Mendelian randomization of blood lipids for coronary heart disease

Michael V. Holmes[1,2]*†, Folkert W. Asselbergs[3,4,5]†, Tom M. Palmer[6]†, Fotios Drenos[7,8], Matthew B. Lanktree[9], Christopher P. Nelson[10,11], Caroline E. Dale[12], Sandosh Padmanabhan[13], Chris Finan[1], Daniel I. Swerdlow[1], Vinicius Tragante[3,14], Erik P.A. van Iperen[4,15], Suthesh Sivapalaratnam[16], Sonia Shah[17], Clara C. Elbers[14,18], Tina Shah[1], Jorgen Engmann[1], Claudia Giambartolomei[17], Jon White[17], Delilah Zabaneh[17], Reecha Sofat[1], Stela McLachlan[19], on behalf of the UCLEB consortium‡, Pieter A. Doevendans[3], Anthony J. Balmforth[20], Alistair S. Hall[20], Kari E. North[21], Berta Almoguera[18], Ron C. Hoogeveen[22], Mary Cushman[23,24], Myriam Fornage[25], Sanjay R. Patel[26], Susan Redline[26], David S. Siscovick[27,28], Michael Y. Tsai[29], Konrad J. Karczewski[30], Marten H. Hofker[31], W. Monique Verschuren[32], Michiel L. Bots[33], Yvonne T. van der Schouw[33], Olle Melander[34], Anna F. Dominiczak[13], Richard Morris[35], Yoav Ben-Shlomo[36], Jackie Price[19], Meena Kumari[1], Jens Baumert[37], Annette Peters[37], Barbara Thorand[37], Wolfgang Koenig[38], Tom R. Gaunt[8], Steve E. Humphries[7], Robert Clarke[39], Hugh Watkins[40], Martin Farrall[40], James G. Wilson[41], Stephen S. Rich[42], Paul I.W. de Bakker[14,33], Leslie A. Lange[43], George Davey Smith[8], Alex P. Reiner[44], Philippa J. Talmud[7], Mika Kivimäki[45], Debbie A. Lawlor[8], Frank Dudbridge[12], Nilesh J. Samani[10,11]†, Brendan J. Keating[2,18,46]†, Aroon D. Hingorani[1]†, and Juan P. Casas[1,12]†

[1]Genetic Epidemiology Group, Institute of Cardiovacular Science, Faculty of Population Healh Sciences, University College London, 1–19 Torrington Place, London WC1E 6BT, UK; [2]Division of Transplant Surgery, Perelman School of Medicine, University of Pennsylvania, 3400 Spruce St, Philadelphia, PA 19104, USA; [3]Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands; [4]Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, The Netherlands; [5]Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, UK; [6]Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK; [7]Centre for Cardiovascular Genetics, Institute of Cardiovascular Science, University College London, Rayne Building, London WC1E 6JF, UK; [8]MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK; [9]Department of Medicine, McMaster University, Hamilton, ON, Canada L8S 4L8; [10]Department of Cardiovascular Sciences, University of Leicester, Leicester, UK; [11]NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, UK; [12]Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK; [13]Institute of Cardiovascular and Medical Sciences, University of Glasgow, 126 University Place, Glasgow G12 8TA, UK; [14]Department of Medical Genetics, Division of Biomedical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands; [15]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, Amsterdam, The Netherlands; [16]Department of Vascular Medicine, Academic Medical Center, Amsterdam 1105 AZ, The Netherlands; [17]Department of Genetics Environment and Evolution, UCL Genetics Institute, 2nd Floor, Darwin Building, Gower Street, London WC1E 6BT, UK; [18]Center for Applied Genomics, Abramson Research Center, The Children's Hospital of Philadelphia, Philadelphia, PA, USA; [19]Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK; [20]Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, UK; [21]Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA; [22]Baylor College of Medicine, Department of Medicine, Division of Atherosclerosis and Vascular Medicine, Houston, TX 77030, USA; [23]Department of Medicine, University of Vermont, 208 South Park Dr, Colchester, VT 05446, USA; [24]Department of Pathology, University of Vermont, 208 South Park Dr, Colchester, VT 05446, USA; [25]The University of Texas Health Science Center at Houston, Houston, TX, USA; [26]Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA; [27]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA; [28]Department of Epidemiology, University of Washington, Seattle, WA, USA; [29]Department of Laboratory Medicine and Pathology, University of Minnesota; [30]Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA; [31]Department Pathology and Medical Biology, Medical Biology

* Corresponding author. Tel: +1 215 615 6578, Fax: +1 267 426 0363, Email: mvholmes@gmail.com

† Contributed equally (joint first and senior authors).

‡ UCLEB members are listed in Appendix.

Division, Molecular Genetics, University Medical Center Groningen and Groningen University, Groningen, The Netherlands; [32]Centre for Prevention and Health Services Research at National Institute for Public Health and the Environment, Utrecht, The Netherlands; [33]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands; [34]Lund University, Malmö, Sweden; [35]Department of Primary Care and Population Health, UCL, Royal Free Campus, Rowland Hill St, London, UK; [36]School of Social and Community Medicine, University of Bristol; [37]Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; [38]Department of Internal Medicine II-Cardiology, University of Ulm Medical Center, Ulm, Germany; [39]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK; [40]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK; [41]Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS 39216, USA; [42]Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA; [43]Department of Genetics, University of North Carolina School of Medicine at Chapel Hill, Chapel Hill, NC27514, USA; [44]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA; [45]Department of Epidemiology and Public Health, University College London, 1–19 Torrington Place, London WC1E 6BT, UK; and [46]Department of Paediatrics, School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Aims	To investigate the causal role of high-density lipoprotein cholesterol (HDL-C) and triglycerides in coronary heart disease (CHD) using multiple instrumental variables for Mendelian randomization.
Methods and results	We developed weighted allele scores based on single nucleotide polymorphisms (SNPs) with established associations with HDL-C, triglycerides, and low-density lipoprotein cholesterol (LDL-C). For each trait, we constructed two scores. The first was unrestricted, including all independent SNPs associated with the lipid trait identified from a prior meta-analysis (threshold $P < 2 \times 10^{-6}$); and the second a restricted score, filtered to remove any SNPs also associated with either of the other two lipid traits at $P \leq 0.01$. Mendelian randomization meta-analyses were conducted in 17 studies including 62,199 participants and 12,099 CHD events. Both the unrestricted and restricted allele scores for LDL-C (42 and 19 SNPs, respectively) associated with CHD. For HDL-C, the unrestricted allele score (48 SNPs) was associated with CHD (OR: 0.53; 95% CI: 0.40, 0.70), per 1 mmol/L higher HDL-C, but neither the restricted allele score (19 SNPs; OR: 0.91; 95% CI: 0.42, 1.98) nor the unrestricted HDL-C allele score adjusted for triglycerides, LDL-C, or statin use (OR: 0.81; 95% CI: 0.44, 1.46) showed a robust association. For triglycerides, the unrestricted allele score (67 SNPs) and the restricted allele score (27 SNPs) were both associated with CHD (OR: 1.62; 95% CI: 1.24, 2.11 and 1.61; 95% CI: 1.00, 2.59, respectively) per 1-log unit increment. However, the unrestricted triglyceride score adjusted for HDL-C, LDL-C, and statin use gave an OR for CHD of 1.01 (95% CI: 0.59, 1.75).
Conclusion	The genetic findings support a causal effect of triglycerides on CHD risk, but a causal role for HDL-C, though possible, remains less certain.
Keywords	Lipids • Heart disease • Mendelian randomization • Aetiology • Epidemiology

Introduction

The association of elevated low-density lipoprotein cholesterol (LDL-C) with coronary heart disease (CHD) events in observational studies has been established as causal based on randomized trials of LDL-C-lowering drugs.[1,2] In contrast, uncertainty exists on the causal relevance of high-density lipoprotein cholesterol (HDL-C) and triglycerides. Whereas observational studies indicate unambiguous associations of triglycerides and HDL-C with CHD (the association being positive for triglycerides and inverse for HDL-C),[3] randomized trials of HDL-C or triglyceride modifying drugs have not, so far, shown the anticipated benefit.[4–6]

These inconsistent findings may have arisen because the observational studies are affected by reverse causality[7] or by confounding (the latter would arise if HDL-C or triglyceride levels mark another causal risk factor without being causal themselves). Alternatively, the negative findings from clinical trials may have arisen from inadequate selection of drug targets or drug molecules.[6,8] Given this uncertainty, it remains unclear whether elevating HDL-C or reducing triglycerides by different means may still have utility for prevention of CHD events.

