Mendelian randomization to assess causal effects of blood lipids on coronary heart disease: lessons from the past and applications to the future

Stephen Burgess and Eric Harshfield

Purpose of review
Mendelian randomization is a technique for judging the causal impact of a risk factor on an outcome from observational data using genetic variants. Although evidence from Mendelian randomization for the effects of major lipids and lipoproteins on coronary heart disease (CHD) risk has been around for a relatively long time, new data resources and new methodological approaches have given fresh insight into these relationships. The lessons from these analyses are likely to be highly relevant when it comes to lipidomics and the analyses of lipid subspecies whose biology is less well understood.

Recent findings
Although analyses of low-density lipoprotein cholesterol and lipoprotein(a) are unambiguous as there are genetic variants that associate exclusively with these risk factors and have well understood biology, analyses for triglycerides, and high-density lipoprotein cholesterol (HDL-c) are less clear. For example, a subset of genetic variants having specific associations with HDL-c are not associated with CHD risk, but an allele score including all variants associated with HDL-c does associate with CHD risk. Recently developed methods, such as multivariable Mendelian randomization, Mendelian randomization-Egger, and a weighted median method, suggest that the relationship between HDL-c and CHD risk is null, thus confirming experimental evidence.

Summary
Robust methods for Mendelian randomization have important utility for understanding the causal relationships between major lipids and CHD risk, and are likely to play an important role in judging the causal relevance of lipid subspecies and other metabolites measured on high-dimensional phenotyping platforms.

Keywords
blood lipids, coronary heart disease, instrumental variables, Mendelian randomization, metabolomics

INTRODUCTION

Mendelian randomization is the use of genetic variants as proxies for increased or decreased exposure to a modifiable phenotype (hereafter referred to as a risk factor) to help judge whether clinical or pharmaceutical interventions on the risk factor are likely to lead to changes in a disease outcome [1,2▪]. The most straightforward application of Mendelian randomization involves taking a single genetic variant that is associated with the risk factor, but not associated with other risk factors that may represent confounders or alternative causal pathways to the outcome [3]. Such a genetic variant may be hard to find, but for protein biomarkers such as fibrinogen or C-reactive protein, genetic variants in or near the relevant coding region (in these cases, the *FGB* and *CRP* gene regions, respectively [4,5]) have been shown to have good specificity of association with the risk factor, at least for measured confounders. An association between such a genetic variant and the outcome is indicative of a causal effect of the risk factor on the outcome [6]. In other cases, such as for IL-1 [7], genetic variants may be associated with alternative risk factors (in this case, C-reactive

Department of Public Health and Primary Care, Cardiovascular Epidemiology Unit, University of Cambridge, Cambridge, UK

Correspondence to Stephen Burgess, Department of Public Health and Primary Care, Cardiovascular Epidemiology Unit, University of Cambridge, Cambridge, UK. E-mail: sb452@medschl.cam.ac.uk

Curr Opin Endocrinol Diabetes Obes 2016, 23:124–130
DOI:10.1097/MED.0000000000000230

KEY POINTS

- Mendelian randomization is a technique to determine the causal impact of a risk factor on an outcome from observational data using genetic variants.
- Robust methods for Mendelian randomization have important utility for understanding the causal relationships between major lipids and CHD risk.
- Multivariable Mendelian randomization, Mendelian randomization-Egger, and a weighted median method for Mendelian randomization are important recently developed methods that are likely to play an important role in judging the causal relevance of lipid subspecies and other metabolites measured on high-dimensional phenotyping platforms.
- There is tremendous scope and untapped potential to apply Mendelian randomization in investigating plausible novel causal pathways of high-dimensional phenotypic traits with diseases and risk factors.

protein and IL-6), but so long as these risk factors represent a single causal pathway (i.e., they are up or downstream of the target risk factor and there is no alternative causal pathway from the genetic variants to the outcome that does not go via the target risk factor), the assumptions necessary for Mendelian randomization would not be violated [8].

