Mendelian randomization studies in coronary artery disease

Henning Jansen1,2, Nilesh J. Samani3,4, and Heribert Schunkert1,2*

1Deutsches Herzzentrum München and Technische Universität München, Munich, Germany; 2DZHK (German Research Centre for Cardiovascular Research), Munich Heart Alliance, Munich, Germany; 3Department of Cardiovascular Sciences, University of Leicester, Leicester, UK; and 4National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, UK

Received 30 October 2013; revised 11 April 2014; accepted 30 April 2014; online publish-ahead-of-print 10 June 2014

Epidemiological research over the last 50 years has discovered a plethora of biomarkers (including molecules, traits or other diseases) that associate with coronary artery disease (CAD) risk. Even the strongest association detected in such observational research precludes drawing conclusions about the causality underlying the relationship between biomarker and disease. Mendelian randomization (MR) studies can shed light on the causality of associations, i.e. whether, on the one hand, the biomarker contributes to the development of disease or, on the other hand, the observed association is confounded by unrecognized exogenous factors or due to reverse causation, i.e. due to the fact that prevalent disease affects the level of the biomarker. However, conclusions from a MR study are based on a number of important assumptions. A prerequisite for such studies is that the genetic variant employed affects significantly the biomarker under investigation but has no effect on other phenotypes that might confound the association between the biomarker and disease. If this biomarker is a true causal risk factor for CAD, genotypes of the variant should associate with CAD risk in the direction predicted by the association of the biomarker with CAD. Given a random distribution of exogenous factors in individuals carrying respective genotypes, groups represented by the genotypes are highly similar except for the biomarker of interest. Thus, the genetic variant converts into an unconfounded surrogate of the respective biomarker. This scenario is nicely exemplified for LDL cholesterol. Almost every genotype found to increase LDL cholesterol level by a sufficient amount has also been found to increase CAD risk. Pending a number of conditions that needed to be fulfilled by the genetic variant under investigation (e.g. no pleiotropic effects) and the experimental set-up of the study, LDL cholesterol can be assumed to act as the functional component that links genotypes and CAD risk and, more importantly, it can be assumed that any modulation of LDL cholesterol—by whatever mechanism—would have similar effects on disease risk. Therefore, MR analysis has tremendous potential for identifying therapeutic targets that are likely to be causal for CAD. This review article discusses the opportunities and challenges of MR studies for CAD, highlighting several examples that involved multiple biomarkers, including various lipid and inflammation traits as well as hypertension, diabetes mellitus, and obesity.

Keywords
Coronary artery disease ● Mendelian randomization

Introduction

Identification and therapeutic targeting of causal cardiovascular risk factors such as hypertension and hypercholesterolaemia has translated into enormous improvements in prevention and therapy of coronary artery disease (CAD). Motivated by these fundamental achievements of modern medicine, epidemiological research has investigated hundreds of biomarkers (including circulating molecules, physiological traits, and other diseases) for association with CAD.1,2

Given the strength and reproducibility of many associations, and potential mechanisms linking the biomarker to CAD risk, several of the novel biomarkers were considered to be causal.3,4 Subsequently, programmes were initiated aiming at medical interventions to block the action of such ‘risk factors’, but in many cases the results did not confer the benefit as predicted from the epidemiological observations.5

The large number of biomarkers, the uncertainty about their causal role in the disease process, and the costs of drug development programmes highlight the need to develop methods that allow discrimination between ‘guilt by causation’ from ‘guilt by association’.6 In selected cases, Mendelian randomization (MR) studies may offer such discrimination.
Genomics of biomarkers and cardiovascular disease

Improved genotyping platforms analysing millions of single nucleotide polymorphisms (SNPs) and globally acting consortia involving tens of thousands of patients and controls have provided sufficient genetic information as well as statistical power to uncover even small genetic effects on various phenotypes. Specifically, genome-wide association (GWA) studies have uncovered multiple variants affecting, for example, LDL or HDL cholesterol, triglycerides, C-reactive protein, interleukin-6 (IL6), blood pressure, obesity, and diabetes mellitus, to name a few of the most intensively studied cardiovascular risk markers.7–12 In parallel, the genetic architecture of CAD has also been defined using a similar approach.13