A further approach to evaluating the causal relevance of biomarkers that addresses these limitations is to exploit the natural randomized allocation of allelic variation in genes affecting their level (Mendelian randomization, outlined in Supplementary material online, *Figure S1*).[9,10] Unlike the directly observed associations of a risk factor with CHD events, genetic associations are protected from reverse causation because genotype is an invariant characteristic determined at conception and unmodified by the development of disease. Moreover, at a population level the randomized allocation of parental alleles at conception tends to balance confounding factors among groups of differing genotypes.[9,10] Where a polymorphism is associated with both risk factor concentration and CHD risk, this supports a causal role for the risk factor, providing certain other assumptions are met.[9]

Several Mendelian randomization studies have investigated the role of LDL-C,[11,12] HDL-C,[13–16] and triglycerides[17] in CHD. Most have used a single nucleotide polymorphism (SNP) from a single locus with weak, non-exclusive effects on the target lipid,[13–15,17] apart from a recent investigation of HDL-C.[16] For example, the association of SNPs in the APOA5 gene with CHD risk has been interpreted as implying a causal role for triglycerides;[17] however, it is more

informative on apolipoprotein A5 as a potential therapeutic target and the association of SNPs in the same gene with HDL-C and LDL-C leaves room for uncertainty.[18] Mendelian randomization analyses based on a single SNP with a non-exclusive association with a biomarker of interest may also lack generalizability. As one of several potential examples, the null association with CHD of an apparently HDL-C-specific SNP in the *LIPG* gene[16] only provided evidence that endothelial lipase (encoded for by *LIPG*) may not be a suitable drug target for CHD prevention, but it does not rule out the possibility that elevating HDL-C through a different drug target might reduce CHD risk.

Recent genetic association studies based on genotyping arrays that capture variation across many thousands of genes, or the whole genome, have indicated that SNPs associated with the major blood lipid fractions are distributed across many genetic loci, each inherited independently and affecting lipid levels approximately additively.[19−21] This provides a new opportunity to undertake Mendelian randomization analyses using multiple SNPs as instrumental variables (described in Supplementary material online, *Figure S2*). This should increase power, because each additional SNP contributes incrementally to the explained variance in the lipid fraction of interest and reduces the lack of specificity often observed with single SNPs, because the effects on traits other than the lipid fraction of interest should be small, non-systematic, and attenuate with the addition of SNPs to the instrument.[22]

In this study, we used multiple independent SNPs as instrumental variables in a Mendelian randomization approach. SNP selection was based on a previous study that we conducted to discover SNPs robustly associated with each blood lipid trait using the ITMAT Broad Institute CARe consortium (IBC) CardioChip array.[21,23] We summed values for individual SNPs to construct two types of allele scores. First, unrestricted allele scores were generated that included all SNPs that were associated with the target lipid trait at a pre-specified *P*-value threshold of $P < 2.4 \times 10^{-6}$. Secondly, restricted allele scores were generated in which SNPs were excluded if they were also associated with either of the other two lipid traits beyond a pre-specified *P*-value threshold of $P \leq 0.01$. Our study incorporates individual participant data, investigates all three lipid traits, and use of lipid-lowering medication in the same data set for their associations with clinically defined and validated CHD events, compares and contrasts associations of both unrestricted and restricted allele scores, which has different underlying assumptions, and applies newly developed methods for instrumental variables meta-analysis that enables inclusion of case−control studies and adjustment for other covariates in the analysis.[22,24]

Methods

Included studies

We analysed data from 17 studies including 62 199 individuals of European origin: 13 longitudinal population studies, 1 case-cohort study, 1 nested case−control study, and 2 case−control studies. Characteristics of the study participants are provided in Supplementary material online, *Table S1*. Altogether there were 12 099 incident or prevalent CHD cases in the study sample.

Single nucleotide polymorphism selection and construction of the allele scores for Mendelian randomization

We based SNP selection on a large-scale gene-centric discovery meta-analysis of blood lipid traits that included 66 240 individuals[21] genotyped with the IBC CardioChip array.[23] We identified all SNPs that met the pre-defined array-wide threshold value of $P < 2.4 \times 10^{-6}$ for the target lipid in the original report.[21] To avoid co-linearity between SNPs, if more than one SNP was present at a gene locus, only the SNP with the lowest *P*-value for the target lipid trait was included in the allele score.

All SNPs passing the *P*-value threshold in the discovery analysis that were in unique loci were incorporated into the analysis. These selected SNPs were used to generate allele scores (summed values of genetic variants, also termed 'genetic instruments') for each individual in the participating studies for the blood lipid traits HCL-C, triglycerides, and LDL-C. We followed this process in order to be able to conduct a Mendelian randomization analysis of blood lipid traits. The advantage of our approach for the identification of SNPs for the genetic instruments was that identified SNPs would be hypothesis-free rather than being selected on a candidate basis through biological understanding. Combining multiple SNPs together increases power of the Mendelian randomization analysis,[25] but additionally helps to address questions of causality for traits that are not directly encoded by any particular gene. We weighted SNPs in each allele score by the published summary beta coefficients from the discovery gene-centric meta-analysis[21] and selected the 'risk' allele such that the associations with the target lipid trait were directionally concordant. The use of weighting was to increase precision of the genetic instrument with the intermediate trait.[25] The weighted values of SNPs were summed to generate an allele score value for each individual.

Blood lipid traits share common genetic variants resulting in overlap of the SNPs identified in the discovery analysis (Supplementary material online, *Figure S3*). This means that allele scores generated for, e.g. HDL-C using all identified SNPs from the discovery analysis would also include SNPs that associate with LDL-C and triglycerides. This could be interpreted as non-specificity of the genetic instrument for the target blood lipid trait. To try and resolve this issue, we took the following approach. First, we generated what we termed an 'unrestricted allele score' that included all SNPs that were associated with the target lipid trait, regardless of any association with other blood lipid traits. Secondly, we generated a 'restricted' allele score that included SNPs exhibiting an association with the target lipid trait but which did not show an association with the other two lipid traits at $P < 0.01$. We compared the estimates derived from Mendelian randomization analysis using unrestricted and restricted allele scores as instrumental variables in order to try and decipher the individual role of blood lipid traits in CHD pathogenesis. The analytical pipeline for construction of the allele scores is outlined in Supplementary material online, *Figure S4*.

For HDL-C, 48 SNPs in 48 independent genes/loci showed association with HDL-C, 29 of which also showed association with triglycerides or LDL-C. The unrestricted score for HDL-C therefore consisted of all 48 SNPs and the restricted allele score comprised the 19 SNPs that did not associate with either triglycerides or LDL-C. Sixty-seven SNPs associated with triglycerides, of which 40 also showed association for LDL-C or HDL-C. The unrestricted triglyceride allele score therefore consisted of 67 SNPs and the restricted allele score contained 27 SNPs that did not associate with HDL-C or LDL-C. Forty-two SNPs associated with LDL-C, of which 23 also associated with triglycerides or HDL-C. The unrestricted LDL-C allele score therefore consisted of 42 SNPs and the restricted LDL-C allele score 19 SNPs. Full details of the SNPs used in

each of the unrestricted and restricted lipid scores are presented in Supplementary material online, *Table S2* and the allele frequencies are displayed in Supplementary material online, *Figures S5−7*. Allele score distributions were normal in each study (Supplementary material online, *Figure S8* and *Table S3*).

Platforms used for genotyping

In 13 of the 17 studies, genotyping was conducted with the IBC Cardio-Chip array and the four remaining studies were genotyped using the MetaboChip[26] (Supplementary material online, *Table S1*). In these studies, we used Metabochip SNPs in linkage disequilibrium (LD) ($R^2 > 0.8$) with those derived from the IBC CardioChip using pair-wise LD calculated from the European subset of the 1000 Genomes Project[27] (http://www.1000genomes.org). Suitable proxies were identified for 135 of 157 total SNPs used to construct the allele scores (Supplementary material online, *Table S2*).

Outcomes

The principal outcome of interest for Mendelian randomization analysis was the combination of incident or prevalent CHD events, but we conducted a subsidiary analysis limited to incident CHD cases (i.e. cases accrued during the follow-up, predominantly after the measurement of blood lipid traits). As a secondary endpoint measure, we analysed carotid intima media thickness (cIMT), which is associated with CHD risk and has been used as a surrogate endpoint in phase II randomized trials of lipid-lowering therapies.[28] Details on outcome ascertainment for each study are provided in Supplementary material online, *Table S4*.

Data handling

Non-normally distributed traits (e.g. triglycerides and cIMT) were log transformed prior to analysis and summary estimates were exponentiated and converted to a percentage difference in the geometric mean. Missing values for genotype or phenotype data were not imputed.

Analysis

The analysis was standardized and run in individual participant data in all contributing studies (Supplementary material online, *Figure S9* for the data analysis pipeline).

Quantifying the association of the allele scores with blood lipid traits

In the 11 general population cohorts that were genotyped using the IBC CardioChip array (Supplementary material online, *Table S1*), to quantify the magnitude of the association between the allele scores and lipid traits, the mean difference and standard error for each lipid trait was estimated comparing the top quintile of each allele score to the bottom quintile. The proportion of variance (R^2) of the allele scores for each lipid trait was estimated within each study, with the 95% confidence interval (CI) of the R^2 obtained through bootstrapping. Estimates were pooled using fixed-effects meta-analysis.

Multiple single nucleotide polymorphism instrumental variable analysis

Instrumental variable analysis is a statistical method used to obtain unbiased estimates between an exposure and an outcome, which exploits the characteristic of the instrument, which is assumed to be free from common confounding. Use of SNPs as instrumental variables is an established technique termed Mendelian randomization that has been used to investigate the causal relationship between many biomarkers and outcomes (outlined in Supplementary material online, *Figure S1*). Our analysis here extends this to incorporate multiple SNPs in combination, an

emerging approach that is gaining traction as a means of investigating non-protein traits and to increase power (described in Supplementary material online, *Figure S2*).

