**PCSK9 genetic variants and risk of type 2 diabetes: a Mendelian randomisation study**

Amand F Schmidt\*$, Daniel I Swerdlow$, Michael V Holmes$, Riyaz S. Patel, Zammy Fairhurst-Hunter, Donald Lyall, Fernando Pires Hartwig, Bernardo Lessa Horta, Elina Hyppönen, Christine Power, Max Moldovan, Erik van Iperen, Kees Hovingh, Ilja Demuth, Kristina Norman, Elisabeth Steinhagen-Thiessen, Juri Demuth, Lars Bertram, Tian Liu, Stefan Coassin, Johann Willeit, Stefan Kiechl, Karin Willeit, Dan Mason, John Wright, Richard Morris, Goya Wanamethee, Peter Whincup, Yoav Ben-Shlomo, Stela McLachlan, Jackie F. Price, Mika Kivimaki, Catherine Welch, Adelaida Sanchez-Galvez, Pedro Marques-Vidal, Andrew Nicolaides, Andrie G. Panayiotou, N. Charlotte Onland-Moret, Yvonne T. van der Schouw, Giuseppe Matullo, Giovanni Fiorito, Simonetta Guarrera, Carlotta Sacerdote, Nicholas J Wareham, Claudia Langenberg, Robert Scott, Jian'an Luan, Martin Bobak, Sofia Malyutina, Andrzej Pająk, Ruzena Kubinova, Abdonas Tamosiunas, Hynek Pikhart, Lise Lotte Nystrup Husemoen, Niels Grarup, Oluf Pedersen, Torben Hansen, Allan Linneberg, Kenneth Starup Simonsen, Jackie Cooper, Steve E Humphries, Murray Brilliant, Terrie Kitchner, Hakon Hakonarson, David S. Carrell, Catherine A. McCarty, Kirchner, H Lester, Eric B. Larson, David R. Crosslin, Mariza de Andrade, Dan M Roden, Joshua C Denny, Cara Carty, Stephen Hancock, John Attia, Elizabeth Holliday, Martin O'Donnell, Salim Yusuf, Michael Chong, Guillaume Pare, Pim van der Harst, Abdullah M. Said, Ruben N. Eppinga, Niek Verweij, Harold Snieder for the LifeLines Cohort study see online appendix, Tim Christen, Dennis O Mook-Kanamori, Stefan Gustafsson, Lars Lind, Erik Ingelsson, Raha Pazoki, Oscar Franco, Albert Hofman, Andre Uitterlinden, Abbas Dehghan, Alexander Teumer, Sebastian Baumeister, Marcus Dörr, Markus M. Lerch, Uwe Völker, Henry Völzke, Joey Ward, Jill P Pell, Daniel J Smith, Tom Meade, Anke H. Maitland-van der Zee, Ekaterina V. Baranova, Robin Young, Ian Ford, Archie Campbell, Sandosh Padmanabhan, Michiel L Bots, Diederick E. Grobbee, Philippe Froguel, Dorothée Thuillier, Beverley Balkau, Amélie Bonnefond, Bertrand Cariou, Melissa Smart, Yanchun Bao, Meena Kumari, Anubha Mahajan, Paul M Ridker, Daniel I. Chasman, Alex P. Reiner, Leslie A Lange, Maryllyn D Ritchie, Folkert W Asselbergs, Juan-Pablo Casas, Brendan J Keating#, David Preiss#, Aroon D Hingorani#, for the UCLEB consortium see online appendix, Naveed Sattar\*#.

\*Corresponding authors: [amand.schmidt@ucl.ac.uk](mailto:amand.schmidt@ucl.ac.uk) and naveed·sattar@glasgow.ac.uk

$Joint First authors.

#Joint Senior authors.