Thereby, recent large-scale genetic research has laid the foundation for MR studies of CAD. Such studies merge genetic information on both biomarkers and disease (CAD) phenotypes. The goal is to unravel whether the risk marker is causally involved in the disease process (Figure 1A–D).14,15

Mendel’s ‘randomization’ to risk alleles

Most SNPs displaying signals in GWA studies affect quantitatively the level of a phenotype (biomarker) depending on whether an individual had inherited 0, 1, or 2 copies of the variant. As an example, Figure 2 shows the average LDL cholesterol level in a European population, depending on the numbers of a genetic variant (rs2228671) in the LDL cholesterol receptor (LDLR) gene.16 It shows that LDL cholesterol level decreases with the copy number (0, 1, or 2) of the minor allele. The figure also shows that the allele associated with lower LDL cholesterol is indeed likewise associated with a lower risk of CAD.16 Pending a number of conditions that need to be fulfilled by the genetic variant under investigation (see below) and the experimental set-up of the study, such association strongly supports the conclusion that the risk factor (LDL cholesterol) is an intermediary and thus causal step in the observed relationship between the genetic variant and CAD risk.17

Reverse causation

In epidemiological settings, the direction of observed associations cannot be determined. In this case, MR studies might provide further information to elucidate the mechanisms behind this association. For example, similar to LDL cholesterol, C-reactive protein plasma level strongly associates with CAD risk18 and is affected by multiple genetic variants.19–21 However, in contrast to LDL cholesterol, variants affecting C-reactive protein neither individually nor jointly resulted in an increased risk of CAD.20,21 Thus, other explanations must be taken into account for interpreting the observed epidemiological correlation between increased C-reactive protein levels and CAD. The most likely cause for raised C-reactive protein levels in CAD patients might result from inflammatory processes in atherosclerotic plaques. Alternatively, unrecognized inflammatory diseases might influence C-reactive protein levels and CAD risk in parallel. Hence, in this case genetic variants helped to distinguish between ‘guilt’ and ‘innocence’.

Mendelian randomization studies vs. randomized clinical trials

Mendelian randomization studies and randomized clinical trials (RCTs) share many features.22 In a RCT, randomization should result in an equal distribution of clinical features in the study groups (e.g. age, sex, disease severity, social factors etc.) to minimize the chances of these affecting the outcome of a given intervention. This is best reflected by the baseline characteristics of study participants that ideally should be identical in the intervention and control groups. In other words, randomization should result in comparable patient groups except for the drug-mediated modulation of the biomarker. Likewise, Mendel’s second law of independent assortment of alleles should result in an overall equal genetic background of individuals carrying the alleles under investigation. Perhaps more importantly, if cases and controls are drawn from the same population, social and environmental factors should be equally distributed in respective genotype groups. Thereby, groups represented by the respective genotypes should be rather comparable except for the biomarker that is modulated by the genetic variant.

These similarities between the two study designs allow concluding that a MR study can predict the outcome of a RCT that leads to a similar modulation of a biomarker as long as it strictly meets several criteria (see below). The advantage of the MR design is its much lower cost. Indeed, once a population has been genotyped on a genome-wide level, basically every biomarker which is at least in part genetically modulated can be studied in silico by exploration of the data set.

Caveats of Mendelian randomization studies

Pleiotropy

Perhaps the most important limitation to MR is pleiotropy whereby a genetic variant has other effects beyond its effect on the specific biomarker being studied. Such pleiotropy can affect the interpretation of MR studies in multiple ways. First, pleiotropic effects can counteract any effect of the variant on the disease acting via the biomarker, thus giving a null finding even when there is a true causal relationship between biomarker and disease. Alternately, a positive association between the genetic variant and disease may be due to pleiotropic effects and may be mistakenly interpreted as a causal association with the biomarker. As shown in Figure 3 the genetic variant rs964184 displayed genome-wide significant associations with LDL, triglycerides and HDL as well as CAD risk. All associations with the biomarkers could hypothetically explain the increase in CAD risk. Due to such pleiotropic effects it remains unclear which of the lipoproteins actually explains the association with CAD risk.