Our instrumental variable analysis took two forms:

(1) Instrumental variable analysis: incorporating data from all studies

For the main analysis, we used an approach that allowed us to incorporate data to maximize power from all 17 studies: 15 prospective studies with measures of blood lipid traits and 16 studies (including two case−control studies) with CHD events (one study, CARDIA, did not contribute to CHD events, Supplementary material online, *Table S1*). For this, we investigated the association of the allele scores for each target lipid trait. This was limited to the 15 prospective cohort studies in which blood lipids were measured at baseline, when most individuals were free from established disease (since the disease process may distort the association of the allele scores with blood lipid levels). We pooled the estimates of the allele scores with blood lipid traits across studies using fixed-effects meta-analysis and used this pooled summary estimate for the second stage of the instrumental variable analysis. This technique assumed a constant effect of the allele score on the target lipid trait. For the second stage, we generated associations between each allele score and CHD in each study. The instrumental variable estimate was then obtained by dividing the allele score−CHD association by the pooled allele score−lipid estimate.[29] This analysis took into account the uncertainty in both the allele score−CHD and allele score−lipid associations using the delta method to estimate standard errors of instrumental variable ratio estimates.[30] These values were then pooled across studies using fixed-effects meta-analysis. This approach was conducted using both unrestricted and restricted allele scores as the instrumental variables for the lipid traits.

(2) Instrumental variable analysis with sequential adjustments using longitudinal cohorts.

Separately, we conducted another Mendelian randomization analysis in an additional attempt to address the lack of specificity of the *unrestricted* allele scores. For this, we conducted an instrumental variable Mendelian randomization analysis using the logistic control function estimator[24] in each study using the unrestricted allele scores as the instrumental variable. The logistic control function estimator is a two-stage process: first, a linear regression analysis is conducted with the target lipid trait as the dependent variable and the unrestricted allele score as the independent variable. The residuals from this first step, along with the target lipid trait, are then incorporated into a logistic regression model in the second stage in which incident/prevalent CHD is the dependent variable. Robust standard errors are specified in the second stage to incorporate the uncertainty in the first-stage residuals. We pooled study-specific instrumental variable estimates across studies using fixed-effects meta-analysis.

Initially, the instrumental variable analyses using this method were conducted unadjusted. We then made sequential adjustments for non-target lipid traits (e.g. for LDL-C we adjusted for HDL-C, triglycerides, and statin use). This approach required that contributing studies had the co-variables of interest, so case−control studies or longitudinal studies without this information were not included, meaning that the sample size was reduced. Thus, the analysis was limited to 14 longitudinal studies.

For cIMT (measured in four prospective cohorts, Supplementary material online, *Table S4*), we used two-stage least squares analysis using the unrestricted and restricted allele scores as instrumental variables in separate models. We pooled study-specific instrumental variable estimates across studies using fixed-effects meta-analysis.

The summary instrumental variable estimates for both the main and subsidiary Mendelian randomization analyses provided an odds ratio

(OR) for CHD or percentage difference in cIMT per 1 unit increase in a genetically instrumented blood lipid trait (i.e. per 1 mmol/L increase in HDL-C or LDL-C, which both have a normal distribution and for a 1 log-unit increase in triglycerides which has a log-normal distribution).

Fasting status was noted for blood lipid measures (Supplementary material online, *Table S5*). To investigate the influence of fasting status on the association of the allele scores with the lipid traits, we conducted a sensitivity analysis by stratifying on fasting status. Furthermore, we excluded non-fasting studies from the first stage of the instrumental variable analysis to examine whether this influenced the instrumental variable estimates for CHD.

Analyses were conducted using Stata v13.1 (StataCorp, TX, USA). We took two-sided *P*-values ≤0.05 to denote evidence against the null hypothesis.

Results

Across 17 studies with 62 199 individuals of European ancestry, there were 12 099 combined incident and prevalent CHD events of which 7339 were incident, and 9942 measures of cIMT, a non-invasive measure of atherosclerosis (Supplementary material online, *Table S1*). For the prospective cohorts, mean values of blood lipid traits, the proportion of individuals receiving lipid-lowering therapy, and whether samples were obtained when individuals were fasting are reported in Supplementary material online, *Table S5*. As expected, each SNP in the allele scores was associated individually with directionally concordant effects on the target lipid in prospective cohorts genotyped using the IBC CardioChip (Supplementary material online, *Figure S10*). There was a partial overlap of SNPs among the three unrestricted allele scores (Supplementary material online, *Figure S3*). By definition, SNPs in the restricted allele scores were non-overlapping.

The associations of each allele score for the target and non-target lipid traits are shown in *Figure 1*. The unrestricted allele scores consistently showed a larger magnitude of effect and explained more variance for the target lipid than the corresponding restricted allele scores. For example, the HDL-C unrestricted allele score was associated with higher HDL-C by 0.23 mmol/L (95% CI: 0.22, 0.24, comparing top to bottom quintiles of the allele score), explaining 3.8% of its variance. The comparable difference for the restricted HDL-C
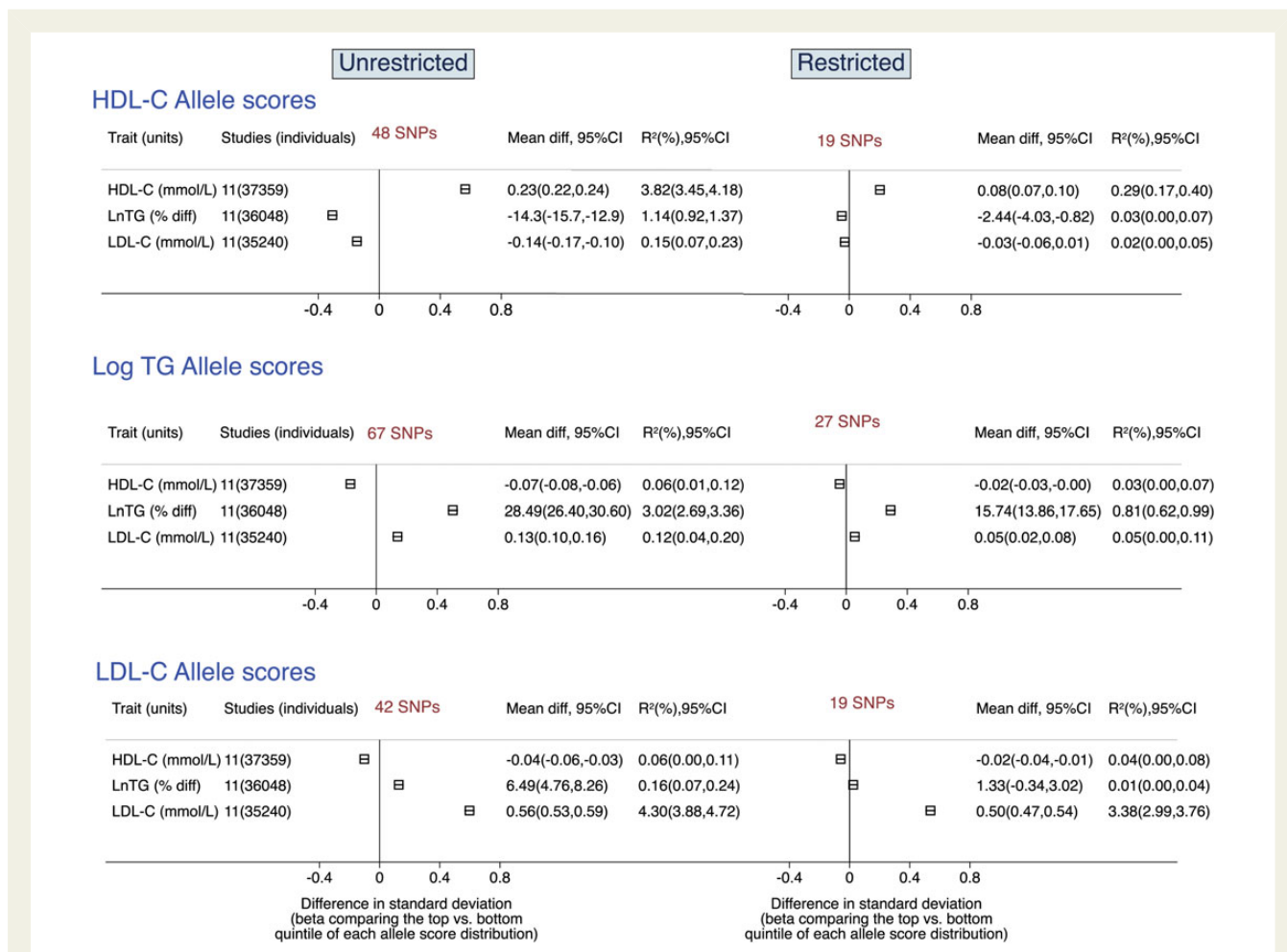
Figure 1 Meta-analysis pooled estimates of the association of the unrestricted and restricted allele scores with target and non-target lipid traits. Estimates were obtained from prospective cohorts genotyped using the ITMAT Broad Institute CARe consortium CardioChip array (detailed in Supplementary material online, *Table S1*). A lower limit of 0 was imposed on the R^2 values. Mean diff, mean difference comparing top to bottom quintile of each allele score. R^2 = proportion of variable of the lipid traits explained by each allele score. TG, triglycerides.

allele score was 0.08 mmol/L (95% CI: 0.07, 0.10), explaining only 0.3% of the variance. Corresponding values for triglycerides and LDL-C allele scores are presented in *Figure 1*. In addition to the association with the target lipid traits, each of the three unrestricted allele scores also showed association with non-target lipid traits (values reported in *Figure 1*). In contrast, the restricted allele scores consistently explained a smaller proportion of variance for *non-target* lipid traits. Stratification of the association of the allele scores with blood lipid traits by fasting status did not show heterogeneity in the estimates with the exception of the values for the restricted allele score for LDL-C; however, this did not influence the overall estimate (Supplementary material online, *Figure S11*).