Under further parametric assumptions (including linearity), an estimate of the causal effect of the risk factor on the outcome can be obtained [9]. The causal estimate from Mendelian randomization is likely to differ from the impact of intervening on the risk factor in practice for many reasons (e.g., the genetic effect is lifelong) [10]. Hence the magnitude of the causal estimate should not be taken too literally, but the causal estimate is a valid test statistic for testing the causal null hypothesis. This enables information on multiple genetic variants to be combined into a single causal estimate, which has greater power to detect a causal effect than a test of the association of any of the individual genetic variants with the outcome [11]. One recent innovation is the use of summarized data on genetic associations with the risk factor and with the outcome to obtain a causal estimate [12,13]. These associations can come from a single dataset (one-sample setting), or from separate datasets (two-sample setting) [14]. A practical advantage of the use of summarized data is the ability to analyse publicly available data from large consortia [15] – such as the Global Lipids Genetics Consortium [16], who have made associations of genetic variants with low-density lipoprotein cholesterol (LDL-c), high-density lipoprotein cholesterol (HDL-c), and triglycerides in over 188 000 individuals available (http://www.sph.umich.edu/csg/abecasis/public/lipids2013/), and CARDIoGRAMplusC4D [17], who have made associations with coronary heart disease (CHD) risk available in over 60 000 cases and 125 000 controls (http://www.cardiogramplusc4 d.org/downloads/). These methods and data resources (in particular, the large sample sizes of consortium data and the ease of obtaining genetic association estimates) have revolutionized the practice and power of Mendelian randomization investigations [18].

In this review, we first consider Mendelian randomization analyses of major lipids and lipoproteins – LDL-c, HDL-c, triglycerides, and lipoprotein(a) – on CHD risk. We discuss methodological innovations in Mendelian randomization (motivated in part by these analyses), and their application to the assessment of the causal effects of these lipids on CHD risk. Finally, we explore the impact of technological advances, such as differentiating major lipids into lipid subspecies using high-dimensional phenotyping platforms, and the potential utility of these advances in a Mendelian randomization framework.

MENDELIAN RANDOMIZATION ANALYSES OF MAJOR LIPIDS AND LIPOPROTEINS ON CORONARY HEART DISEASE RISK

Genetic evidence for a link between hypercholesterolemia and CHD risk has a long history [19] that precedes the popularization of Mendelian randomization. Links between LDL-c and CHD risk are well established for both common and rare genetic variants [20], and formal approaches for Mendelian randomization have clearly shown a deleterious causal effect of increased LDL-c on CHD risk [21,22]. In many ways, LDL-c is an ideal risk factor for use in Mendelian randomization. Several genetic variants associated with LDL-c are located in gene regions that also have corresponding pharmaceutical interventions, such as the *HMGCR* gene region for statins [23], and the *PCSK9* gene region for proprotein convertase subtilisin/kexin type 9-inhibitors [24,25]. Indeed, a Mendelian randomization analysis using variants in the *NPC1L1* gene region [26▪▪] was published in advance of a large trial of ezetimibe (a Niemann-Pick C1-Like 1-inhibitor) [27], and correctly predicted its result. The possible benefit of combination therapy by statin and ezetimibe has been considered in a factorial Mendelian randomization analysis, comparing individuals with genetically lowered LDL-c because of *HMGCR* variants alone, because of *NPC1L1* variants alone, and because of the presence of variants in both gene regions [28]. Genetic variants in different gene regions, as well as genetic variants with varying strengths of association