**Author affiliations**

**Institute of Cardiovascular Science, University College London, UK** (AF Schmidt PhD, DI Swerdlow PhD, R Patel MD, Professor AD Hingorani MD, Prof FW Asselbergs MD), **Department of Medicine, Imperial College London, London, UK** (DI Swerdlow), **Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, UK** (MV Holmes PhD, D PreissMD), **Medical Research Council Population Health Research Unit at the University of Oxford, UK** (MV Holmes, D Preiss), **The Barts Heart Centre, St Bartholomew’s Hospital** (R Patel), **Wellcome Trust Centre for Human Genetics, University of Oxford** (Z Fairhurst-Hunter MSc, A Mahajan PhD), **Institute of Health & Wellbeing, University of Glasgow, Scotland, UK** (D Lyall, PhD, J Ward PhD, JP Pell PhD, DJ Smith PhD) **Postgraduate Program in Epidemiology, Federal University of Pelotas, Pelotas, Brazil** (FP Hartwig MSc, Prof BL Horta PhD), **Centre for Population Health Research, Sansom Institute for Health Research, University of South Australia, Australia** (Professor E Hyppönen PhD), **Population, Policy and Practice, UCL GOS Institute of Child Health, London, UK** (E Hyppönen, Prof C Power PhD), **South Australian Health and Medical Research Institute, Adelaide, Australia** (E Hyppönen), **School of Health Sciences, University of South Australia, Australia** (M Moldovan PhD), **South Australian Health and Medical Research Institute - EMBL Australia, Adelaide, Australia**  (M Moldovan), **Durrer Center for Cardiovascular Research, Netherlands Heart Institute, Utrecht, The Netherlands** (E van Iperen MSc), **Department of clinical epidemiology, biostatistics and bioinformatics, Academic medical center Amsterdam, the Netherlands** (E van Iperen), **Department of vascular medicine, Academic medical center Amsterdam, the Netherlands** (Professor K Hovingh PhD), **Charité Research Group on Geriatrics** (I Demuth PhD, K Norman PhD, Prof E Steinhagen-Thiessen MD) **and Institute of Medical and Human Genetics** (I Demuth), **Charité – Universitätsmedizin Berlin, Germany**; **E.CA Economics GmbH, Berlin, Germany** (J Demuth PhD); **Lübeck Interdisciplinary Platform for Genome Analytics (LIGA), Institutes of Neurogenetics and Integrative and Experimental Genomics, University of Lübeck, Lübeck, Germany** (Prof L Bertram MD) **and Neuroepidemiology and Ageing Research Unit, School of Public Health, Faculty of Medicine, The Imperial College of Science, Technology, and Medicine, London , UK** (L Bertram); **Max Planck Institute for Human Development, Berlin, Germany; Max Planck Institute for Molecular Genetics, Berlin, Germany** (T Liu PhD), **Division of Genetic Epidemiology Innsbruck, Department of Medical Genetics, Molecular and Clinical Pharmacology, Medical University of Innsbruck, Innsbruck, Austria** (S Coassin PhD), **Department of Neurology, Medical University Innsbruck, Innsbruck, Austria** (Prof J Willeit PhD, Prof S Kiechl MD, K Willeit MD), **Bradford Institute for Health Research, Bradford Royal Infirmary, Bradford, UK** (D Mason PhD, Prof J Wright FRCP), **School of Social and Community Medicine, University of Bristol, Bristol, UK** (Prof R Morris PhD, Prof Y Ben-Shlomo PhD), **Dept Primary Care & Population Health, University College London, UK** (Prof G Wannamethee PhD), **Population Health Research Institute, St George’s, University of London, UK** (Prof P Whincup FRCP), **Centre for Population Health Sciences, The Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, UK** (S McLachlan PhD, Prof JF Price MD), **Department of Epidemiology and Public Health, UCL Institute of