Such confounding by pleiotropy is least likely where the genetic variant being studied directly lies near the gene for the biomarker under study and affects its level (e.g. C reactive protein gene variants modulating C reactive protein level). However, pleiotropy becomes
increasingly possible where the relationship between a variant and the biomarker is more complex (e.g. for a variant in a gene that codes for a protein that is only part of a complex lipoprotein or for a variant that affects a non-protein phenotype such as blood pressure).

**Linkage disequilibrium**

Genomic loci in close proximity on a given chromosome are usually inherited together. The closer the distance on a chromosome, the higher is the resulting linkage disequilibrium. As an example, a SNP affecting the expression of gene A may be in linkage disequilibrium with a SNP that affects expression of gene B. If the product of gene B is causally related to the disease outcome it would be wrong to conclude that gene A—or the dependent biomarker—is responsible for the phenotype, although such association could be found. Because of this inconsistency of Mendel’s second law (independent inheritance of different traits), gene A and its product may be only indirectly associated with the disease phenotype in a MR approach. To avoid potential misinterpretations, it would be ideal to use only SNPs for MR studies that lie in genomic regions without any further proximity to loci that might circumvent the association of the SNP and the disease.

**Quantitative effect of single nucleotide polymorphism on biomarker and statistical power**

The effect of a SNP on a complex disease can be diluted at least on two levels. First, there are usually multiple genetic variants as well as environmental factors influencing the variability of a biomarker, such that the effect of an individual SNP may be small. Therefore, the variability of the biomarker based on the genetic variant under investigation should be sufficient to affect the disease phenotype, as any further assumptions made in MR studies are based on the quantitative effect of the genetic variant on the presumably intermediary phenotype. Studying single variants comes with the benefit that functional links between variant, affected gene and intermediary phenotype
may be more evident limiting the chance for unexpected pleiotropic effects. If the effect of a single genetic variant is too small, additive effects can be studied by analysing several genetic variants in combination. However, this raises the possibility of potentially adding further pleiotropic effects. Second, any risk factor acts in concert with multiple others and explains only a fraction of the inherited or environmental component underlying a disease such as CAD. Thus, study populations have to be sufficiently large to detect small effects. Moreover, multiple analyses for the CAD association have been carried out on a single CAD GWAS meta-analysis. As this dataset contains most of current GWAS information on common SNPs and CAD risk, appropriate replications in independent samples are difficult to obtain.

**Population stratification**

If the population under investigation is not homogenous, but rather based on two or more substrata, any disease that runs at higher prevalence in one of these subpopulations may display association with all SNPs that are predominantly found in this group. This potential limitation merits specific attention if the genetic variant—biomarker—disease relationships are not studied in the same population, which is considered to be the ideal scenario for a MR study. Needing to combine different datasets to achieve enough power, e.g. findings from a GWAS meta-analysis on a biomarker with another GWAS meta-analysis on CAD, brings its own challenges.

**Canalization**

A potential difference between a genetic variant and clinical biomarker is that the first may affect the biomarker already during childhood (or even earlier) and the second may only be of relevance in the adult individual. Thus, counter-regulatory mechanisms that compensate for the effects related to the SNP in utero or during childhood may blur the association with the genetic variant and the disease. Vice versa, a lifetime exposure to a risk factor may amplify its effects as compared to an epidemiological assessment, as it has been shown for SNPs affecting LDL cholesterol or blood pressure.

**Biomarkers, traits and diseases studied for association with coronary artery disease in Mendelian randomization studies**

The rapid speed with which MR studies are being applied to CAD is demonstrated by a literature search in PubMed, using the search terms MR and CAD (or myocardial infarction), which produced >110 hits. Considering studies, which involved analysis on the relationship between SNP—biomarker as well as the relationship between SNP—and CAD (with a minimum of 5000 cases), our search identified 20 biomarkers, traits, and diseases. These are summarized in Figure 4 and provide details in Supplementary material online, Table S1.
The value of MR studies is best illustrated by adequately powered analyses, in which the effects of a SNP on a biomarker predicted results from large-scale randomized controlled trials, in which a drug modulating the very same biomarker was tested. In this respect, SNPs decreasing serum type secretory phospholipase A2 (sPLA2-IIa) activity displayed no beneficial effects on coronary event rates, which is consistent to what has been shown for varespladip, a clinically tested sPLA2-IIa inhibitor.27 Vice versa, SNPs reducing 3-hydroxy-3-methylglutaryl-coenzyme-A (HMG-CoA) reductase activity—like statins—associated with lower coronary event rates.28 Below we discuss some of the main findings, and for important biomarkers, as well as their implications in terms of therapeutic targeting.

