资料和方法

研究设计和数据来源

采用两样本 MＲ 分析来研究 ＲC 与 CHD 和 MI 的潜在因果关系。本研究中使用的数据全部来自 公共数据，这些数据在数据库网站上公开获得，具 体的伦理批准在原始的全基因组关联分析( genomewide association study，GWAS) 文章中已有表述。本 研究遵循了 MＲ 研究的三个假设: ( 1) 遗传变异与 危险因素密切相关( 本研究中的 ＲC) ; ( 2) 遗传变异 与混杂因素无关; ( 3) 遗传变异仅通过危险因素影 响结局( 本研究中的心血管疾病) ［13］。使用 Ｒ 程序 ( 版本 4． 1． 0) 的 Two Sample MＲ 包进行统计分析。 从 IEU OpenGWAS project( https: / /gwas． mrcieu． ac． uk /) 网站获得 ＲC( met-d-Ｒemnant\_C) 、CHD( ieua-7) 和 MI( ieu-a-798) 的 GWAS 数据( 表 1) ，网站访 问时间为 2022-10-18。ＲC 的汇总数据包括 115 078 例参与者，SNP 数量为 12 321 875 个，均来自欧洲人 群，男女不限。CHD 和 MI 的汇总数据来自 CAＲDIoGＲAMplusC4D 联盟的 GWAS 荟萃分析，该研究 包括多达 184 305 例参与者( 77%为欧洲血统; 60 801 例 CHD 和 123 504 例对照样本，SNP 数量为 9 455 779 个; 43 676 例 MI 和 128 199 例对照样本， SNP 数量为 9 289 492 个) ［14］。

工具变量

工具变量 从 ＲC 的 GWAS 汇总数据中筛选出有意义的 SNP( 筛选条件: P＜5E－08，连锁不平衡 r 2 为 0. 001， 连锁不平衡区域宽度为 10 000 kb，以保证各个 SNP 相互独立，排除基因多效性对结果的影响) 。我们 计算了 F 统计量以量化遗传变异的强度，F＜10 为 弱工具变量，须剔除。F 统计量计算公式为: F = N－K－1 K × Ｒ2 1－Ｒ2，其中 N 为暴露数据库的样本量，K 为 SNP 的个数，Ｒ2 为暴露数据库中由 SNP 解释的变异 所占的比例。Ｒ2 计算 公 式 为: Ｒ2 = 2 × MAF × ( 1 － MAF) ×β2 ，其中 MAF 为次要等位基因频率，β 为等 位基因效应值。然后从 CHD 和 MI 的 GWAS 汇总 数据中提取上述筛选出的 ＲC 相 关 SNP; 缺 失 的 SNP 用与其具有高度连锁性( r 2 ＞0. 8) 的 SNP 代替， 剔除无合适替代位点的 SNP。分别汇总 ＲC 与 CHD 及 ＲC 与 MI 数据集的信

统计学分析

本研究以逆方差加权法( inverse variance weighted，IVW) ［16］为主要分析方法，该方法结合了 Wald 比率估计值，以获得暴露对结果因果影响的一致估 计，其特点是回归时不考虑截距项的存在。本文还 使用了 其 他 的 MＲ 分 析 方 法，包 括 MＲ-Egger 回 归［17］、加 权 中 位 数 法 ( weight median estimator， WME) ［18］、简单模 式( simple mode) ［19］、加 权 模 型weighted mode) ［19］来验证暴露( ＲC) 与结局( CHD) 的因果关系。MＲ-Egger 回归法与 IVW 分析方法最 大的区别就是加多了截距项，主要用来判断有无水 平多效性。WME 是假设分析中至少有一半的 IV 有 效，然后获得因果关系的一致估计。简单模式根据 因果效应对 SNP 进行分类，并将相似值划分为一个 簇，使用 SNP 数量最多的簇来评估估计的因果效 应。加权模式将每个集群中 SNP 的因果效应值与 模块 SNP 数量进行加权，返回的结果是具有最大 SNP 数权重的临时估计。使用加权模式的一致性 因果效应估计要求满足“ZEMPA 假设”( 零模态多 效性假设)