Epidemiology and Health Care, University College London, UK** (Prof M Kivimaki PhD, C Welch PhD, A Sanchez-Galvez PhD, Prof M Bobak PhD, H Pikhart PhD), **Department of Medicine, Internal Medicine, Lausanne university hospital** (P Marques-Vidal PhD), **Department of Vascular Surgery, Imperial College, London, United Kingdom** (A Nicolaides PhD), **Department of Surgery, Nicosia Medical School, University of Nicosia, Nicosia, Cyprus** (A Nicolaides), **Cyprus International Institute for Environmental and Public Health, Cyprus University of Technology, Limassol, Cyprus** (AG Panayiotou PhD), **Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands** (NC Onland-Moret PhD, Prof YT van der Schouw PhD, FW Asselbergs, Prof ML Bots MD, Prof DE Grobbee PhD), **Human Genetics Foundation, HuGeF, Turin, Italy** (G Matullo PhD, G Fiorito PhD, S Guarrera PhD), **Department of Medical Sciences, University of Turin, Turin, Italy** (G Matullo PhD, G Fiorito PhD, S Guarrera PhD) **Cancer Epidemiology Unit, San Giovanni Battista Hospital, Turin, Italy; Centre for Oncology Prevention, CPO Piemonte, Turin, Italy** (C Sacerdote PhD), **MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Addenbrooke’s Hospital, Cambridge, UK** (NJ Wareham PhD, C Langenberg PhD, Prof R Scott PhD, J Luan PhD), **Novosibirsk State Medical University, Novosibirsk, Russian Federation; and Institute of Internal and Preventive Medicine, Siberian Branch of the Russian Academy of Medical Sciences, Novosibirsk, Russian Federation** (Prof S Malyutina PhD), **Collegium Medicum, Jagiellonian University, Krakow, Poland** (A Pająk PhD), **National Institute of Public Health, Prague, Czech Republic** (R Kubinova PhD), **Lithuanian University of Health Sciences, Kaunas, Lithuania** (Prof A Tamosiunas PhD), **Research Centre for Prevention and Health, the Capital Region of Denmark, Denmark** (LLN Husemoen PhD, KS Simonsen PhD), **The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark** (N Grarup PhD, O Pedersen PhD, T Hansen PhD) **Research Centre for Prevention and Health, the Capital Region of Denmark, Denmark; Department of Clinical Experimental Research, Rigshospitalet, Denmark; and Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark** (Prof A Linneberg PhD), **Centre for Cardiovascular Genetics, Institute Cardiovascular Science, University College London, UK** (J Cooper MSc, Prof SE Humphries PhD), **Center for Human Genetics, Marshfield Clinic Research Foundation** (M Brilliant PhD, T Kitchner PhD), **Children's Hospital of Philadelphia** (H Hakonarson PhD), **Essentia Institute of Rural Health** (DS Carrell PhD, CA McCarty PhD), **Geisinger** (KH Lester PhD, MD Ritchie PhD), **Group Health** (EB Larson MD, DR Crosslin PhD, **Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA** (Prof M de Andrade PhD), **Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA Department of Pharmacology, Vanderbilt University School of Medicine, Nashville, TN, USA** (DM Roden), **Vanderbilt University** (JC Denny PhD), **WHI** (C Carty PhD), **University of Newcastle, Newcastle NSW Australia** (S Hancock PhD, J Attia PhD, E Holliday PhD), **Population Health Research Institute, Hamilton, Ontario, Canada** (M O'Donnell PhD, S Prof Yusuf D Phil, M Chong MSc, Prof G Pare MD), **University of Groningen, University Medical Center Groningen, Department of Cardiology, Groningen, The Netherlands** (Prof P van der Harst PhD, AM Said BSc, RN. Eppinga PhD, N Verweij PhD), **University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands** (P van der Harst), **Durrer Center for Cardiovascular