**LDL cholesterol**

The causal role for LDL in promoting CAD was well-established long before MR studies were conducted. In fact, the experience from families carrying LDL receptor mutations documented that markedly increased LDL cholesterol levels increase CAD risk.29 The same conclusion can be drawn from multiple epidemiological and interventional studies and—more recently—from MR studies.7,16,17 As clinical trials often follow their patients only for a few years,30,31 a particular advantage of the MR approach is that it can provide information on the impact of a lifetime modulation of a biomarker.7,16,17 In this respect, MR studies demonstrated that frequent variants in the LDL receptor gene, which increase LDL cholesterol already in childhood by 4 mg/dL, result in stronger effects on CAD risk than predicted by epidemiological or clinical studies for such degree of LDL variability.76 Vice versa, individuals carrying a rare PCSK9 allele, which lowers LDL way below population average (by 21–38 mg/dL), showed a marked 40–80% reduced incidence of myocardial infarction.26 Such MR studies clearly demonstrated the biological relevance of the biomarker LDL without performing a long-term interventional study. Moreover, the genetic studies encouraged the development of new drugs, including monoclonal antibodies against PCSK9, which displayed remarkable effects on LDL cholesterol levels and are now being tested in clinical trials for their potential to decrease coronary risk.32–34

**HDL cholesterol**

Another long-standing drug target has been HDL cholesterol, since a low HDL cholesterol level has been a widely replicated biomarker for CAD risk.35 However, despite consistent evidence from prospective cohort studies and extensive experimental research, a causal interference between HDL cholesterol and CAD has never been proven. Great efforts have been made to medically increase HDL cholesterol level, for example, by CETP-inhibitors, fibrates or nicotinic acid,36 but no consistent effects on CAD prevention have been shown for any of these drugs. Several genetic variants have been identified by GWAS studies to influence HDL cholesterol concentrations in plasma.7 A MR project conducted in Copenhagen used a single SNP, rs4986970 in the LCAT gene, which was associated with a 13% decrease of HDL cholesterol but showed no association with MI risk.37 Comparable results were obtained for four SNPs encoding...
for apolipoprotein-A-I, a major component of HDL. Voight and colleagues recently reported that carriers of another SNP in the endothelial lipase gene, Asn396Ser, had higher HDL cholesterol (0.14 mmol/L or 3 mg/dL compared with non-carriers) but no reduction in incident CAD (OR: 0.99, 95% CI: 0.88 – 1.11, P = 0.85). In contrast, an increase of LDL cholesterol per 1 SD based on a respective change mediated by LDL SNPs in a genetic risk score, had profound effects on incident CAD (OR: 2.13, 95% CI: 1.69 – 2.69, P = 2 × 10⁻⁶). Given these consistent MR results with plasma HDL cholesterol, attempts to decrease CAD risk by measures solely aimed at raising plasma HDL cholesterol may prove futile. Indeed, recent studies using two CETP inhibitors that have markedly raised plasma HDL level have not shown any benefit in terms of CAD risk reduction.

**Cholesteryl ester transfer protein**

Cholesteryl ester transfer protein facilitates the transfer of cholesteryl esters from HDL to LDL cholesterol and it is activity affects plasma concentrations of the two lipoproteins in an inverse manner. In Japanese families, a splicing defect of the CETP gene resulted in CETP deficiency and increased HDL cholesterol levels. Genetic studies on common SNPs revealed association between moderate reduction of CETP mass and activity (≏5–10%) and slightly reduced risk for CAD (5%). In a more recent large-scale MR analysis, Voight et al. found a reduced risk of myocardial infarction by 4% for a variant at the CETP locus, which increases HDL and decreases LDL cholesterol levels. Whether the observed effects on coronary events are due to the modulation of HDL or LDL cholesterol, or other effects mediated by CETP activity, remains obscure at this time.