with LDL-c concentrations (including rare gain-of-function and loss-of-function mutations with large effects on LDL-c) have been shown to have associations with CHD risk that are proportional to their association with LDL-c [29], both strengthening the argument that LDL-c is the relevant causal risk factor, and suggesting that all mechanisms of LDL-c-lowering seem to have similar effects on CHD risk. However, the magnitude of the genetically-predicted causal effect of LDL-c on CHD risk is much larger than the observed reduction in CHD risk from taking statins; the Mendelian randomization estimate is 3.5 times larger than the estimate from trials [10] (based on taking statins for 5+ years in primary prevention [30]). One explanation for this is that genetically predicted variation in LDL-c concentrations is lifelong, and so the Mendelian randomization estimate represents the effect of long-term reduction in LDL-c. Genetic studies have corroborated the slight increases in type 2 diabetes (T2D) risk that are observed in statin trials [31], with several LDL-c-lowering variants showing suggestive associations with increased T2D risk [32▪]. This suggests that the increase in T2D risk is likely to be an on-target effect of statin drugs, rather than an off-target effect; also that it may be a consequence of LDL-c-lowering more widely rather than a specific effect of intervention on the 3-hydroxy-3-methylglutaryl-CoA reductase pathway.

A similar story can be told for lipoprotein(a). The kringle IV type 2 size polymorphism (a copy number variant) is highly predictive of lipoprotein(a) concentrations, explaining 21% of variation in lipoprotein(a) [33]. In contrast, no genetic variant for LDL-c, HDL-c, or triglycerides explains more than 1% of variation [16]. This polymorphism (and also single nucleotide polymorphisms in the *LPA* region [34]) are also associated with CHD risk, suggesting a deleterious causal effect of increased lipoprotein(a) on CHD risk [33]. Similarly to that for LDL-c, the effect estimate from the Mendelian randomization analysis is 2.5 times greater than that from a standard observational analysis [33]. Another explanation for this, which may also be relevant to other Mendelian randomization analyses, is that genetic variants may be associated with another aspect of lipid biology (such as particle size or activity) and not just concentration.

For triglycerides, the story is less clear because of a lack of genetic variants associated with triglycerides that do not also associate with LDL-c and/or HDL-c. A methodological development to address this is multivariable Mendelian randomization, in which the causal effects of multiple risk factors can be estimated simultaneously [35▪]. This requires genetic variants to be associated with one or more of the risk factors, but not associated with other risk factors that may represent confounders of any risk factor – outcome association or alternative causal pathways to the outcome that are not via one of the target risk factors. Multivariable Mendelian randomization analyses have suggested a deleterious causal effect of increased triglycerides on CHD risk [36,37]. However, there is little consistency in the associations of individual triglyceride-related variants with CHD risk [38], with some variants being associated with CHD risk [39,40], and others showing no clear association. This may reflect genuine heterogeneity among different triglycerides.

Although there are genetic variants that appear to have specific associations with HDL-c, these variants are not associated with CHD risk [41]. However, an allele score based on all the genetic variants known to be associated with HDL-c at a genome-wide level of significance is associated with CHD risk, suggesting a protective causal effect of HDL-c (if the Mendelian randomization assumptions are satisfied – see later) [42▪]. Holmes *et al.* demonstrated an inverse association with CHD risk for an unrestricted score that explained 3.8% of the variance in HDL-c, but no association for a restricted score omitting variants additionally associated with LDL-c or triglycerides that explained 0.3% of the variance in HDL-c. One explanation for the null finding with the restricted score is that the analysis lacked the power to detect a causal effect. Multivariable Mendelian randomization is a useful tool in this case, as a multivariable analysis can include genetic variants that have pleiotropic associations with either LDL-c or triglycerides. This provides robustness to pleiotropy but still reasonable power to detect a causal effect. A multivariable Mendelian randomization analysis using a limited number of genetic variants did not reveal a causal effect of HDL-c [35▪], and neither did an initial analysis, including all genome-wide significant variants [36]. Although a more principled multivariable Mendelian randomization analysis taking into account the relative weights of the genetic variants did suggest a protective effect of HDL-c [37], the magnitude of the effect was much smaller (4.5 times smaller) than that for LDL-c; there is also the potential of some residual bias because of pleiotropic associations of the 185 genetic variants.