1． 4 敏感性分析 本研究使用 Cochran’s Q 检验来判断 IVW 模型 的异质性; 当 P＜0. 05 时认为存在异质性，此时使用 IVW 的随机效应模型来进行因果推断。用 MＲ-Egger 回归分析来检测基因多效性是否存在，当截距项 与 0 差异很大时，说明存在水平多效性。最后进行 逐个剔除检验，评估合并的 IVW 估计值是否由任何 单个 SNP 所影响。

2 结 果

2． 1 工具变量

met-d-Ｒemnant\_C 中共筛选出 56 个与 ＲC 高 度相关的 SNP。经过几次筛选，CHD 数据集获得 52 个 SNP，剔除与结局相关的 4 个 SNP ( rs12740374、 rs12151108、rs115478735 和 rs653178) ，同时经 MＲPＲESSO 分 析 后 剔 除 3 个 离 群 值 ( rs1500188、 rs3777411 和 rs77960347) ，最终纳入 45 个 SNP。Ｒ2 为 3. 73%，单一 SNP 对应的 F 统计量分布范围为 24. 50～377. 72，表明在 MＲ 分析中没有使用弱工具 变量。MI 数据集获得 52 个 SNP，剔除与结局相关 的 3 个 SNP( rs12740374、rs115478735 和 rs653178) ， 同时 经 MＲ-PＲESSO 分 析 后 剔 除 2 个 离 群 值 ( rs1500188 和 rs3777411) ，最终纳入 47 个 SNP。Ｒ2 为 4. 28%，单一 SNP 对应的 F 统计量分布范围为 24. 50～570. 33，表明在 MＲ 分析中没有使用弱工具 变量。SNP 基本信息见表 2。

2． 2 ＲC 与 CVD 的孟德尔随机化分析 遗传预测的 ＲC 升高与 CHD 和 MI 风险增加 有关，五种方法得到的因果效应方向一致( 表 3、图 1 和图 2) 。IVW 法分析结果显示，ＲC 与 CVD 发 生风险之间关联具有统计学 意 义 ( CHD: OＲ = 1. 57，95%CI: 1. 40 ～ 1. 76，P = 2. 01E－ 14; MI: OＲ = 1. 59，95%CI: 1. 40～1. 79，P = 1. 26E－13) 。

2． 3 敏感性分析 漏斗图显示纳入的所有 SNP 基本对称( 图 3) ， 表明工具变量之间差异较小。虽然 Cochran’s Q 检 验( P＜0. 05) 发现存在异质性，但随机效应 IVW 分 析方法的结果表明 ＲC 与 CHD( P = 2. 01E－ 14) 和 MI( P = 1. 26E－13) 确实存在因果关系，并且 ＲC 升 高后 CHD( b = 0. 45) 和 MI( b = 0. 46) 发病风险也相 应增加。MＲ-Egger 回归结果提示筛选出的 SNP 与 CHD 和 MI 不存在基因多效性( CHD 和 MI 的 P 值 分别为 0. 924 1、0. 740 5) 。经逐个剔除检验后，CHD 和 MI 分别余下的 45 和 47 个 SNP 都在无效线的右 侧，表明本研究的 MＲ 分析结果是稳健的( 图 4) 。