So far, four CETP inhibitors—Torcetrapib, Dalcetrapib, Anacetrapib, and Evacetrapib—have been tested in large RCTs. Torcetrapib and Dalcetrapib failed to reduce the incidence of CAD.

In fact, Torcetrapib increased the rate of coronary events, potentially mediated by a pleiotropic drug effect, namely an increase in blood pressure. The remaining agents in this class being tested are Anacetrapib and Evacetrapib. In contrast to Torcetrapib, these drugs have more profound effects on LDL levels. Thus, these agents cannot address the effects of an isolated change of HDL on CAD risk. Moreover, at present time it remains open as to whether MR and CRT studies are congruent in showing that lower CETP activity goes along with a mild beneficial effect on coronary event rates.

**Lipoprotein (a)**

Lipoprotein (a) [LP(a)] has been shown to associate with CAD in both cross-sectional and longitudinal studies. With the ability of delivering cholesterol to atherosclerotic lesions and promoting thrombosis by inhibiting plasmin, LP(a) represents an interesting target for interventions. However, due to a lack of specific medication a RCT with the intention to lower LP(a) has not been conducted so far. Interestingly, the LP(a) gene locus was among the first for which genome-wide significant association with CAD was established and a SNP was found to increase both LP(a) levels and CAD risk. A subsequent well-conducted MR study further clearly revealed that LP(a)-associated genetic variants also affect myocardial infarction risk, such that nowadays there is little doubt that LP(a) is a causal risk factor. Therefore drugs that target plasma LP(a) level, unlike plasma HDL cholesterol level, may prove to be very efficacious in preventing CAD.

**Triglycerides**

Remnant lipoproteins compromise a heterogeneous group of triglyceride-rich particles such as very low-density lipoproteins and intermediate-density lipoproteins. These proteins share the ability to accumulate in atherosclerotic plaques and contribute their cholesterol-enriched content to the lesions. Fasting plasma triglyceride concentration was observed to associate with prevalent CAD many years ago. Non-fasting triglycerides—as a direct marker of elevated remnant cholesterol—also associated with the risk of CAD in prospective cohort studies. Bansal et al. directly compared fasting with non-fasting triglycerides for predicting CAD and revealed in ~26,000 American women, after adjustments, that both were associated with risk of CAD.

To test whether elevated non-fasting triglycerides and calculated remnant cholesterol associates with CAD, Jørgensen et al. looked in a classical MR setting at a variant in the APOA5 gene which has been shown repeatedly to be one of the strongest genetic determinants of plasma triglyceride levels. The variant displayed a significant association with increased non-fasting triglycerides and remnant particles and also an increased odds ratio (1.87, P < 0.001) for the risk of a myocardial infarction. In addition, Sarwar et al. assessed the variant rs662799 within the same gene in relation to risk of CAD and found a significant association with coronary heart disease as well [OR: 1.18 (95% CI: 1.11 – 1.26; P = 2.6 × 10⁻⁷)]. Similar findings have been made when multiple SNPs affecting triglyceride levels were studied jointly. For example, Do et al. found a significant association (P = 1 × 10⁻⁷) between triglyceride SNPs and CAD even after adjusting for LDL- and HDL-related effects. Furthermore, they confirmed their finding while restricting the analysis to 44 SNPs with moderate-to-strong effects on triglyceride levels but minimal effect on LDL levels (P = 3 × 10⁻⁰⁵ for association between triglycerides and CAD). Therefore, the totality of the MR analyses to date indicate that triglycerides are a causal risk factor for CAD although further work needs to be done on triglyceride subtypes regarding their specific risk contribution.