For LDL-C, in 16 cohort/case−control studies with 11 826 combined incident/prevalent CHD cases, a 1 mmol/L genetically instrumented increment in LDL-C gave an OR for CHD of 1.78 (95% CI: 1.58, 2.01) for the unrestricted, and 1.92 (95% CI: 1.68, 2.19) for the restricted allele score (*Figure 2*). For HDL-C, using the unrestricted allele score a 1 mmol/L genetically instrumented increment in HDL-C yielded an OR for CHD of 0.53 (95% CI: 0.40, 0.70), but the comparable estimate for the restricted allele score was 0.91 (95% CI: 0.42, 1.98). For triglycerides, a genetically instrumented 1 log-unit increment in triglycerides yielded similar estimates for CHD events: an OR of 1.62 (95% CI: 1.24, 2.11) for the unstricted score and 1.61 (95% CI: 1.00, 2.59) for the restricted score. Estimates derived from instrumental variable analysis using incident-only CHD cases were comparable in effect size and direction to those from the analyses incorporating the combined incident and prevalent events (*Figure 2*). There was a similar inconsistency in the effect estimate of the unrestricted allele score for HDL-C and risk of incident-only CHD (OR: 0.68 per 1 mmol/L lower HDL-C; 95% CI: 0.47, 0.97) and that for the restricted HDL-C allele score with incident-only CHD (OR: 1.33; 95% CI: 0.49, 3.59).

For each of the restricted and unrestricted allele scores, no difference was identified when the analysis was limited to fasted samples for the first stage of the instrumental variable analysis (Supplementary material online, *Figure S12*).

Sequential adjustment of the unrestricted LDL-C allele score for HDL-C, triglycerides, and statin use only moderately diminished the estimate for the association with CHD events (*Figure 3*), but comparable adjustments had more marked effects on the estimates for the HDL-C allele score. The association of the unrestricted HDL-C allele score with incident/prevalent CHD was shifted from an OR for CHD of 0.55 (95% CI: 0.38, 0.79) on unadjusted analysis to an OR of 0.79 (95% CI: 0.47, 1.32) with adjustment for triglycerides alone (*Figure 3*). In contrast, adjustment for LDL-C alone did not influence the estimate (OR: 0.52; 95% CI: 0.34, 0.78). When adjusted for triglycerides, LDL-C, and statin therapy, the OR for the association of the unrestricted HDL-C allele score with incident and prevalent CHD was 0.81 (95% CI: 0.44, 1.46), which was comparable with the estimates derived from the restricted allele score (OR: 0.91; 95% CI: 0.42, 1.98, *Figure 2*). For triglycerides, adjustment for HDL-C diminished the estimate for CHD risk from an OR of 1.38 (95% CI: 0.98, 1.94) for the unadjusted allele score to an OR of 0.97 (95% CI: 0.64. 1.49). Adjustment for LDL-C produced only a small alteration in the summary estimate for CHD risk: OR: 1.31 (95% CI: 0.86, 1.98). With adjustment for HDL-C, LDL-C,

and statin use the OR estimate for the unrestricted triglyceride allele score with incident and prevalent CHD was 1.01 (95% CI: 0.59, 1.75).

Only the LDL-C allele scores showed association with cIMT. A 1 mmol/L genetically instrumented increment in LDL-C was associated with higher cIMT by 2.49% (95% CI: 0.45, 4.57) and 3.81% (1.48, 6.19) for the unrestricted and restricted allele scores, respectively. Estimates for other lipid traits are provided in *Figure 4*.

Discussion

This Mendelian randomization analysis was based on individual participant level data including 62 199 individuals from 17 studies and used a multiple SNP instrumental variable meta-analysis approach. We reconfirmed the causal role of LDL-C in CHD risk and provided additional support for a causal role of triglycerides in CHD. The causal association of HDL-C with CHD remains possible, but less certain.

A key problem in trying to understand the causal relevance of HDL-C and triglycerides in CHD risk has been the close epidemiological and biological interrelationship between the two. Both associate with CHD events in observational studies, yet statistical adjustment for one attenuates the association of the other.[3] Incomplete biological understanding makes interpretation of this observational evidence challenging. Multiple instrument Mendelian randomization studies utilizing SNPs affecting the levels of these two traits offer a new route to understand their causal relevance and many such SNPs have been identified by recent genome-wide and gene-centric association studies,[19,21] including the set of SNPs used in the present analysis.[21] Although, multiple instruments Mendelian randomization analysis reduces the non-specificity, it does not abolish it. For this reason, we generated two different allele scores. First, an unrestricted score that includes all genetic determinants of each lipid trait, which can be conceived as being more comprehensive in biological terms, as well as more powerful (e.g. R^2 of unrestricted score for HDL-C was 3.8%). In contrast, the restricted score, though substantially increasing specificity for the target lipid, is both less biologically comprehensive and statistically less powerful (e.g. R^2 of restricted score for HDL-C was 0.3%). Owing to these limitations, we also undertook instrumental variable analyses using the unrestricted scores in which adjustments were made for the non-target lipids. We then compared the effect estimates from these different approaches to draw inferences on the causal role of HDL-C and triglycerides using LDL-C, whose aetiological role in CHD is established, as a positive control. This strategy, comparing the consistency of potentially causal estimates derived from instrumental variable analysis that used three different approaches, each of them with different underlying assumptions, in individual participant data sets, we believe has not been employed before and thus represents a novel aspect of the current analysis.

