 的异质性, 若检验 $P < 0.05$, 则表示 SNPs 之间存在异质性; 另外, 采用 MR-Egger 截距检测水平多效性, 若检验 $P < 0.05$, 则表示存在多效性。

以上分析均使用 R 4.2.2 软件中的“Two Sample MR”“MRPRESSO”“MRInstruments”及“Mendelian Randomization”软件包实现。检验水准 $\alpha = 0.05$ 。

2 结果

2.1 工具变量 根据筛选标准, 成功匹配出 LDL-C, HDL-C, TG, ApoA-I, ApoB 的 SNPs 个数分别为 175, 431, 373, 354, 200 个; R^2 分别为 7.68%, 11.92%, 10.32%, 10.14%, 9.23%; 单一 SNP 的 F 值分布范围分别为 28.60 ~ 3 119.06, 21.79 ~ 3 941.42, 22.46 ~ 2 465.32, 22.46 ~ 2 496.76, 27.94 ~ 2 523.57, 表明不存在弱工具变量。

2.2 孟德尔随机化分析 异质性检验结果显示, 统计量 $Q_{LDL-C} = 409.190$, $Q_{HDL-C} = 636.443$, $Q_{TG} = 632.353$, $Q_{ApoA-I} = 568.530$, $Q_{ApoB} = 440.091$, P 值均 < 0.001 , 提示 SNP 间存在异质性, 故使用随机效应 IVW 模型。

在单变量 MR 分析中, IVW 方法显示 HDL-C ($OR = 0.774$, 95% CI : 0.708 ~ 0.845, $P < 0.001$) 及

ApoA-I ($OR = 0.878$, 95% CI : 0.796 ~ 0.968, $P = 0.009$) 能降低 NAFLD 的发病风险, TG ($OR = 1.323$, 95% CI : 1.201 ~ 1.458, $P < 0.001$) 可增加 NAFLD 的发病风险, MR-Egger、WME 以及 MR-PRESSO 方法均得到与 IVW 一致的结果。由于 MR-Egger 截距检验显示 LDL-C 与 NAFLD 间因果关联受到水平多效性的影响, 故更关注 MR-Egger 结果 ($OR = 0.788$, 95% CI : 0.650 ~ 0.956, $P = 0.016$), 即 LDL-C 可降低 NAFLD 的发病风险, IVW、WME 以及 MR-PRESSO 方法结果均与 MR-Egger 方法方向一致。见图 1。留一法逐个剔除各 SNP 后, 未发现对结果有较大影响的 SNP (见图 2), MR 分析结果是稳健的。

从未调整 BMI、T2D 的 MVMR 分析的 4 个模型中 (见图 3) 可以看出, LDL-C ($OR = 0.870$, 95% CI : 0.777 ~ 0.975, $P = 0.016$; $OR = 0.875$, 95% CI : 0.783 ~ 0.978, $P = 0.018$)、HDL-C ($OR = 0.875$, 95% CI : 0.785 ~ 0.977, $P = 0.017$; $OR = 0.852$, 95% CI : 0.765 ~ 0.949, $P = 0.004$) 可能降低 NAFLD 的发病风险, 而 TG ($OR = 1.296$, 95% CI : 1.148 ~ 1.463, $P < 0.001$; $OR = 1.365$, 95% CI : 1.226 ~ 1.519, $P < 0.001$; $OR = 1.266$, 95% CI : 1.121 ~ 1.430, $P < 0.001$; $OR = 1.341$, 95% CI : 1.202 ~ 1.496, $P < 0.001$) 可能增加 NAFLD 的发病风险。

在调整 BMI、T2D 后的 MVMR 分析的 4 个模型 (见图 4) 显示, HDL-C ($OR = 0.866$, 95% CI : 0.775 ~ 0.967, $P = 0.011$; $OR = 0.853$, 95% CI : 0.763 ~ 0.952, $P = 0.005$) 可能降低 NAFLD 的发病风险, 而 TG ($OR = 1.234$, 95% CI : 1.093 ~ 1.393, $P = 0.001$; $OR = 1.317$, 95% CI : 1.183 ~ 1.466, $P < 0.001$; $OR = 1.212$, 95% CI : 1.073 ~ 1.370, $P = 0.002$; $OR = 1.283$, 95% CI : 1.149 ~ 1.433, $P < 0.001$) 可能增加 NAFLD 的发病风险。