METHODOLOGICAL ADVANCES IN MENDELIAN RANDOMIZATION AND RELATION TO MENDELIAN RANDOMIZATION ANALYSIS OF MAJOR LIPIDS

Two other methodological advances that have relevance to assessing the causal relevances of major

lipids are: Mendelian randomization-Egger [43▪▪] and a weighted median method [44▪▪]. Mendelian randomization-Egger is a method adapted from the meta-analysis literature on publication bias [45]. In a Mendelian randomization setting, each genetic variant contributes an estimate of the causal effect, and a pooled estimate is calculated based on all the genetic variants (genetic variants are treated similarly to studies in a meta-analysis). However, if even one of the genetic variants violates the Mendelian randomization assumptions, then the causal estimate from that variant will be biased, and the usual pooled estimate (known as the inverse-variance weighted estimate [13]) will be biased and have an inflated type 1 error rate. This may lead to false positive findings when genetic variants are pleiotropic [46]. Rather than the standard approach, which assesses whether genetic variants associated with the risk factor are also associated with the outcome, Mendelian randomization-Egger assesses whether there is a dose-response relationship in the genetic associations with the risk factor and with the outcome. This is a higher standard of proof than demanded in a standard Mendelian randomization analysis, and so Mendelian randomization-Egger has reduced type 1 error rates [43▪▪]. Mendelian randomization-Egger enables a test of 'directional pleiotropy' (whether pleiotropic associations of genetic variants are likely to bias causal estimates in one particular direction). Additionally, under the assumption that genetic variants may have pleiotropic effects on the outcome, but that these pleiotropic effects are uncorrelated with instrument strength [47], Mendelian randomization-Egger provides a consistent estimate of the causal effect [43▪▪].

The weighted median method is a simple idea: rather than taking a pooled estimate that is a weighted mean of the causal estimates based on each genetic variant individually, to report a pooled estimate that is a weighted median [48]. The median is not affected by outlying results, and so the weighted median estimate is not sensitive to a handful of pleiotropic genetic variants. Formally, it is a consistent estimate of the causal effect if at least half of the genetic variants (by weight) are valid instruments [44▪▪]. Both approaches are worthwhile sensitivity analyses for Mendelian randomization when some genetic variants are suspected to be pleiotropic. The Mendelian randomization-Egger estimate has the advantage that it allows all genetic variants to be pleiotropic, although it makes an assumption on the distribution of these pleiotropic effects; however, it may be imprecise, and it is highly influenced if there are one or two strong variants. The weighted median estimate is more precise and more stable, but relies on the majority of evidence in the analysis being reliable.

The application of these methods to major lipids is very revealing: using all genome-wide significant variants, all analyses (standard Mendelian randomization, Mendelian randomization-Egger, weighted median) suggest causal effects of LDL-c and triglycerides on CHD risk, with no evidence of directional pleiotropy [44▪▪]. However, although the standard Mendelian randomization analysis using all genome-wide significant variants for HDL-c suggests a protective effect of HDL-c on CHD risk, the Mendelian randomization-Egger and weighted median analyses suggest a null effect, with evidence of directional pleiotropy in the Mendelian randomization-Egger analysis [44▪▪]. This null finding is supported by trial evidence on cholesteryl ester transfer protein inhibitors, which raise HDL-c levels, but do not lower CHD incidence [49].

The conclusion from this is that Mendelian randomization analyses can be simple or not, depending on the available genetic variants and their specificity of association with the risk factor under analysis. A naive Mendelian randomization analysis, particularly one using a large number of genetic variants, can be misleading. However, the development of new methods can help either to add confidence in the finding from a Mendelian randomization analysis, or to call it into question [50▪].