The estimates of LDL-C from instrumental variable analysis showed that a long-term genetically increased LDL-C, regardless of the analytical strategy used (unrestricted, restricted, or unrestricted score plus sequential adjustments) resulted in an increased causal OR for CHD, which is similar in magnitude to that reported in randomized trials of statin-lowering therapies in individuals at low risk of vascular disease[1] and is further evidence of the validity of our
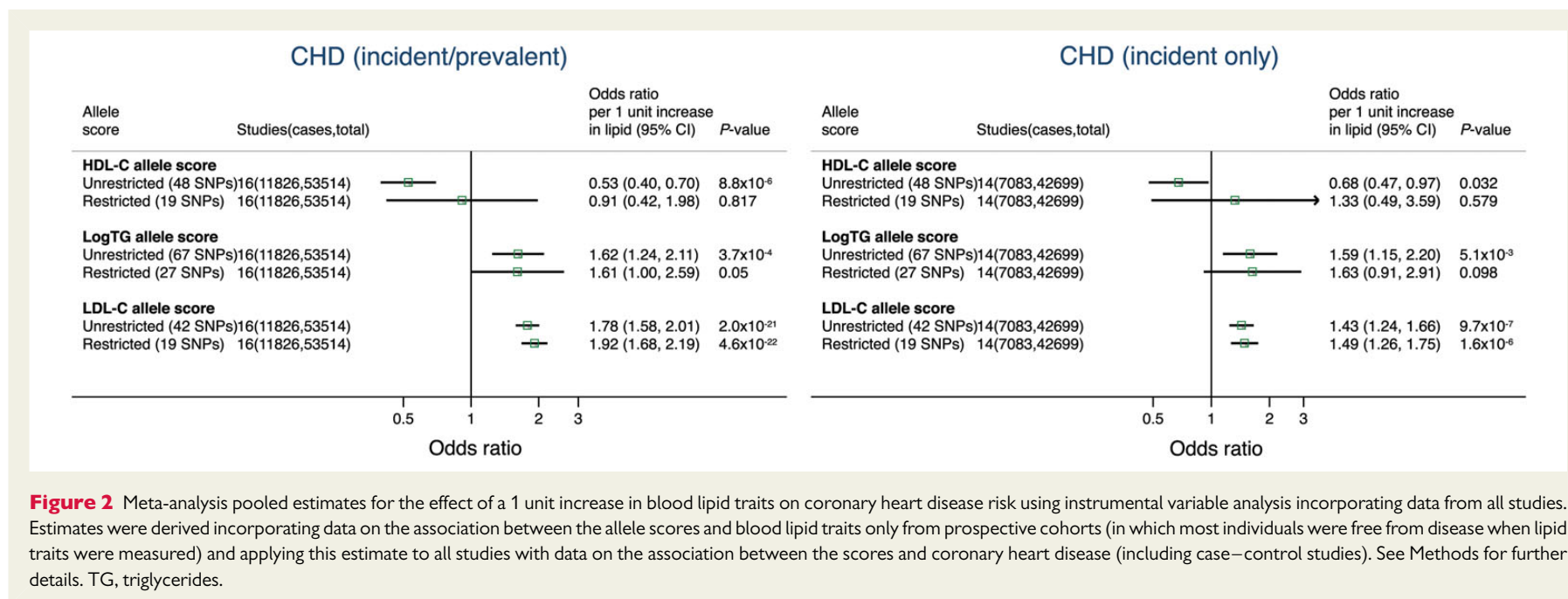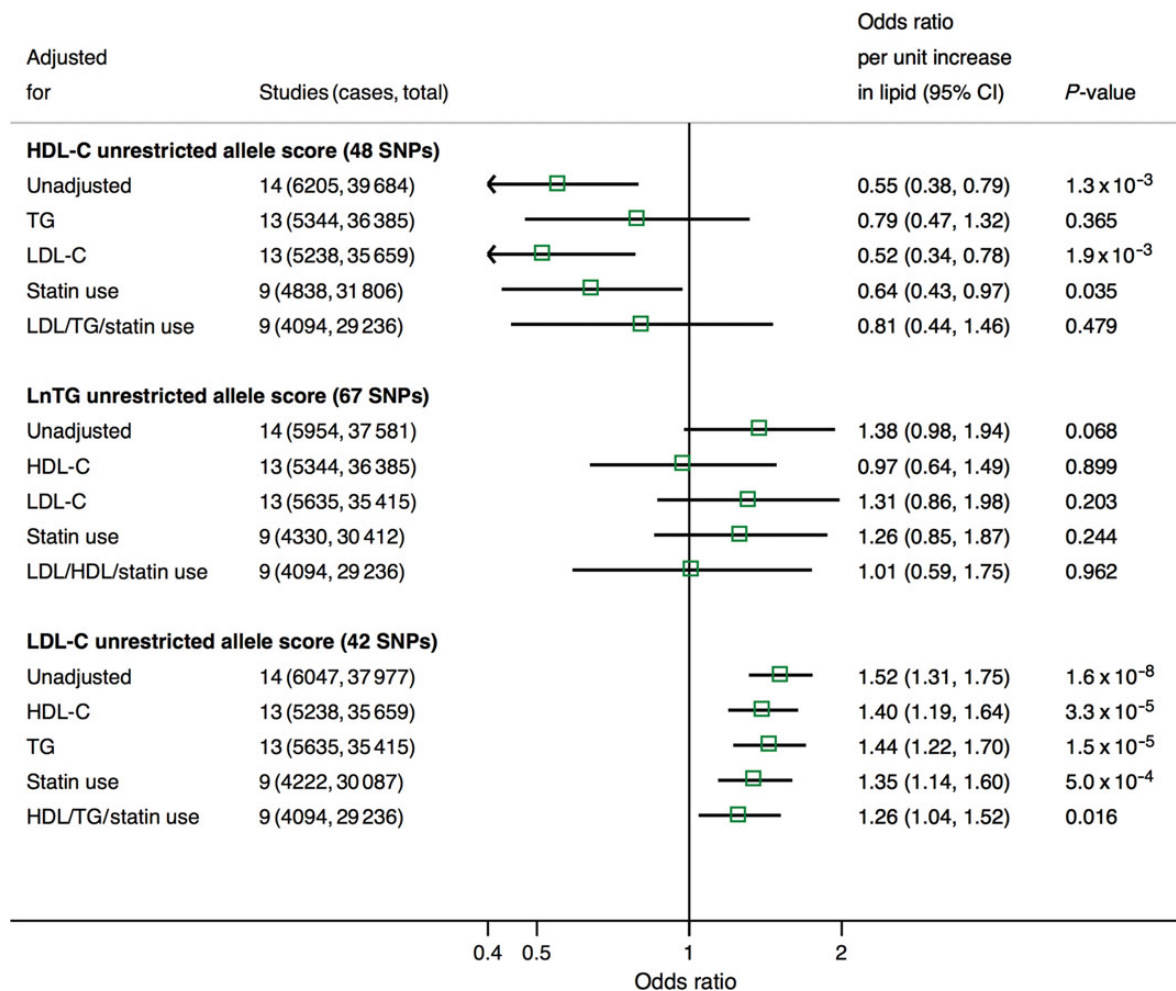
Figure 2 Meta-analysis pooled estimates for the effect of a 1 unit increase in blood lipid traits on coronary heart disease risk using instrumental variable analysis incorporating data from all studies. Estimates were derived incorporating data on the association between the allele scores and blood lipid traits only from prospective cohorts (in which most individuals were free from disease when lipid traits were measured) and applying this estimate to all studies with data on the association between the scores and coronary heart disease (including case–control studies). See Methods for further details. TG, triglycerides.

Figure 3 Meta-analysis pooled estimates for the effect of a 1 unit increase in blood lipid traits on combined incident/prevalent coronary heart disease risk using instrumental variable analysis with the unrestricted allele score, adjusted for non-target traits and statin use. Analysis was conducted in prospective cohorts with instrumental variables regression analysis. TG, triglycerides.

various analytical approaches. The instrumental variable analysis of LDL-C on cIMT is also in keeping with recent findings,[31] and supports the use of cIMT as an appropriate surrogate marker of therapies that modulate LDL-C.

For triglycerides, the findings for the unrestricted and restricted allele scores were concordant, with both showing association with CHD. However, the unrestricted score adjusted for HDL-C diminished the association to null. Thus, two out of the three approaches provided evidence of a causal role of triglycerides in CHD, making it likely that triglycerides are causally related to CHD. It is intriguing that the association of the unrestricted score for triglycerides with CHD events diminished to null when adjusted for HDL-C. This could mean that a treatment that targets a triglyceride pathway that has no effect on HDL-C may not be beneficial, whereas a treatment that targets a triglyceride pathway that both reduces triglycerides and increases HDL-C could have a role in prevention of CHD events. An alternative explanation is that HDL-C could mark long-term triglyceride concentrations, but this hypothesis requires further investigation. As recently suggested by Wurtz *et al.*[32] in response to a Mendelian

randomization analysis of remnant cholesterol by Varbo *et al.*,[33] access to metabolomics data will enable partitioning of triglyceride containing lipoproteins according to size and composition (e.g. apolipoprotein B content) and facilitate investigation of the role of these subcomponents individually in CHD pathogenesis.

For HDL-C, only one of the approaches provided evidence that genetic determinants of HDL-C are causally related to CHD. The unrestricted HDL-C allele score (which did not impose constraints on the pathways that the genes in the allele score encode for) showed strong evidence of an association with CHD. But this unrestricted HDL-C allele score also showed association with triglycerides (and to a lesser extent LDL-C). In contrast, the restricted HDL-C allele score did not show an association with CHD. The restricted HDL-C allele score was more selective for HDL-C (showing only a very weak association with triglycerides and no effect on LDL-C), but also explained less of the variance of the index trait, HDL-C (even when compared with other restricted scores), so it remains uncertain if this attenuation in the effect estimate implies that an intervention that solely modifies HDL-C would not reduce risk of CHD,
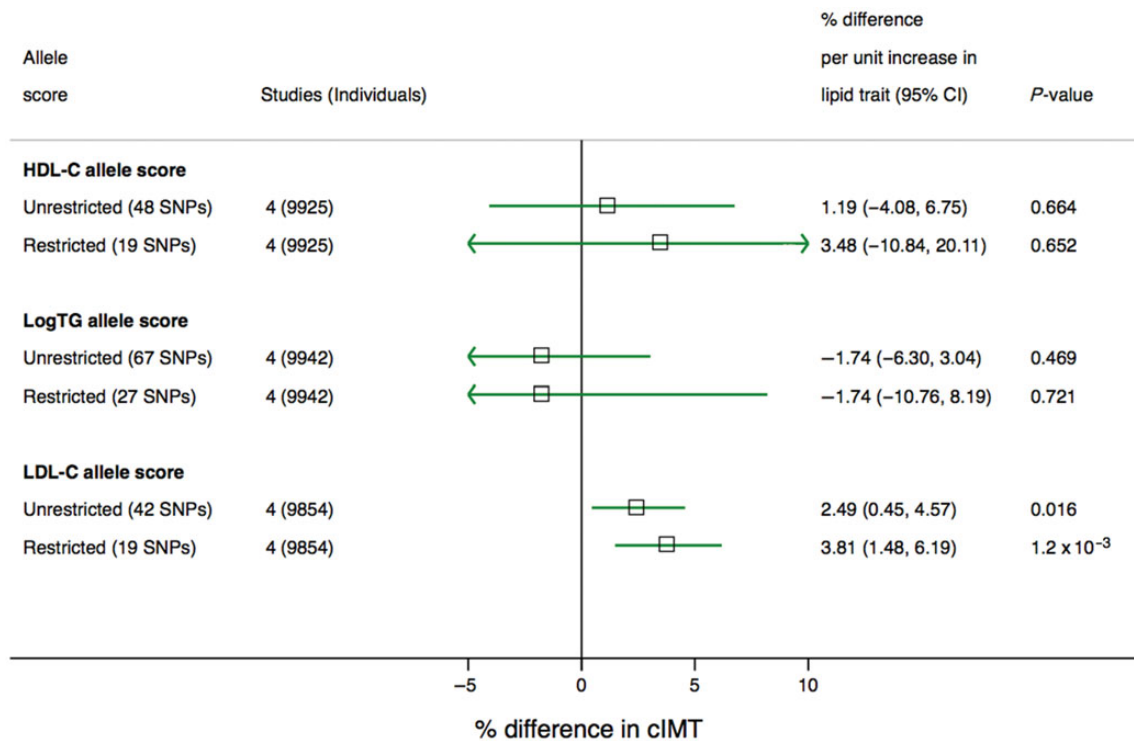
Figure 4 Meta-analysis pooled estimate for the effect of a 1 unit increase in blood lipid traits on carotid intima medial thickness. The four population-based prospective cohorts with carotid intima medial thickness traits were CHS, FHS, MESA, and Whitehall II (Supplementary material online, *Table S4*).