3 讨论

本研究利用欧洲人群的大样本 GWAS 汇总数据, 运用单变量孟德尔随机化和多变量孟德尔随机化方法系统分析了 LDL-C、HDL-C、TG、ApoA-I、ApoB 五项血脂指标与 NAFLD 之间的因果关系。此前 MR 研究探讨了 LDL-C、TC、TG 与 NAFLD 之间的因果关系, 并未发现因果关联^[10], 而本研究单变量 MR 发现 TG 是 NAFLD 的危险因素, 并额外发现了 HDL-C、ApoA-I 是 NAFLD 的保护因素, 这可能是因为本研究采用的 GWAS 汇总数据样本量更大, SNPs 数量更多。另外, 本研究为探讨各脂质指标与 NAFLD 之间的独立关系采用了多变量 MR, 结果进一步支持了 HDL-C 和 TG 对 NAFLD 的因果效应。研究提示在

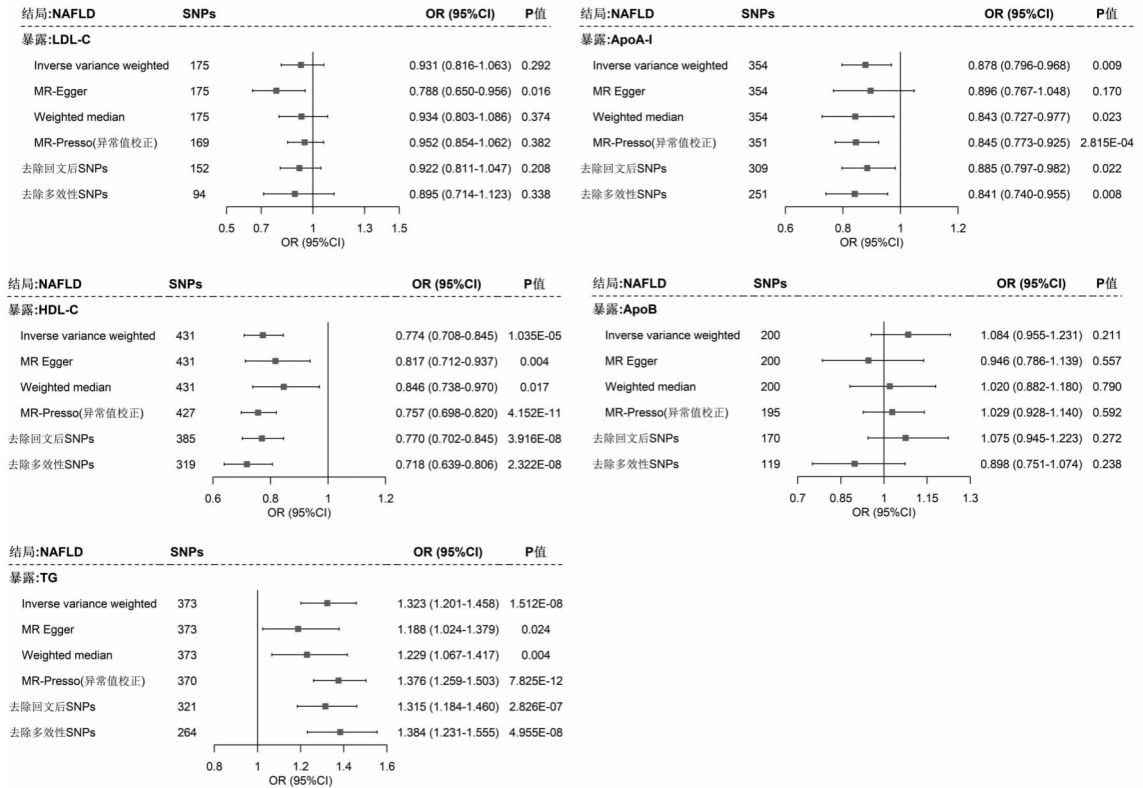
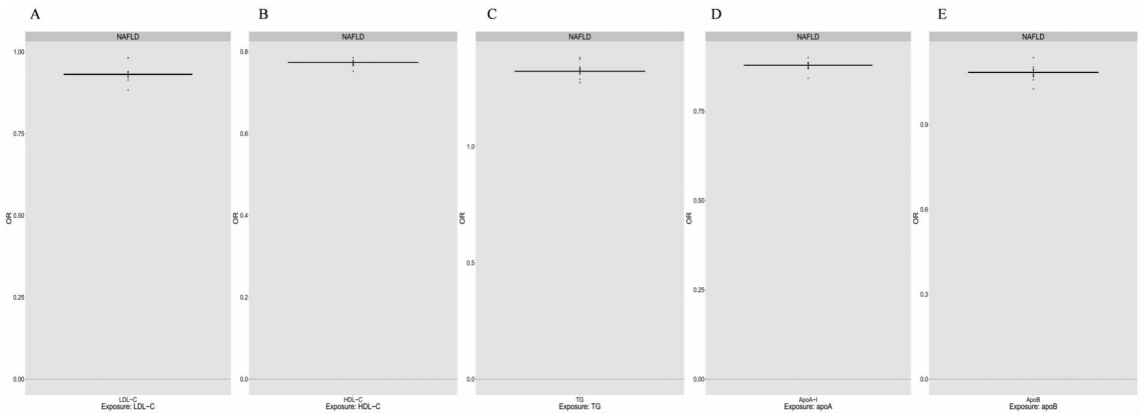