**Lipoprotein-associated phospholipase A2**

Lipoprotein-associated phospholipase A2 is an enzyme produced mainly by inflammatory cells, including macrophages and lymphocytes. It circulates bound to LDL particles and it produces pro-apoptotic and pro-inflammatory mediators (e.g. precursors of arachidonic acid), which may influence vascular function, growth of atherosclerotic plaques, and inflammation in plaques as well. Furthermore, LP-PLA2 mass and activity has been described to accumulate in unstable and ruptured plaques.58 In epidemiological settings, LP-PLA2 mass and activity has been found to associate with CAD risk.59–62 However, a gain-of-function mutation in the PLASG7 gene encoding for LP-PLA2 was without effect on coronary atherosclerosis or CHD events in a recent meta-analysis including 26,000 Europeans. Likewise, a loss-of-function mutation (V279F) was without consistent effects in a meta-analysis on seven predominantly Asian case–control studies (albeit in south Koreans, this loss-of-function mutation led to a reduced risk of CAD). Thus, the majority of genetic data argued against a causal role of LP-PLA2. Indeed, darapladib, a medication
which inhibits LP-PLA2 activity, failed to meet the endpoints in a trial enrolling 16,000 patients with a history of an acute coronary syndrome.\textsuperscript{64} Such negative effect might have predicted by the MR studies. Another study on darapladib (SOLID TIMI 52) investigates recurrent coronary events in 13,000 patients and is still ongoing.\textsuperscript{65}

**Serum type secretory phospholipase A\(_2\)**

Another enzyme, serum type secretory phospholipase A\(_2\) (sPLA\(_2\)-IIa), has been associated with the risk cardiovascular events in prospective settings as well.\textsuperscript{66,67} Serum type secretory phospholipase A\(_2\) has various potential roles in the atherosclerotic process. For example, it hydrolyses phospholipids on lipoproteins, which leads to an increased binding of LDL to preotoglycans in the arterial wall and accelerates the formation of atherosclerotic plaques. Additionally, it enhances the amount of oxidative stress by generating arachidonic acid, lysophospholipids, and non-esterified fatty acids.\textsuperscript{58} Breitling \textit{et al.} assessed in a single-cohort approach the association between variants encoding sPLA\(_2\)-IIa and the serum concentration of sPLA\(_2\)-IIa and secondary CAD events.\textsuperscript{68} Holmes \textit{et al.} conducted a MR meta-analysis involving 19 population studies. Despite the fact that the allele of interest led to a remarkable reduction of sPLA\(_2\)-IIa enzyme activity and sPLA\(_2\)-IIa mass, they found no association with incident major vascular events, which raises doubt about a causal role of this enzyme.\textsuperscript{69} In fact, a trial testing a medication (varespladib in VISTA-16) designed to selectively block sPLA\(_2\)-IIa was halted for lack of efficacy.\textsuperscript{27}

**C-reactive protein**

Multiple studies have shown that plasma C-reactive protein level is robustly associated with prevalent and future CAD events.\textsuperscript{18,70} Moreover, studies employing statins not only demonstrated that elevated C-reactive protein is a good marker for identification of patients with elevated risk but also that a decrease of C-reactive protein levels by this medication goes along with a reduced incidence of cardiovascular events.\textsuperscript{71} If C-reactive protein acts as a causal factor for CAD, drugs that lower C-reactive protein levels should also reduce the risk of CAD. In fact, such drugs are currently under development.\textsuperscript{72} However, several large MR studies have now convincingly excluded a role of plasma C-reactive protein level in CAD.\textsuperscript{19–21,73} Thus, medical lowering of C-reactive protein levels is unlikely to be successful in preventing CAD.