or whether it is due to a reduction in statistical power. This former interpretation is in agreement with findings from our unrestricted allele score adjusted for triglycerides, and with a previous multiple SNPs Mendelian randomization analysis that, using different genetic instruments (Supplementary material online, *Figure S13*), also failed to identify a clear causal role of HDL-C in CHD.[16]

Our study has a number of possible limitations. First, of the 17 contributing studies, 13 were a subsample of the 32 studies that contributed towards the gene-centric discovery meta-analysis on blood lipid traits.[21] Thus, it is theoretically possible that using a partially overlapping set of studies for the discovery and Mendelian randomization analysis may potentially result in model over-fitting. Secondly, our allele scores were designed to proxy total levels of blood lipid and lipoprotein traits, and therefore do not address whether there are subtypes of these traits (e.g. HDL subparticles)[34] that could play contrasting roles in vascular disease. For example, we cannot exclude the possibility that the restricted HDL-C allele score may have lacked genes that are present in the unrestricted allele score that encode subparticles of HDL that *do* have a causal role in CHD. This requires further investigation with Mendelian randomization using SNPs or allele scores that are specific for HDL subtypes. Thirdly, it is possible that some of the null findings could be due to limited power, including the analysis for cIMT. Examination of these findings in other data sets is therefore warranted.

In conclusion, the findings from a multiple SNP Mendelian randomization analysis in over 62 000 participants with >12 000 CHD events support a causal effect of triglycerides but evidence on the causal role, if any, of HDL-C on CHD risk remains uncertain.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Appendix: The UCLEB Consortium

T.S. (Department of Epidemiology and Public Health, UCL Institute of Epidemiology and Health Care, University College London, 1-19 Torrington Place, London WC1E 6BT, UK); J.E. (Department of Epidemiology and Public Health, UCL Institute of Epidemiology and Health Care, University College London, 1-19 Torrington Place, London WC1E 6BT, UK); C.E.D. (Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK); S.S. (University College London Genetics Institute, Department of Genetics, Environment and Evolution, Gower St, London WC1E 6BT, UK); J.W. (University College London Genetics Institute, Department of Genetics, Environment and Evolution, Gower St, London WC1E 6BT, UK); C.G. (University College London Genetics Institute, Department of Genetics, Environment and Evolution, Gower St, London WC1E 6BT, UK); S.M.L. (Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK); D.Z. (University College London Genetics Institute, Department of Genetics, Environment and Evolution, Gower St, London WC1E 6BT, UK); Alana Cavadino (MRC Centre of Epidemiology for Child Health, Department of Population Health Sciences, UCL; Institute of Child Health, University College London, 30 Guilford Street, London WC1N 1EH, UK); C.F. (Department of Epidemiology and Public Health, UCL Institute of Epidemiology and Health Care, University College London, 1-19 Torrington Place, London WC1E 6BT, UK); Andrew Wong (MRC Unit for Lifelong Health and Ageing, 33 Bedford Place, London WC1B 5JU, UK); Antoinette Amuzu (Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK); Ken Ong (MRC Unit for Lifelong Health and Ageing, 33 Bedford Place, London WC1B 5JU, UK; MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Box 285, Cambridge CB2 0QQ, UK); T.R.G. (MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK); M.V.H. (Department of Epidemiology and Public Health, UCL Institute of Epidemiology and Health Care, University College London, 1-19 Torrington Place, London WC1E 6BT, UK); Helen Warren (Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK); D.I.S. (Department of Epidemiology and Public Health, UCL Institute of Epidemiology and Health Care, University College London, 1-19 Torrington Place, London WC1E 6BT, UK); Teri-Louise Davies (MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community; Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK); F.D. (Centre for Cardiovascular Genetics, Department of Medicine, British Heart Foundation Laboratories, Rayne Building, Royal Free and University College Medical School, 5 University Street, London, WC1E 6JF, UK); J.C. (Centre for Cardiovascular Genetics, Department of Medicine, British Heart Foundation Laboratories, Rayne Building, Royal Free and University College Medical School, 5 University Street, London, WC1E 6JF, UK); R.S. (Centre for Clinical Pharmacology, University College London, London WC1E 6JF, UK), M.C. (William Harvey Research Institute, Barts and the London Queen Mary's School of Medicine and Dentistry, John Vane Building, Charterhouse Square, London EC1M 6BQ, UK); Shah Ebrahim (Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK); D.A.L. (MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community; Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK); P.J.T. (Centre for Cardiovascular Genetics, Department of Medicine, British Heart Foundation Laboratories, Rayne Building, Royal Free and University College Medical School, 5 University Street, London, WC1E 6JF, UK); S.E.H. (Centre for Cardiovascular Genetics, Department of Medicine, British Heart Foundation Laboratories, Rayne Building, Royal Free and University College Medical School, 5 University Street, London, WC1E 6JF, UK); Christine Power (MRC Centre of Epidemiology for Child Health, Department of Population Health Sciences, UCL; Institute of Child Health, University College London, 30 Guilford Street, London WC1N 1EH, UK); Elina Hypponen (MRC Centre of Epidemiology for Child Health, Department of Population Health Sciences, UCL; Institute of Child Health, University College London, 30 Guilford Street, London WC1N 1EH, UK); R.M. (MRC Unit for Lifelong Health and Ageing, 33 Bedford Place, London WC1B 5JU, UK); Rebecca Hardy (MRC Unit for Lifelong Health and Ageing, 33 Bedford Place, London WC1B 5JU, UK); Diana Kuh (MRC Unit for Lifelong Health and Ageing, 33 Bedford Place, London WC1B 5JU, UK); Nicholas Wareham (MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Box 285, Cambridge CB2 0QQ, UK); Claudia Langenberg (MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Box 285, Cambridge CB2 0QQ, UK; Department of Epidemiology and Public Health, UCL Institute of Epidemiology and Health Care, University College London, 1-19 Torrington Place, London WC1E 6BT, UK); Yoav BenShlomo (School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK); I.N.D. (MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community; Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK); Peter Whincup (Division of Population Health Sciences and Education, St George's, University of London, Cranmer Terrace, London SW17 0RE, UK); R.M. (Department of Primary Care and Population Health, University College London, Royal Free Campus, Rowland Hill Street, London NW3 2PF, UK); J.P. (Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK); ME.K. (Department of Epidemiology and Public Health, UCL Institute of Epidemiology and Health Care, University College London, 1-19 Torrington Place, London WC1E 6BT, UK); MI.K. (Department of Epidemiology and Public Health, UCL Institute of Epidemiology and Health Care, University College London, 1-19 Torrington Place, London WC1E 6BT, UK); V.P. (University College London Genetics Institute, Department of Genetics, Environment and Evolution, Gower St, London WC1E 6BT, UK); F. D. (Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK); J.C.W. (Genetics Division, Research and Development, GlaxoSmithKline, NFSP, Harlow CM19 5AW, UK); J.P.C. (Department of Epidemiology and Public Health; UCL Institute of Epidemiology and Health

Care, University College London, 1-19 Torrington Place, London WC1E 6BT, UK; Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK); A.D.H. (Department of Epidemiology and Public Health, UCL Institute of Epidemiology and Health Care, University College London, 1-19 Torrington Place, London WC1E 6BT, UK; Centre for Clinical Pharmacology, University College London, London WC1E 6JF, UK).

References

1. Mihaylova B, Emberson J, Blackwell L, Keech A, Simes J, Barnes EH, Voysey M, Gray A, Collins R, Baigent C. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. *Lancet* 2012;**380**:581–590.

2. Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, Kirby A, Sourjina T, Peto R, Collins R, Simes R. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet* 2005;**366**:1267–1278.

3. Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray KK, Thompson A, Wood AM, Lewington S, Sattar N, Packard CJ, Collins R, Thompson SG, Danesh J. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA* 2009;**302**:1993–2000.

4. Schwartz GG, Olsson AG, Abt M, Ballantyne CM, Barter PJ, Brumm J, Chaitman BR, Holme IM, Kallend D, Leiter LA, Leitersdorf E, McMurray JJ, Mundl H, Nicholls SJ, Shah PK, Tardif JC, Wright RS. Effects of dalcetrapib in patients with a recent acute coronary syndrome. *N Engl J Med* 2012;**367**:2089–2099.