FUTURE DIRECTIONS FOR MENDELIAN RANDOMIZATION

High-throughput phenotyping approaches to collecting 'omics' data, including genomics, transcriptomics, metabolomics, and proteomics, have recently been gaining traction; new approaches are constantly being developed to measure an ever-widening number of phenotypic traits on larger and larger populations. The measurement of such a vast array of high-dimensional phenotypic traits brings novel opportunities to perform genome-wide association studies (GWAS) that can examine the associations of millions of genetic variants with thousands of metabolites or proteins [51▪▪]. Lipidomics is a subset of metabolomics concerned with the study of lipid profiles derived from mass spectrometry or NMR platforms, which produces information on the composition and abundance of lipids in the body, thereby contributing to an understanding of how lipids function in a biological system [52]. Although numerous metabolomics GWAS have been performed in recent years [53▪▪], very few high-dimensional phenotyping studies have used a Mendelian randomization approach to assess whether the associated phenotypic traits that they identified could have causal effects on diseases or risk factors.

The studies that have employed Mendelian randomization on high-throughput data have taken either of two approaches: to determine the causal role of conventional risk factors on levels of high-dimensional phenotypes (e.g., metabolites), or to determine whether high-dimensional phenotypes have a causal effect on diseases or traits. As an example of the first approach, a meta-analysis of four Finnish population cohorts obtained levels of 82 different metabolites and metabolic measures using nuclear magnetic resonance, including lipoprotein lipids, fatty acids, and amino acids [54]. The authors found evidence that strongly supports causal effects of adiposity on 24 metabolites that are potential cardiometabolic risk factors [54]. Another study using mass spectrometry in a British population determined that gene expression levels derived from expression quantitative trait loci in fat, skin, and lymphoblastoid cell lines could play a causal role on levels of a wide range of metabolites [55]. The authors identified two loci (*THEM4* and *CYP3A5*) where the allele associated with increased metabolite levels was significantly associated with decreased gene expression in one or more tissues, supporting the notion that the underlying causal variants at these two loci could have regulatory consequences [55]. To illustrate the second approach, a prospective cohort study that conducted mass spectrometry used summarized CHD association results from CARDIoGRAMplusC4D to find four lipid-related metabolites (lysophosphatidylcholines 18 : 1 and 18 : 2, monoglyceride 18 : 2, and sphingomyelin 28 : 1) with evidence for a causal role in CHD development [56▪].

Metabolomics and proteomics particularly stand to benefit from the availability of summarized data for Mendelian randomization and a two-sample setting, where the associations of high-dimensional phenotypic traits with genetic variants are measured in one population (usually a small cross-sectional study of healthy individuals) and the associations of those variants with diseases and risk factors are measured in another population, such as the large consortia mentioned earlier (for disease outcomes, usually a consortium of case-control studies) [15]. Furthermore, the multivariable Mendelian randomization approach will be particularly relevant to high-dimensional platforms, as it may be difficult to find genetic variants having a specific association with a single variable (and in lipidomics in particular [57]). However, it is important to distinguish between pleiotropy and mediation (also called 'horizontal' and 'vertical' pleiotropy) [15]: if several metabolites are on the same causal pathway, then a genetic variant associated with all of these metabolites is not truly pleiotropic, as the associations reflect a single causal pathway. In this case, a Mendelian randomization analysis can assess the causal effect of the entire pathway, but it cannot address the question of causation for any of the individual metabolites on the pathway without incorporating additional biological information. The particular challenge of Mendelian randomization with high-dimensional assays lies in identifying a suitable set of genetic variants for a particular metabolite or protein (or a small set of metabolites or proteins for multivariable Mendelian randomization) that will not violate Mendelian randomization assumptions. Thus, the Mendelian randomization-Egger and weighted median methods could be especially important to provide some robustness against pleiotropic variants.