Research, Netherlands Heart Institute, Utrecht, the Netherlands** (P van der Harst), **Department of Epidemiology, University of Groningen, University Medical Center Groningen, Netherlands** (Prof H Snieder PhD), **Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands** (T Christen MSc, DO Mook-Kanamori PhD), **Department of Medical Sciences, Molecular Epidemiology, Uppsala University, Uppsala, Sweden** (S Gustafsson PhD, Prof Lars Lind PhD), **Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA; and Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden** (Prof E Ingelsson PhD), **Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands** (R Pazoki PhD, O Franco PhD, Prof A Hofman PhD, A Dehghan PhD), **Department of Biostatistics and Epidemiology, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London** (A Dehghan), **Department of Internal Medicine, Erasmus University Medical center, Rotterdam, The Netherlands** (A Uitterlinden PhD), **Institute for Community Medicine, University Medicine Greifswald** (A Teumer PhD, Prof H Völzke PhD, S Baumeister PhD), **DZHK (German Centre for Cardiovascular Research), partner site Greifswald** (A Teumer, H Völzke, Prof M Dörr MD, Prof U Völker PhD), **Department of Epidemiology and Preventive Medicine University of Regensburg** (S Baumeister), **Department of Internal Medicine B, University Medicine Greifswald** (M Dörr), **Department of Medicine A, University Medicine Greifswald** (Prof MM Lerch PhD), **Interfaculty Institute of Genetics and Functional Genomics, University Medicine Greifswald** (U Völker), **Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK** (Prof T Meade FRS), **Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute of Pharmaceutical Sciences, Faculty of Science, Utrecht University, Utrecht, The Netherlands** (Prof AH Maitland-van der Zee PhD, EV Baranova MSc), **Respiratory Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands** (AH Maitland-van der Zee), **Robertson Centre for Biostatistics,University of Glasgow, Glasgow,UK** (R Young PhD, I Ford PhD), **Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK** (A Campbell MA), **Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK** (Prof S Padmanabhan PhD, Prof N Sattar PhD), **CNRS UMR 8199, European Genomic Institute for Diabetes (EGID), Institut Pasteur de Lille, University of Lille, 59000 Lille, France** (Prof P Froguel PhD, D Thuillier PhD, A Bonnefond PhD), **Department of Genomics of Common Disease, Imperial College London, W12 0NN London, United Kingdom** (P Froguel, A Bonnefond), **Centre de Recherche en Epidémiologie et Santé des Populations, CESP, INSERM U1018, Renal and cardiovascular epidemiology, Villejuif, France** (B Balkau PhD), **l'institut du Thorax, INSERM, CNRS, UNIV Nantes, CHU Nantes, Nantes, France** (Prof B Cariou MD), **Institute for Social and Economic Research, University of Essex, Colchester, Essex, CO4 3SQ, UK** (M Smart PhD, Y Bao PhD, Prof M Kumari PhD), **Harvard Medical School Center for Cardiovascular Disease Prevention Brigham and Women's Hospital** (Prof PM Ridker MD, DI Chasman PhD), **UWash** (AP Reiner MD),CARe/WHI (LA Lange PhD), **Farr Institute of Health Informatics Research, UCL Institute of Health Informatics, University College London, London, UK** (FW Asselbergs, Prof JP Casas PhD), **Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, the Netherlands** (FW Asselbergs), **Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, the Netherlands** (FW Asselbergs), **Department of Surgery, University of Pennsylvania, Philadelphia, PA 19104, USA** (BJ Keating PhD).