本研究采用两样本 MＲ 方法来评估 ＲC 与 CHD 和 MI 风险增加之间的因果关系，得到了 ＲC 与 CHD 和 MI 风险增加之间因果关系的有力证据。具体来 说，因果推断得到了各种 MＲ 方法( 包括 IVW、MＲEgger、加权中位数法、简单模型和加权模型) 中一致 的效应估计值和方向的支持，这将有助于我们更好 地理解 ＲC 对 CVD 的遗传影响。本研究发现 ＲC 每 增加 一 个 标 准 差，CHD 风 险 将 增 加 57% ( OＲ = 1. 57，95%CI: 1. 40 ～ 1. 76，P = 2. 01E－14) ，MI 风险 将增加 59% ( OＲ = 1. 59，95% CI: 1. 40 ～ 1. 79，P = 1. 26E－13) 。本研究结果和既往的研究结果一致。 针对哥本哈根的 73 513 例丹麦血统的白人受试者 进行的研究发现，非空腹残余胆固醇增加 1 mmol /L 与缺血性心脏病的 2. 8 倍因果风险相关，并独立于 HDLC，这意味着富含甘油三酯的脂蛋白颗粒的胆 固醇含 量 升 高 会 导 致 缺 血 性 心 脏 病［20］。Quispe 等［21］评估了非 ASCVD 患者中 LDLC 和载脂蛋白 B ( apolipoprotein B，ApoB) 以外的 ＲC 相关风险后发 现，经过包括 LDLC 和 ApoB 在内的多变量调整后， logＲC 与较高的 ASCVD 风险相关( HＲ 1. 65，95%CI: 1. 45～1. 89) ，即在无 ASCVD 个体中，升高的 ＲC 水 平与 ASCVD 相关，独立于传统的危险因素，如 LDLC 和 ApoB 水平。Doi 等［22］使用预定义的 ＲC 升高 的切点计算了从低于 5%、7. 5%和( 或) 10%的 10 年 MI 和缺血性心脏病发生的净重分类指数( net reclassification index，NＲI) ，该指数定义为缺血性心脏 病、MI 和冠状动脉血运重建的复合死亡。结果发现 ＲC 水平升高可显著改善 MI 和缺血性心脏病的风 险预测。总体而言，降低残余胆固醇水平可能会减 少冠状动脉粥样硬化和 CVD 发生风险，并且需要进 一步研究靶向降低 ＲC 在一级预防中的潜在价值。 从机制上讲，ＲC 可能具有促炎［23］、促进动脉粥 样硬化和促血栓形成的作用，这都增加 CHD 和 MI 的风险。ＲC 携带更大的胆固醇负荷，在动脉内膜 中积累，比 LDL 更有效地诱导巨噬细胞泡沫化，促 进泡沫细胞的形成和动脉粥样硬化斑块的炎症变 化［24］。有研究发现，内皮表面 ＲC 脂解过程中释放 的游离脂肪酸也可以对内皮细胞和巨噬细胞发挥 促炎作用［25-26］。此外，低水平的 HDL 可能与泡沫细 胞胆固 醇 流 出 减 少 有 关，从而加剧动脉粥样硬 化［27］。Varbo 等［9］发现 ＲC 水平升高与低度炎症相 关，而 LDL 与低度炎症关系不大，动脉粥样硬化的 炎症程度主要由 ＲC 所驱动。Elshazly 等［28］分析了 5 754 例接受连续血管内超声检查的 CHD 患者的数 据后发现，在他汀类药物治疗的动脉粥样硬化性心 血管疾病患者中，残余胆固醇与冠状动脉粥样硬化 进展相关且独立于传统的血脂指标、C 反应蛋白或 临床危险因素，并且较高的残余胆固醇水平也与较 高的主要不良心血管事件相关。ＲC 能上调纤溶酶 原激活物抑制剂 1 基因和纤溶酶原激活物抑制剂 1 抗原和活性的表达，促进血小板聚集和凝块形成; 活化的血小板的表面膜还支持凝血酶原复合物的 组装和活性，从而进一步产生凝血酶并扩增凝血级 联反应［29］。所有这些过程都有助于形成不稳定的 斑块、斑块侵蚀和血栓形成［30］，从而导致心血管疾 病死亡。

本研究有几个优势。首先，MＲ 设计研究可以 模拟观察环境中的随机对照试验。相较于随机对 照试验，更能节省时间和成本，这使得孟德尔随机 化分析在评估和筛选潜在因果关联方面具有一定 的优势; 并且 MＲ 研究降低了混杂偏倚以及避免了 反向因果关系。其次，本研究采用严格的质量控制 条件和合理的分析方法，包括运用多种模型来评价 因果效应。因此，本研究的结果是可靠和稳定的。 第三，我们使用的数据来自大型 GWAS，为高统计功 效提供了足够的样本量。 同时，本研究也存在一些局 限。首 先，所 有 GWAS 数据都主要来自欧洲人群，本研究的发现不 能适用于其他人群。其次，与性状相关的遗传变异 仅解释了危险因素变异的一小部分，且这种关系是 理论上的，仍须进一步在随机对照试验中来证实。 综上所述，本研究应用两样本 MＲ 分析方法，以 基因变异为工具变量对 ＲC 与 CHD 和 MI 患病风险 进行了探讨，提供了 ＲC 升高对 CHD 和 MI 风险增 加的因果关系证据，提示检测血浆 ＲC 水平对 ASCVD 评估及剩留风险管控方面具有重要价值。