CONCLUSION

There is tremendous scope and untapped potential to apply Mendelian randomization in investigating plausible novel causal pathways of high-dimensional phenotypic traits with diseases and risk factors. Mendelian randomization is a tool that can provide additional evidence to prioritize further research and clinical applications, or just as importantly, to discourage additional resource allocation toward a specific pathway. Over the next few years, Mendelian randomization is likely to be applied with increasing regularity to high-dimensional phenotypic data where concomitant genetic information is available, and in lipidomics in particular.

Acknowledgements

The authors would like to thank Robert Scott (MRC Epidemiology Unit, University of Cambridge) and Adam Butterworth (Cardiovascular Epidemiology Unit, University of Cambridge) for helpful comments on an earlier draft of this work.

Financial support and sponsorship

S.B. is supported by the Wellcome Trust (grant number 100114).

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- ▪ of special interest
- ▪▪ of outstanding interest

1. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 2003; 32:1–22.

2. Burgess S, Thompson SG. Mendelian randomization: methods for using ■ genetic variants in causal estimation. Chapman & Hall; 2015.
Comprehensive introduction to Mendelian randomization. Available at all good bookshops.
3. Hingorani A, Humphries S. Nature's randomised trials. Lancet 2005; 366:1906–1908.
4. Keavney B, Danesh J, Parish S, *et al.* Fibrinogen and coronary heart disease: test of causality by 'Mendelian randomization'. Int J Epidemiol 2006; 35:935–943.
5. CRP CHD Genetics Collaboration. Association between C reactive protein and coronary heart disease: Mendelian randomisation analysis based on individual participant data. Brit Med J 2011; 342:d548.
6. VanderWeele T, Tchetgen Tchetgen E, Cornelis M, Kraft P. Methodological challenges in Mendelian randomization. Epidemiology 2014; 25:427–435.
7. The Interleukin-1 Genetics Consortium. Cardiometabolic consequences of genetic up-regulation of the interleukin-1 receptor antagonist: Mendelian randomisation analysis of more than one million individuals. Lancet Diabetes Endocrinology 2015; 3:243–253.
8. Burgess S, Butterworth AS, Thompson JR. Beyond Mendelian randomization: how to interpret evidence of shared genetic predictors. J Clin Epidemiol 2016; 69:208–216.
9. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. Stat Methods Med Res 2007; 16:309–330.
10. Burgess S, Butterworth A, Malarstig A, Thompson S. Use of Mendelian randomisation to assess potential benefit of clinical intervention. Brit Med J 2012; 345:e7325.
11. Pierce B, Ahsan H, VanderWeele T. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. Int J Epidemiol 2011; 40:740–752.
12. Johnson T. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Technical Report, Queen Mary University of London; 2011. http://webspace.qmul.ac.uk/tjohnson/gtx/outline2.pdf.
13. Burgess S, Butterworth A, Thompson S. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013; 37:658–665.
14. Pierce B, Burgess S. Efficient design for Mendelian randomization studies: subsample and two-sample instrumental variable estimators. Am J Epidemiol 2013; 178:1177–1184.
15. Burgess S, Scott R, Timpson N, *et al.,* EPIC-InterAct Consortium. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. Eur J Epidemiol 2015; 30:543–552.
16. The Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013; 45:1274–1283.
17. The CARDIoGRAMplusC4D Consortium. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 2013; 45:25–33.
18. Burgess S, Timpson NJ, Ebrahim S, Smith GD. Mendelian randomization: where are we now and where are we going? Int J Epidemiol 2015; 44:379–388.
19. Austin MA, Hutter CM, Zimmern RL, Humphries SE. Familial hypercholesterolemia and coronary heart disease: a HuGE association review. Am J Epidemiol 2004; 160:421–429.