5. Keech A, Simes RJ, Barter P, Best J, Scott R, Taskinen MR, Forder P, Pillai A, Davis T, Glasziou P, Drury P, Kesaniemi YA, Sullivan D, Hunt D, Colman P, d'Emden M, Whiting M, Ehnholm C, Laakso M, investigators Fs. Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study): randomised controlled trial. *Lancet* 2005;**366**:1849–1861.

6. Barter PJ, Caulfield M, Eriksson M, Grundy SM, Kastelein JJ, Komajda M, Lopez-Sendon J, Mosca L, Tardif JC, Waters DD, Shear CL, Revkin JH, Buhr KA, Fisher MR, Tall AR, Brewer B, Investigators I. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med* 2007;**357**:2109–2122.

7. Wattanasuwan N, Khan IA, Gowda RM, Vasavada BC, Sacchi TJ. Effect of acute myocardial infarction on cholesterol ratios. *Chest* 2001;**120**:1196–1199.

8. Goldberg AS, Hegele RA. Cholesteryl ester transfer protein inhibitors for dyslipidemia: focus on dalcetrapib. *Drug Des Devel Ther* 2012;**6**:251–259.

9. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008;**27**:1133–1163.

10. Hingorani A, Humphries S. Nature's randomised trials. *Lancet* 2005;**366**:1906–1908.

11. Ference BA, Yoo W, Alesh I, Mahajan N, Mirowska KK, Mewada A, Kahn J, Afonso L, Williams KA Sr, Flack JM. 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.

12. Kathiresan S, Melander O, Anevski D, Guiducci C, Burtt NP, Roos C, Hirschhorn JN, Berglund G, Hedblad B, Groop L, Altshuler DM, Newton-Cheh C, Orho-Melander M. Polymorphisms associated with cholesterol and risk of cardiovascular events. *N Engl J Med* 2008;**358**:1240–1249.

13. Frikke-Schmidt R, Nordestgaard BG, Stene MC, Sethi AA, Remaley AT, Schnohr P, Grande P, Tybjaerg-Hansen A. Association of loss-of-function mutations in the ABCA1 gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. *JAMA* 2008;**299**:2524–2532.

14. Johannsen TH, Kamstrup PR, Andersen RV, Jensen GB, Sillesen H, Tybjaerg-Hansen A, Nordestgaard BG. Hepatic lipase, genetically elevated high-density lipoprotein, and risk of ischemic cardiovascular disease. *J Clin Endocrinol Metab* 2009;**94**:1264–1273.

15. Haase CL, Tybjaerg-Hansen A, Grande P, Frikke-Schmidt R. Genetically elevated apolipoprotein A-I, high-density lipoprotein cholesterol levels, and risk of ischemic heart disease. *J Clin Endocrinol Metab* 2010;**95**:E500–E510.

16. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, Hindy G, Holm H, Ding EL, Johnson T, Schunkert H, Samani NJ, Clarke R, Hopewell JC, Thompson JF, Li M, Thorleifsson G, Newton-Cheh C, Musunuru K, Pirruccello JP, Saleheen D, Chen L, Stewart A, Schillert A, Thorsteinsdottir U, Thorgeirsson G, Anand S, Engert JC, Morgan T, Spertus J, Stoll M, Berger K, Martinelli N, Girelli D, McKeown PP, Patterson CC, Epstein SE, Devaney J, Burnett MS, Mooser V, Ripati S, Surakka I, Nieminen MS, Sinisalo J, Lokki ML, Perola M, Havulinna A, de Faire U, Gigante B, Ingelsson E, Zeller T, Wild P, de Bakker PI, Klungel OH, Maitland-van der Zee AH, Peters BJ, de Boer A, Grobbee DE, Kamphuisen PW, Deneer VH, Elbers CC, Onland-Moret NC, Hofker MH, Wijmenga C, Verschuren WM, Boer JM, van der Schouw YT, Rasheed A, Frossard P, Demissie S, Willer C, Do R, Ordovas JM, Abecasis GR, Boehnke M, Mohlke KL, Daly MJ, Guiducci C, Burtt NP, Surti A, Gonzalez E, Purcell S, Gabriel S, Marrugat J, Peden J, Erdmann J, Diemert P, Willenborg C, Konig IR, Fischer M, Hengstenberg C, Ziegler A, Buysschaert I, Lambrechts D, Van de Werf F, Fox KA, El Mokhtari NE, Rubin D, Schrezenmeir J, Schreiber S, Schafer A, Danesh J, Blankenberg S, Roberts R, McPherson R, Watkins H, Hall AS, Overvad K, Rimm E, Boerwinkle E, Tybjaerg-Hansen A, Cupples LA, Reilly MP, Melander O, Mannucci PM, Ardissino D, Siscovick D, Elosua R, Stefansson K, O'Donnell CJ, Salomaa V, Rader DJ, Peltonen L, Schwartz SM, Altshuler D, Kathiresan S. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* 2012;**380**:572–580.

17. Triglyceride Coronary Disease Genetics C, Emerging Risk Factors C, Sarwar N, Sandhu MS, Ricketts SL, Butterworth AS, Di Angelantonio E, Boekholdt SM, Ouwehand W, Watkins H, Samani NJ, Saleheen D, Lawlor D, Reilly MP, Hingorani AD, Talmud PJ, Danesh J. Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies. *Lancet* 2010;**375**:1634–1639.

18. Waterworth DM, Ricketts SL, Song K, Chen L, Zhao JH, Ripatti S, Aulchenko YS, Zhang W, Yuan X, Lim N, Luan J, Ashford S, Wheeler E, Young EH, Hadley D, Thompson JR, Braund PS, Johnson T, Struchalin M, Surakka I, Luben R, Khaw KT, Rodwell SA, Loos RJ, Boekholdt SM, Inouye M, Deloukas P, Elliott P, Schlessinger D, Sanna S, Scuteri A, Jackson A, Mohlke KL, Tuomilehto J, Roberts R, Stewart A, Kesaniemi YA, Mahley RW, Grundy SM, McArdle W, Cardon L, Waeber G, Vollenweider P, Chambers JC, Boehnke M, Abecasis GR, Salomaa V, Jarvelin MR, Ruokonen A, Barroso I, Epstein SE, Hakonarson HH, Rader DJ, Reilly MP, Witteman JC, Hall AS, Samani NJ, Strachan DP, Barter P, van Duijn CM, Kooner JS, Peltonen L, Wareham NJ, McPherson R, Mooser V, Sandhu MS. Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler Thromb Vasc Biol* 2010;**30**:2264–2276.

19. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, Johansen CT, Fouchier SW, Isaacs A, Peloso GM, Barbalic M, Ricketts SL, Bis JC, Aulchenko YS, Thorleifsson G, Feitosa MF, Chambers J, Orho-Melander M, Melander O, Johnson T, Li X, Guo X, Li M, Shin Cho Y, Jin Go M, Jin Kim Y, Lee JY, Park T, Kim K, Sim X, Twee-Hee Ong R, Croteau-Chonka DC, Lange LA, Smith JD, Song K, Hua Zhao J, Yuan X, Luan J, Lamina C, Ziegler A, Zhang W, Zee RY, Wright AF, Witteman JC, Wilson JF, Willemsen G, Wichmann HE, Whitfield JB, Waterworth DM, Wareham NJ, Waeber G, Vollenweider P, Voight BF, Vitart V, Uitterlinden AG, Uda M, Tuomilehto J, Thompson JR, Tanaka T, Surakka I, Stringham HM, Spector TD, Soranzo N, Smit JH, Sinisalo J, Silander K, Sijbrands EJ, Scuteri A, Scott J, Schlessinger D, Sanna S, Salomaa V, Saharinen J, Sabatti C, Ruokonen A, Rudan I, Rose LM, Roberts R, Rieder M, Psaty BM, Pramstaller PP, Pichler I, Perola M, Penninx BW, Pedersen NL, Pattaro C, Parker AN, Pare G, Oostra BA, O'Donnell CJ, Nieminen MS, Nickerson DA, Montgomery GW, Meitinger T, McPherson R, McCarthy MI, McArdle W, Masson D, Martin NG, Marroni F, Mangino M, Magnusson PK, Lucas G, Luben R, Loos RJ, Lokki ML, Lettre G, Langenberg C, Launer LJ, Lakatta EG, Laaksonen R, Kyvik KO, Kronenberg F, Konig IR, Khaw KT, Kaprio J, Kaplan LM, Johansson A, Jarvelin MR, Janssens AC, Ingelsson E, Igl W, Kees Hovingh G, Hottenga JJ, Hofman A, Hicks AA, Hengstenberg C, Heid IM, Hayward C, Havulinna AS, Hastie ND, Harris TB, Haritunians T, Hall AS, Gyllensten U, Guiducci C, Groop LC, Gonzalez E, Gieger C, Freimer NB, Ferrucci L, Erdmann J, Elliott P, Ejebe KG, Doring A, Dominiczak AF, Demissie S, Deloukas P, de Geus EJ, de Faire U, Crawford G, Collins FS, Chen YD, Caulfield MJ, Campbell H, Burtt NP, Bonnycastle LL, Boomsma DI, Boekholdt SM, Bergman RN, Barroso I, Bandinelli S, Ballantyne CM, Assimes TL, Quertermous T, Altshuler D, Seielstad M, Wong TY, Tai ES, Feranil AB, Kuzawa CW, Adair LS, Taylor HA Jr, Borecki IB, Gabriel SB, Wilson JG, Holm H, Thorsteinsdottir U, Gudnason V, Krauss RM, Mohlke KL, Ordovas JM, Munroe PB, Kooner JS, Tall AR, Hegele RA, Kastelein JJ, Schadt EE, Rotter JI, Boerwinkle E, Strachan DP, Mooser V, Stefansson K, Reilly MP, Samani NJ, Schunkert H, Cupples LA, Sandhu MS, Ridker PM, Rader DJ, van Duijn CM, Peltonen L, Abecasis GR, Boehnke M, Kathiresan S. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010;**466**:707–713.