图 1 单变量 MR 森林图
Fig. 1 The forest plot of univariable MR analysis



注:图 A、B、C、D、E 分别表示 LDL - C、HDL - C、TG、ApoA - I、ApoB。

图 2 “leave - one - out” 敏感性分析图

Fig. 2 The “leave - one - out” sensitivity analysis plot

高密度脂蛋白胆固醇或甘油三酯异常人群中开展 NAFLD 筛查以及生活方式干预可能是有益的。

既往观察性研究中血脂各指标与 NAFLD 之间的关系结果并不一致,一项针对中国人群的横断面研究仅发现 HDL - C 是 NAFLD 的独立保护因素 ($OR = 0.35, 95\% CI: 0.16 \sim 0.76$)^[5],而另一项更大的纳入 110 626 名体检人群的横断面研究发现高 TG ($OR = 2.77, 95\% CI: 2.67 \sim 2.89$)、低 HDL - C ($OR = 1.64, 95\% CI: 1.51 \sim 1.79$)、高 LDL - C ($OR = 1.35, 95\%$

$CI: 1.28 \sim 1.43$)可能会增加 NAFLD 的发生风险^[6]。在分层分析中,一项包含 2 058 名参加体检项目人群的横断面研究中观察到仅高甘油三酯是非肥胖人群 NAFLD 的独立危险因素 ($OR = 1.701, 95\% CI: 1.085 \sim 2.666$)^[26],而另一项回顾性研究发现仅 TG 与 NAFLD 显著相关 ($z = 7.791, P < 0.001$),而在对 BMI 进行分层分析后发现低 HDL - C 是超重和肥胖人群患者 NAFLD 的危险因素 ($z = -2.794, P = 0.005$),高 LDL - C 增加了非肥胖人群发生 NAFLD