**Interleukin-6 receptor**

Despite the doubts about C-reactive protein as a target for CAD prevention, inflammatory markers still remain an interesting field in the search for causative factors and novel drug targets. Interleukin 6 as a pro-inflammatory agent binds to the IL-6 receptor, which is located on hepatocytes, monocytes, and the endothelial wall and mediates inflammatory responses. In a prospective study, IL-6 was associated with adverse cardiovascular prognosis.\textsuperscript{74} In addition, patients with incident CAD had increased concentrations of circulating IL-6 prior to their event such that blockade of IL-6 binding might provide a way to reduce atherosclerosis and future CAD events.\textsuperscript{75,76} Tocilizumab, a monoclonal IL-6 antibody, has already proved its ability to sufficiently alter the course of rheumatoid arthritis by reducing articular inflammation.\textsuperscript{77} The question as to whether Tocilizumab can also prevent CAD has not yet been studied in RCTs. The genetics of the IL-6 receptor provides an adequate tool to investigate such association in a MR study. Patients carrying the allele leading to reduced IL-6 binding not only showed a significant lower inflammatory response but also a remarkable reduction in CAD events compared with individuals carrying the alternate allele.\textsuperscript{78} These findings raise hope that IL-6 receptor inhibition may be suitable for primary or secondary prevention of CAD. The IL-6 receptor variant also showed associations with reduced C-reactive protein and fibrinogen concentrations, most likely due to downstream effects of IL-6 signalling. Thus, although this SNP affects both C-reactive protein and fibrinogen level, it is not suitable to study these biomarkers—rather than IL-6—in a MR study design. This illustrates one of the caveats of MR studies whereby testing and interpretation are most robust when the variant directly affects the gene for the biomarker being evaluated.

**Pentraxin 3**

Pentraxin 3 (PTX3) has been shown to be involved in vascular inflammation.\textsuperscript{79} Dubin \textit{et al.} revealed in a prospective study that PTX3 concentrations associate with increased risk for all-cause-mortality, cardiovascular events, and incident heart failure. Similar findings were made in case-control studies.\textsuperscript{80} Investigating three common PTX3 polymorphisms Barbati \textit{et al.} observed association with PTX3 plasma levels but not with the risk of acute myocardial infarction, suggesting that PTX3 is likely not a causal risk factor.

**Fibrinogen**

Fibrinogen—with its ability to form fibrin—is the major component of blood clots and thus involved in the manifestation of atherothrombotic events.\textsuperscript{81} Moreover, fibrinogen affects multiple inflammatory conditions.\textsuperscript{82} Elevated levels of fibrinogen have been found in patients suffering from CAD and data from prospective studies established association between fibrinogen and risk of coronary heart disease.\textsuperscript{83} To study a possible causal relationship between the variability of fibrinogen serum levels and CAD, Sabater-Uleal \textit{et al.} conducted a multi-ethnic meta-analysis of genome-wide association studies in over 100,000 subjects. The merged effect of the SNPs increasing fibrinogen showed no evidence for an association with CAD suggesting a lack of causality.

**Blood pressure**

Based on overwhelming evidence from epidemiological and intervention studies high blood pressure can be considered as a proven causal factor for CAD.\textsuperscript{87,88} Further evidence comes now from a genetic study.\textsuperscript{25} Specifically, 30 SNPs have been found to associate with systolic and diastolic blood pressure, each of them leading to an increase of 0.5–1.2 mmHg in systolic blood pressure.\textsuperscript{10} Analysing data from almost 22,500 CAD cases and 65,000 controls, the CARDIoGRAM investigators found an average increase of CAD risk by 3% per risk allele. Patients in the highest quintile in terms of numbers and effect sizes of blood pressure alleles had a 70% higher odds of having CAD, when compared with patients in the bottom quintile of a genetic risk score distribution.\textsuperscript{25} These findings were remarkable for the fact that—like LDL SNPs—blood pressure associated SNPs had stronger effects on CAD risk than expected by epidemiological studies. In observational settings, the true risk mediated by blood pressure might be underestimated due to a regression dilution bias.
that may result from inherent inaccuracies in the measurement of blood pressure. In contrast, a SNP may be a more precise denominator of a small variability in blood pressure as in the nowadays GWAS era this association (between a SNP and blood pressure) stems from thousands of measurements. Furthermore, the SNPs under investigation may affect mechanisms, i.e. endothelial function that—aside from blood pressure increase—may have a direct effect on vascular biology. These effects might be more closely related to the development of CAD than the variability in blood pressure itself.

**Body mass index and obesity**

Obesity—widely assessed as elevated body mass index (BMI)—has consistently shown to associate with future risk of cardiovascular complications. Nevertheless, causality of the association has not yet been proved. Three loci (FTO, MC4R, and TMEM18) with the largest known effect on BMI were tested for their association with ischemic heart disease in 75,627 individuals demonstrating that a genetically driven increase in BMI of 0.28 kg/m² per individual allele translates into a significant 3% increase of CAD risk. This study adds to the emerging evidence that higher BMI or the condition leading to this play a causal role in the development of CAD. However, there are pleiotropic effects—namely type 2 diabetes—reported at least for FTO and MC4R risk alleles, which might affect the observed associations. Thus, the mechanism linking obesity-related genetic variants and CAD risk might include additional phenotypes. Further investigations are required to unmask such traits.