20. Holmes MV, Harrison S, Talmud PJ, Hingorani AD, Humphries SE. Utility of genetic determinants of lipids and cardiovascular events in assessing risk. *Nat Rev Cardiol* 2011;**8**:207–221.

21. Asselbergs FW, Guo Y, van Iperen EP, Sivapalaratnam S, Tragante V, Lanktree MB, Lange LA, Almoguera B, Appelman YE, Barnard J, Baumert J, Beitelshees AL, Bhangale TR, Chen YD, Gaunt TR, Gong Y, Hopewell JC, Johnson T, Kleber ME, Langaee TY, Li M, Li YR, Liu K, McDonough CW, Meijs MF, Middelberg RP, Musunuru K, Nelson CP, O'Connell JR, Padmanabhan S, Pankow JS, Pankratz N, Rafelt S, Rajagopalan R, Romaine SP, Schork NJ, Shaffer J, Shen H, Smith EN, Tischfield SE, van der Most PJ, van Vliet-Ostaptchouk JV, Verweij N, Volcik KA, Zhang L, Bailey KR, Bailey KM, Bauer F, Boer JM, Braund PS, Burt A, Burton PR, Buxbaum SG, Chen W, Cooper-Dehoff RM, Cupples LA, Dejong JS, Delles C, Duggan D, Fornage M, Furlong CE, Glazer N, Gums JG, Hastie C, Holmes MV, Illig T, Kirkland SA, Kivimaki M, Klein R, Klein BE, Kooperberg C, Kottke-Marchant K, Kumari M, Lacroix AZ, Mallela L, Murugesan G, Ordovas J, Ouwehand WH, Post WS, Saxena R, Scharnagl

H, Schreiner PJ, Shah T, Shields DC, Shimbo D, Srinivasan SR, Stolk RP, Swerdlow DI, Taylor HA Jr, Topol EJ, Toskala E, van Pelt JL, van Setten J, Yusuf S, Whittaker JC, Zwinderman AH, Anand SS, Balmforth AJ, Berenson GS, Bezzina CR, Boehm BO, Boerwinkle E, Casas JP, Caulfield MJ, Clarke R, Connell JM, Cruickshanks KJ, Davidson KW, Day IN, de Bakker PI, Doevendans PA, Dominiczak AF, Hall AS, Hartman CA, Hengstenberg C, Hillege HL, Hofker MH, Humphries SE, Jarvik GP, Johnson JA, Kaess BM, Kathiresan S, Koenig W, Lawlor DA, Marz W, Melander O, Mitchell BD, Montgomery GW, Munroe PB, Murray SS, Newhouse SJ, Onland-Moret NC, Poulter N, Psaty B, Redline S, Rich SS, Rotter JI, Schunkert H, Sever P, Shuldiner AR, Silverstein RL, Stanton A, Thorand B, Trip MD, Tsai MY, van der Harst P, van der Schoot E, van der Schouw YT, Verschuren WM, Watkins H, Wilde AA, Wolffenbuttel BH, Whitfield JB, Hovingh GK, Ballantyne CM, Wijmenga C, Reilly MP, Martin NG, Wilson JG, Rader DJ, Samani NJ, Reiner AP, Hegele RA, Kastelein JJ, Hingorani AD, Talmud PJ, Hakonarson H, Elbers CC, Keating BJ, Drenos F. Large-scale genecentric meta-analysis across 32 studies identifies multiple lipid loci. *Am J Hum Genet* 2012;**91**:823–838.

22. Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, Davey Smith G, Sterne JA. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res* 2012;**21**:223–242.

23. Keating BJ, Tischfield S, Murray SS, Bhangale T, Price TS, Glessner JT, Galver L, Barrett JC, Grant SF, Farlow DN, Chandrupatla HR, Hansen M, Ajmal S, Papanicolaou GJ, Guo Y, Li M, Derohannessian S, de Bakker PI, Bailey SD, Montpetit A, Edmondson AC, Taylor K, Gai X, Wang SS, Fornage M, Shaikh T, Groop L, Boehnke M, Hall AS, Hattersley AT, Frackelton E, Patterson N, Chiang CW, Kim CE, Fabsitz RR, Ouwehand W, Price AL, Munroe P, Caulfield M, Drake T, Boerwinkle E, Reich D, Whitehead AS, Cappola TP, Samani NJ, Lusis AJ, Schadt E, Wilson JG, Koenig W, McCarthy MI, Kathiresan S, Gabriel SB, Hakonarson H, Anand SS, Reilly M, Engert JC, Nickerson DA, Rader DJ, Hirschhorn JN, Fitzgerald GA. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS One* 2008;**3**:e3583.

24. Palmer TM, Sterne JA, Harbord RM, Lawlor DA, Sheehan NA, Meng S, Granell R, Smith GD, Didelez V. Instrumental variable estimation of causal risk ratios and causal odds ratios in Mendelian randomization analyses. *Am J Epidemiol* 2011;**173**:1392–1403.

25. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol* 2013;**42**:1134–1144.

26. Voight BF, Kang HM, Ding J, Palmer CD, Sidore C, Chines PS, Burtt NP, Fuchsberger C, Li Y, Erdmann J, Frayling TM, Heid IM, Jackson AU, Johnson T, Kilpelainen TO, Lindgren CM, Morris AP, Prokopenko I, Randall JC, Saxena R, Soranzo N, Speliotes EK, Teslovich TM, Wheeler E, Maguire J, Parkin M, Potter S, Rayner NW, Robertson N, Stirrups K, Winckler W, Sanna S, Mulas A, Nagaraja R, Cucca F, Barroso I, Deloukas P, Loos RJ, Kathiresan S, Munroe PB, Newton-Cheh C, Pfeufer A, Samani NJ, Schunkert H, Hirschhorn JN, Altshuler D, McCarthy MI, Abecasis GR, Boehnke M. The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet* 2012;**8**:e1002793.

27. Kaiser J. DNA sequencing. A plan to capture human diversity in 1000 genomes. *Science* 2008;**319**:395.

28. Bots ML, Visseren FL, Evans GW, Riley WA, Revkin JH, Tegeler CH, Shear CL, Duggan WT, Vicari RM, Grobbee DE, Kastelein JJ, Investigators R. Torcetrapib and carotid intima-media thickness in mixed dyslipidaemia (RADIANCE 2 study): a randomised, double-blind trial. *Lancet* 2007;**370**:153–160.

29. Angrist JD, Krueger AB. Split-sample instrumental variables estimates of the return to schooling. *J Bus Econ Stat* 1995;**13**:225–235.

30. Thomas DC, Lawlor DA, Thompson JR. Re: Estimation of bias in nongenetic observational studies using 'Mendelian triangulation' by Bautista LE, Smeeth L, Hingorani AD, Casas JP. *Ann Epidemiol* 2007;**17**:511–513.

31. Shah S, Casas JP, Drenos F, Whittaker J, Deanfield J, Swerdlow DI, Holmes MV, Kivimaki M, Langenberg C, Wareham N, Gertow K, Sennblad B, Strawbridge RJ, Baldassarre D, Veglia F, Tremol E, Gigante B, De Faire U, Kumari M, Talmud PJ, Hamsten A, Humphries SE, Hingorani AD. Causal relevance of blood lipid fractions in the development of carotid atherosclerosis: Mendelian randomization analysis. *Circ Cardiovasc Gen* 2013;**6**:63–72.

32. Wurtz P, Kangas AJ, Soininen P, Lehtimaki T, Kahonen M, Viikari JS, Raitakari OT, Jarvelin MR, Davey Smith G, Ala-Korpela M. 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.

33. Varbo A, Benn M, Tybjaerg-Hansen A, Jorgensen AB, Frikke-Schmidt R, Nordestgaard BG. Remnant cholesterol as a causal risk factor for ischemic heart disease. *J Am Coll Cardiol* 2013;**61**:427–436.

34. Wurtz P, Raiko JR, Magnussen CG, Soininen P, Kangas AJ, Tynkkynen T, Thomson R, Laatikainen R, Savolainen MJ, Laurikka J, Kuukasjarvi P, Tarkka M, Karhunen PJ, Jula A, Viikari JS, Kahonen M, Lehtimaki T, Juonala M, Ala-Korpela M, Raitakari OT. High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. *Eur Heart J* 2012;**33**:2307–2316.