20. Strong A, Rader DJ. Clinical implications of lipid genetics for cardiovascular disease. Curr Cardiovasc Risk Rep 2010; 4:461–468.
21. Linsel-Nitschke P, Götz A, Erdmann J, *et al.* Lifelong reduction of LDL-cholesterol related to a common variant in the LDL-receptor gene decreases the risk of coronary artery disease: a Mendelian randomisation study. PLoS One 2008; 3:e2986.
22. Waterworth D, Ricketts S, Song K, *et al.* Genetic variants influencing circulating lipid levels and risk of coronary artery disease. Arterioscler Thromb Vasc Biol 2010; 30:2264–2276.
23. Pedersen T, Kjekshus J, Berg K, *et al.* Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). Lancet 1994; 344:1383–1389.
24. Cohen J, Boerwinkle E, Mosley T Jr, Hobbs H. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med 2006; 354:1264–1272.
25. Fitzgerald K, Frank-Kamenetsky M, Shulga-Morskaya S, *et al.* Effect of an RNA interference drug on the synthesis of proprotein convertase subtilisin/kexin type 9 (PCSK9) and the concentration of serum LDL cholesterol in healthy volunteers: a randomised, single-blind, placebo-controlled, phase 1 trial. Lancet 2014; 383:60–68.
26. Stitzel NO, Won HH, Morrison AC, *et al.* Inactivating mutations in NPC1L1 ■■ and protection from coronary heart disease. N Engl J Med 2014; 371:2072–2082.
First genetic evidence that ezetimibe is likely to be effective in lowering heart disease risk.
27. Cannon CP, Blazing MA, Giugliano RP, *et al.* Ezetimibe added to statin therapy after acute coronary syndromes. N Engl J Med 2015; 372:2387–2397.
28. Ference BA, Majeed F, Penumetcha R, *et al.* Effect of naturally random allocation to lower low-density lipoprotein cholesterol on the risk of coronary heart disease mediated by polymorphisms in NPC1L1, HMGCR, or both: a 2×2 factorial Mendelian randomization study. J Am Coll Cardiol 2015; 65:1552–1561.
29. Ference BA, Yoo W, Alesh I, *et al.* Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. J Am Coll Cardiol 2012; 60:2631–2639.
30. Taylor F, Ward K, Moore T, *et al.* Statins for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2013; 2013:1.
31. Sattar N, Preiss D, Murray HM, *et al.* Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. Lancet 2010; 375:735–742.
32. Fall T, Xie W, Poon W, *et al.* Using genetic variants to assess the relationship ■ between circulating lipids and type 2 diabetes. Diabetes 2015; doi: 102337/db14-1710 [Epub ahead of print]
Systematic investigation trying to separate more specific lipid-related genetic variants from more general adiposity-related genetic variants and the impact of these sets of variants on T2D risk.
33. Kamstrup P, Tybjaerg-Hansen A, Stefensen R, Nordestgaard B. Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. J Am Med Assoc 2009; 301:2331–2339.
34. Clarke R, Peden J, Hopewell J, *et al.* Genetic variants associated with Lp(a) lipoprotein level and coronary disease. N Engl J Med 2009; 361:2518–2528.
35. Burgess S, Thompson S. Multivariable Mendelian randomization: the use of ■ pleiotropic genetic variants to estimate causal effects. Am J Epidemiol 2015; 181:251–260.
Method for estimating the causal effects of related risk factors that have common genetic predictors such as different lipid fractions.
36. Do R, Willer CJ, Schmidt EM, *et al.* Common variants associated with plasma triglycerides and risk for coronary artery disease. Nat Genet 2013; 45:1345–1352.
37. Burgess S, Freitag D, Khan H, *et al.* Using multivariable Mendelian randomization to disentangle the causal effects of lipid fractions. PLoS One 2014; 9:e108891.
38. Rosenson RS, Davidson MH, Hirsh BJ, *et al.* Genetics and causality of triglyceride-rich lipoproteins in atherosclerotic cardiovascular disease. J Am Coll Cardiol 2014; 64:2525–2540.
39. The TG and HDL Working Group of the Exome Sequencing Project. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. N Engl J Med 2014; 371:22–31.