**Diabetes mellitus**

Although there is a two to three-fold higher risk of CAD in patients with diabetes, there is an ongoing debate about the exact nature of the relationship. Moreover, recent interventional studies have raised doubt about causality as aggressive blood glucose lowering failed to sufficiently reduce cardiovascular events. Genetic variants might provide an adequate tool to further answer the question of a causal interference. About 40 variants have been found to associate with type 2 diabetes in Western-Europeans. These variants were tested for their association with CAD in CARDIoGRAM. Interestingly, diabetes SNPs had only a mild impact on CAD (the average increase in CAD risk observed per individual type 2 diabetes risk allele was 1.0076, \( P = 0.02 \) for OR). Albeit this increase was statistically significant ( \( P = 5.8 \times 10^{-5} \) ), it contrasted also significantly from the effect that was expected ( \( 1.067, P = 7.1 \times 10^{-10} \) for the difference between observed and expected effects), based on the effects of these alleles on diabetes risk and the effect of diabetes on CAD risk as observed in the Framingham Heart Study. This might be in part because of a potential overestimation of the respective SNP effects on diabetes in GWAS (winner's curse). Another explanation refers to a hypothetical overestimation of diabetes effects on CAD in epidemiological settings. This is emphasized by recent findings from clinical trials which disappointed in that aggressive blood glucose lowering failed to further decrease coronary events. Moreover, when compared with other quantitative risk factors, e.g. genetically mediated high blood pressure or LDL cholesterol levels, type 2 diabetes often manifests in the late adulthood and only then starts to affect the risk of CAD. Such patients may be underrepresented in CARDIoGRAM.

**Telomere length**

Several cross-sectional and longitudinal studies have shown an association between shorter mean leucocyte telomere length (LTL) and CAD. Whether this reflects a causal association or is a confounded association due to the known effect on telomere length of other putative CAD risk factors such as oxidative stress and inflammation has been a matter of some debate. Recently, through GWAS, Codd et al. identified seven variants associated with mean LTL, several in genes that are components of the telomerase complex. A genetic risk score analysis combining lead variants at all seven loci in the CARDioGRAM CAD GWAS meta-analysis showed an association of the alleles associated with shorter LTL with an increased risk of CAD [21% (95% confidence interval, 5–35%) per standard deviation in LTL]. These findings suggest a causal association of shorter telomere length with risk of CAD, the mechanism of which merits further investigation.

**Summary**

Mendelian randomization has emerged as a valuable approach in investigating whether an association of a biomarker with CAD is causal or not. Already, the evidence points to several long-held candidates (plasma HDL cholesterol level, C-reactive protein) as not being causal. On the other hand the likely causal involvement of other biomarkers [LP(a), IL-6] has been enhanced providing greater confidence that efforts to target them therapeutically will prove rewarding. The instruments for carrying out MR studies are rapidly improving and will be of great benefit for future decision-making upon the development of novel drug targets. However, despite the convincing concept of MR analysis, several limitations and requirements have to be taken into consideration while designing and interpreting a MR study. In this regard, the MR study adds to established study designs (like RCT) without the ability to fully replace them.

**Supplementary material**

Supplementary material is available at European Heart Journal online.

**Funding**

This work was supported by the EU-funded Integrated Project CIViGenes@target (HEALTH-F2-2013-601456), Transatlantic Networks of Excellence in Cardiovascular Research Program of the Fondation Leducq, and as well as the BMBF-funded German National Genome Network (NGFN-Plus) Project Atherogenomics (FKZ: 01GS0831) and eAtheroSysMed (FKZ: 01ZX1313A). N.J.S. holds a chair funded by the British Heart Foundation and is a NIHR senior investigator.

**References**

Mendelian randomization studies in coronary artery disease


Mendelian randomization studies in coronary artery disease