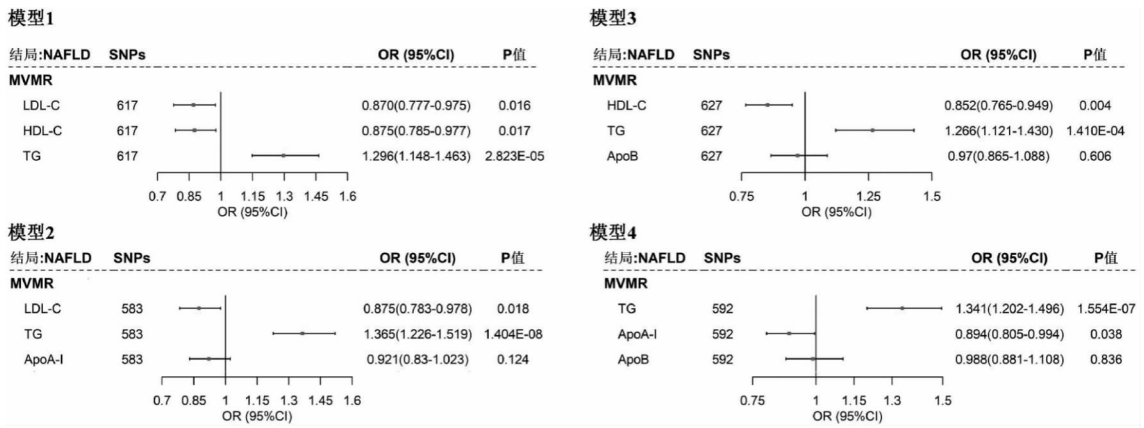


图 3 多变量 MR 森林图
Fig. 3 The forest plot of multivariable MR analysis

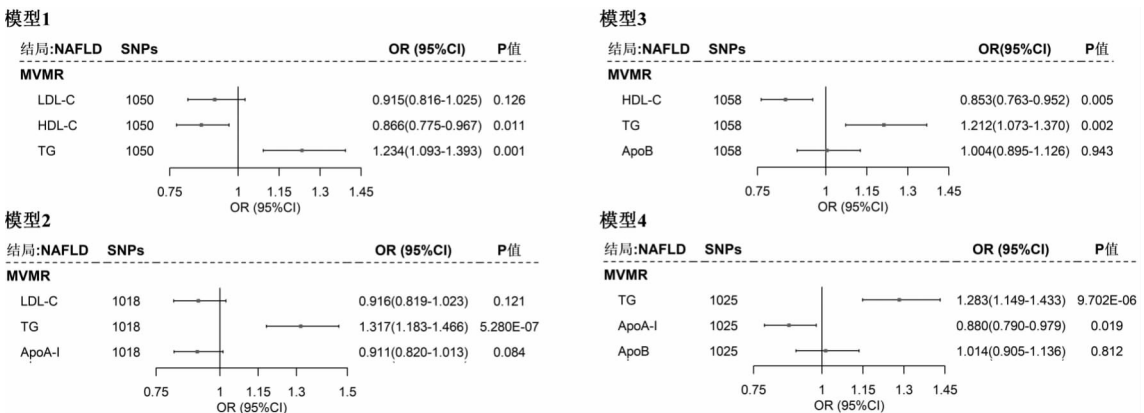


图 4 血脂表型与 NAFLD 因果关联,多变量 MR 森林图(调整 BMI、T2D)
Fig. 4 The forest plot of multivariable MR analysis (after adjusting for BMI、T2D)

的风险($z=2.985, P=0.003$)^[27]。虽然高密度脂蛋白胆固醇和甘油三酯与 NAFLD 关系的机制仍尚不清楚,但是目前多条证据支持了起源于两次打击假说^[28]的脂毒性模型,即甘油三酯通过促进脂毒素如游离脂肪酸、溶血磷脂酸、溶血磷脂酰胆碱等的产生间接促进肝细胞损伤^[29],本研究的结果也支持了这一观点。总体而言,以上观察性研究基本与本研究结果一致。

本研究具有以下优点:(1)本研究建立了更严格的标准来筛选工具变量,使用了多种因果效应估计方法相互验证,一致效应表明了结果的稳健性;(2)为了提高估计的准确性,本研究使用了多变量孟德尔随机化方法并考虑到了 BMI 及 T2D 的混杂影响。

同时本研究也存在局限性:(1)本研究使用的 GWAS 数据仅包括欧洲人群,故研究结果推至其他人群存在局限性;(2)NAFLD 有多种亚型,亚型分析的结果是否与本研究一致还需要进一步的探索。

综上,本研究通过单变量及多变量孟德尔随机化分析,支持 HDL-C 是 NAFLD 的保护因素,而 TG 是

NAFLD 的危险因素。在高密度脂蛋白胆固醇及甘油三酯人群异常人群中开展 NAFLD 的早期筛查和早期干预,可以降低 NAFLD 的疾病负担。

利益冲突声明 本研究不存在任何利益冲突

参考文献

- [1] Estes C, Anstee QM, Arias -Loste MT, et al. Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016 - 2030 [J]. Journal of Hepatology, 2018, 69(4): 896 - 904.
- [2] Cotter TG, Rinella M. Nonalcoholic fatty liver disease 2020: the state of the disease[J]. Gastroenterology, 2020, 158(7): 1851 - 1864.
- [3] 姚诚子,冯巩,宇文思,等.非酒精性脂肪性肝病发生及进展的危险因素[J].临床肝胆病杂志,2020,36(2):433 - 436. Yao CZ, Feng G, Yu Wen S, et al. Risk factors for the occurrence and progression of non - alcoholic fatty liver disease[J]. Journal of Clinical Hepatology, 2020, 36(2): 433 - 436.
- [4] 高欢.非酒精性脂肪性肝病的患病率及相关危险因素分析[D].大连:大连医科大学,2019. Gao H. Analysis of prevalence and related risk factors of