40. Do R, Stitziel NO, Won HH, *et al.* Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. Nature 2014; 518:102–106.
41. Voight B, Peloso G, Orho-Melander M, *et al.* Plasma HDL cholesterol and risk of myocardial infarction: a Mendelian randomisation study. Lancet 2012; 380:572–580.
42. Holmes MV, Asselbergs FW, Palmer TM, *et al.* Mendelian randomization of ■ blood lipids for coronary heart disease. Eur Heart J 2015; 36:539–550.
Epidemiological investigation into the associations of various statistically-motivated genetic risk scores for lipid fractions with CHD risk.
43. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid ■■ instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015; 44:512–525.
Novel sensitivity analysis method for detecting bias from pleiotropy, and potentially correcting for this pleiotropy under a weaker set of assumptions than usually made in a Mendelian randomization investigation.
44. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in ■■ Mendelian randomization with some invalid instruments using a weighted median estimator. 2015; https://www.academia.edu/15479132/Consistent.
Another novel sensitivity analysis method for assessing whether a causal effect is evidenced by the majority of genetic variants, and is not influenced by the presence of a few invalid instrumental variables, with application of this method and the Mendelian randomization-Egger method to the analysis of the causal effects of major lipids on CHD.
45. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Brit Med J 1997; 315:629–634.
46. Pickrell, J. Fulfilling the promise of Mendelian randomization. Technical Report bioRxiv. 2015 [Epub ahead of print]. doi: 10.1101/018150
47. Kolesár M, Chetty R, Friedman J, *et al.* Identification and inference with many invalid instruments. J Bus Econ Stat 2015; 33:474–484.
48. Han C. Detecting invalid instruments using L1-GMM. Econ Lett 2008; 101:285–287.
49. Schwartz GG, Olsson AG, Abt M, *et al.* Effects of dalcetrapib in patients with a recent acute coronary syndrome. N Engl J Med 2012; 367:2089–2099.
50. Burgess S, Bowden J, Fall T, et al. Sensitivity analyses for robust causal ■ inference from Mendelian randomization analyses with multiple genetic variants. Technical Report. University of Cambridge. 2015.
Review of approaches for assessing the instrumental variable assumptions and performing sensitivity analyses for Mendelian randomization.
51. Suhre K, Raffler J, Kastenmüller G. Biochemical insights from population ■■ studies with genetics and metabolomics. Arch Biochem Biophys 2016; 589:168–176.
The study reviews findings from recent GWAS of metabolomics and provides concrete examples of how their results can be interpreted in a biochemical context.
52. Watson AD. Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. J Lipid Res 2006; 47:2101–2111.

53. Kastenmüller G, Raffler J, Gieger C, Suhre K. Genetics of human metabolism: ■■ an update. Hum Mol Genet 2015; 24:R93–R101.
The review provides a summary of the key aspects of GWAS with metabolomics and an update of recently published studies in this area.
54. Würtz P, Wang Q, Kangas AJ, *et al.* Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. PLoS Med 2014; 11:e1001765.
55. Shin SY, Fauman EB, Petersen AK, *et al.* An atlas of genetic influences on human blood metabolites. Nat Genet 2014; 46:543–550.
56. Ganna A, Salihovic S, Sundstrom J, *et al.* Large-scale metabolomic profiling ■ identifies novel biomarkers for incident coronary heart disease. PLoS Genet 2014; 10:e1004801.
In this study, results are presented from a Mendelian randomization analysis to determine whether the levels of four lipid metabolites have a causal role in the development of CHD.
57. Würtz P, Kangas AJ, Soininen P, *et al.* Lipoprotein subclass profiling reveals pleiotropy in the genetic variants of lipid risk factors for coronary heart disease: a note on Mendelian randomization studies. J Am Coll Cardiol 2013; 62:1906–1908.