**Running title:** *PCSK9* and type 2 diabetes.

**Word count text:** 2,988

**Word count abstract:** 237

**Number of references:** 44

**Number of tables and figures:** 4

**(Web) appendix:** 2

**Abstract**

**Background** Statin treatment and variants in the gene encoding HMG-CoA reductase are associated with reductions in both the concentration of low-density lipoprotein cholesterol (LDL-C) and the risk of coronary heart disease, but also with modest hyperglycemia, higher weight and risk of type 2 diabetes mellitus (T2DM). We sought to clarify the associations of LDL-C-lowering *PCSK9* variants with T2DM and related biomarkers to gauge likely effects of PCSK9 inhibitors.

**Methods** Associations of *PCSK9* variants with LDL-C, fasting blood glucose, HbA1c, fasting insulin, weight and T2DM risk, available for over 550,000 individuals and 51,623 T2DM cases, were estimated using a standardised analysis plan, meta-analyses and weighted gene-centric scores (GS).

**Findings** Combined analysis of four independent *PCSK9* variants scaled to 1 mmol/L lower LDL-C resulted in the following associations: fasting glucose (0·09 mmol/L; 95%CI 0·02; 0·15), HbA1c (0·03%; 95%CI -0·01; 0·08), fasting insulin (0·00% 95%CI -0·06; 0·07), body weight (1·03 kg; 95%CI 0·24; 1·82), waist-to-hip ratio (0·006; 95%CI 0·003; 0·010), body mass index (0·11 kg/m2; 95%CI -0·09; 0·30), and an odds ratio of 1·29 for T2DM (95%CI 1·11; 1·50).

**Interpretation** *PCSK9* variants associated with lower LDL-C were also associated with higher levels of glucose, weight, waist-to-hip ratio and increased T2DM risk. Trials of PCSK9 inhibitor drugs should carefully evaluate these safety outcomes and quantify the risks and benefits of PCSK9 inhibitor treatment as previously done for statins.

**Funding** the funding sources are cited at the end of the paper.

**Keywords** Diabetes Mellitus; Genetic Association Studies; Mendelian randomisation; LDL-cholesterol; PCSK9 inhibition.

**Research in context**

*Evidence before this study*

We searched PubMed for “pcsk9[All Fields] AND ("antagonists and inhibitors"[Subheading] OR ("antagonists"[All Fields] AND "inhibitors"[All Fields]) OR "antagonists and inhibitors"[All Fields] OR "inhibitors"[All Fields]) AND ("diabetes mellitus"[MeSH Terms] OR ("diabetes"[All Fields] AND "mellitus"[All Fields]) OR "diabetes mellitus"[All Fields])” up to 8th October 2016 to find studies reporting treatment with PCSK9 inhibitors and/or carriage or genetic variants in *PCSK9* in relation to diabetes. This identified 17 studies, two of which presented novel, yet contrasting findings in relation to genetic variants in *PCSK9* and glycaemic status.

Randomized trials of treatment with statins and carriage of corresponding genetic variants in *HMGCR* that lower low density lipoprotein cholesterol (LDL-C) both increase the risk of diabetes. More recently, genetic predisposition to lower LDL-C concentrations has been linked to higher risk of diabetes, suggesting that dysglyaemia may be a consequence of lowering LDL-C in general. Whether lowering of LDL-C by PCSK9 inhibitors results in increased risk of diabetes is currently unknown. Clinical trials of PCSK9 inhibitors and cardiovascular outcomes are on-going, but reliable evidence is on the association of PCSK9 inhibition and risk of diabetes may take longer to accrue. Mendelian randomization is an established approach that uses randomly-allocated variants in the encoding gene to infer mechanism-based efficacy and safety outcomes from pharmacological perturbation of a drug target.

*Added value of this study*

We used four genetic variants in *PCSK9* in over 550,000 individuals with ~50,000 diabetes cases and found that *PCSK9* genetic variants that associated with lower LDL-C concentrations also showed significant associations with increased concentrations of fasting glucose, body weight and risk of diabetes. This adds robust new evidence to previous studies that identified weak associations of *PCSK9* with risk of diabetes.

*Implications of all the available evidence*

Similar to statin therapy, treatment with PCSK9 inhibitors is likely to increase the risk of diabetes. Current treatment trials and patients treated with PCSK9 inhibitors should be carefully monitored for dysglycaemia. Drugs that lower LDL-C may increase risk of diabetes.

**Introduction**

The benefit of statins in reducing LDL-C and coronary heart disease (CHD) risk is well established. More recently, only after completion of numerous randomised controlled trials (RCTs), was it discovered that statins increase risk of type 2 diabetes mellitus (T2DM)1,2, though this effect is modest and greatly outweighed by the benefits of this drug class. Genetic studies based on common variants in the gene encoding the target of statins, HMG-coA reductase (HMGCR) suggest the effect is mechanism-based (i.e. on-target)3. Genetic studies utilising variants in a broader range of genes indicate a more general link between lower LDL-C and higher T2DM risk4,5 . Consistent with this, patients with autosomal dominant familial hypercholesterolemia (FH) caused by mutations in the LDL-receptor and apolipoprotein B genes are 50% less likely to be diagnosed with T2DM compared to their unaffected relatives6.