- nonalcoholic fatty liver diseases [D]. Dalian: Dalian Medical University, 2019.
- [5] Wu J, He SM, Xu HQ, et al. Non-alcoholic fatty liver disease incidence, remission and risk factors among a general Chinese population with a 6-year follow-up [J]. *Scientific Reports*, 2018, 8(1): 7557.
- [6] Kong LX, Yang Y, Li HD, et al. Prevalence of nonalcoholic fatty liver disease and the related risk factors among healthy adults: A cross-sectional study in Chongqing, China [J]. *Frontiers in Public Health*, 2023, 11: 1127489.
- [7] Burgess S, Thompson SG, CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies [J]. *International Journal of Epidemiology*, 2011, 40(3): 755-764.
- [8] Yu Z, Zhang LJ, Zhang G, et al. Lipids, apolipoproteins, statins, and intracerebral hemorrhage: a Mendelian randomization study [J]. *Annals of Neurology*, 2022, 92(3): 390-399.
- [9] Thomas DG, Wei Y, Tall AR. Lipid and metabolic syndrome traits in coronary artery disease: a Mendelian randomization study [J]. *Journal of Lipid Research*, 2021, 62: 100044.
- [10] Li ZA, Zhang B, Liu QR, et al. Genetic association of lipids and lipid-lowering drug target genes with non-alcoholic fatty liver disease [J]. *EBioMedicine*, 2023, 90: 104543.
- [11] Richardson TG, Sanderson E, Palmer TM, et al. Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis [J]. *PLOS Medicine*, 2020, 17(3): e1003062.
- [12] Ghodssian N, Abner E, Emdin CA, et al. Electronic health record-based genome-wide meta-analysis provides insights on the genetic architecture of non-alcoholic fatty liver disease [J]. *Cell Reports Medicine*, 2021, 2(11): 100437.
- [13] Pulit SL, Stoneman C, Morris AP, et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry [J]. *Human Molecular Genetics*, 2019, 28(1): 166-174.
- [14] Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps [J]. *Nature Genetics*, 2018, 50(11): 1505-1513.
- [15] Li N, Zhang XY, Zhang M, et al. Associations of genetically determined lipid traits and lipid-modifying agents with the risk of diabetic retinopathy: A Mendelian randomization study [J]. *Atherosclerosis*, 2023, 369: 9-16.
- [16] Polyzos SA, Kountouras J, Mantzoros CS. Adipose tissue, obesity and non-alcoholic fatty liver disease [J]. *Minerva Endocrinologica*, 2017, 42(2): 92-108.
- [17] Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of non-alcoholic fatty liver disease: Practice guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association [J]. *Hepatology*, 2012, 55(6): 2005-2023.
- [18] Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome [J]. *eLife*, 2018, 7: e34408.
- [19] Hemani G, Tilling K, Davey smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data [J]. *PLOS Genetics*, 2017, 13(11): e1007081.
- [20] Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data [J]. *Genetic Epidemiology*, 2013, 37(7): 658-665.
- [21] Holmes MV, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality [J]. *Nature Reviews Cardiology*, 2017, 14(10): 577-590.
- [22] Bowden J, Davey smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression [J]. *International Journal of Epidemiology*, 2015, 44(2): 512-525.
- [23] Bowden J, Davey smith G, Haycock PC, et al. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator [J]. *Genetic Epidemiology*, 2016, 40(4): 304-314.
- [24] Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects [J]. *American Journal of Epidemiology*, 2015, 181(4): 251-260.
- [25] Gronau QF, Wagenmakers EJ. Limitations of Bayesian Leave-One-Out Cross-Validation for model selection [J]. *Computational Brain & Behavior*, 2019, 2(1): 1-11.
- [26] Cho HC. Prevalence and factors associated with nonalcoholic fatty liver disease in a nonobese Korean population [J]. *Gut and Liver*, 2016, 10(1): 117-125.
- [27] Tang ZL, Pham M, Hao YM, et al. Sex, age, and BMI modulate the association of physical examinations and blood biochemistry parameters and NAFLD: A retrospective study on 1994 cases observed at Shuguang Hospital, China [J]. *BioMed Research International*, 2019, 2019: 1246518.
- [28] Day CP, James OFW. Steatohepatitis: a tale of two "hits"? [J]. *Gastroenterology*, 1998, 114(4): 842-845.
- [29] Semova I, Biddinger SB. Triglycerides in nonalcoholic fatty liver disease: guilty until proven innocent [J]. *Trends in Pharmacological Sciences*, 2021, 42(3): 183-190.