Gain-of-function mutations in *PCSK9*, the gene encoding proprotein convertase subtilisin/kexin type 9 also cause FH7, while loss-of-function mutations in the same gene lower LDL-C and protect against CHD8. Consequently, monoclonal antibodies (mAbs) inhibiting PCSK9 have been developed9 and are effective in lowering LDL-C-lowering by 50-70%10, with preliminary evidence suggesting that this may be associated with reduced myocardial infarction risk and all-cause mortality9. While large phase 3 RCTs to evaluate the effects of PCSK9 mAbs on cardiovascular events are underway, conclusive evidence on the specific effect of PCSK9 inhibition on T2DM risk from individual or meta-analysis of RCTs may not emerge for some time.

We used the principle of Mendelian randomisation (MR) as a tool for drug target validation, whereby common variants in a gene that encodes a drug target, through effects on expression or activity, are used to predict the on-target effect of pharmacological modification of the same target3,11,12. We evaluated associations of common genetic variants in *PCSK9* with markers of glycaemia, body weight, and with T2DM risk to assess the potential on-target effects of PCSK9 inhibition on these traits. While a recent study provided weak evidence of an association of a single nucleotide polymorphism (SNP) in *PCSK9* and T2DM13, our aim was to provide definitive evidence on the relationship of *PCSK9* genetic variants and risk of T2DM by using multiple SNPs in the *PCSK9* locus in 50 studies supplemented by large genetic consortia.

**Methods**

*Genetic variant selection*

Four SNPs in or near *PCSK9* were selected based on a strong association with LDL-C as reported by the Global Lipids Genetics Consortium (GLGC)14; low pairwise linkage disequilibrium (LD) (r2 ≤0·30) with SNPs within the same and adjacent genes (based on 1,000 Genomes CEU data); high prior probability of being a functional variant based on the combined annotation dependent depletion (CADD) score, and/or the SNP being non-synonymous15; and/or prior reported associations with CHD16. Based on these criteria, SNPs rs11583680 (minor allele frequency [MAF] = 0.14), rs11591147 (MAF = 0.01), rs2479409 (MAF = 0.36) and rs11206510 (MAF = 0.17) were selected (see Appendix Table 1).

*Individual participant-level and summary-level data*

We analysed data from two sources. Participating studies executed a common analysis script (available at http://www.ucl.ac.uk/genetic-epidemiology/TBA) on their own data, submitting summary estimates to the UCL analysis centre. Main effect estimates from the participating studies were then meta-analysed with pooled summary estimates from the public domain data repositories of relevant genetic (GWAS) consortia, but only if the study-level estimates had not previously contributed to consortia results, so as to prevent double counting.

Data were collected on LDL-C, insulin (fasting and non-fasting), glucose (fasting and non-fasting), haemoglobin A1c (HbA1c), insulin resistance and secretion via basal homeostatic model assessments (HOMA-IR and HOMA-B), body weight, body mass index (BMI), waist-to-hip ratio (WHR), and history or incidence of T2DM.

Publicly available summary-level data were available on blood lipids from GLGC14, T2DM-related biomarkers (plasma insulin, glucose, HbA1c, HOMA-IR and HOMA-B) from Meta-analyses of Glucose and Insulin-related traits Consortium (MAGIC)17-19, body weight, BMI and WHR from the Genetic Investigation of Anthropometric Traits consortium (GIANT)20,21, and T2DM from the Diabetes Genetics Replication and Meta-analysis consortium (DIAGRAM22) and Exome chip 80K23. In addition, cross-sectional data were obtained for adiposity traits and the prevalence of T2DM from UK Biobank24.

*Statistical analyses*

In all analyses we assumed an additive allele effect with genotypes coded as 0, 1 and 2, representing the number of minor alleles. Continuous biomarkers were analysed using linear regression models, and the composite endpoint of prevalent or incident T2DM was analysed using logistic regression. Study-specific associations were pooled for each SNP using the inverse variance weighted method for fixed and random effects meta-analysis. Between-study heterogeneity was assessed using the Q-test, and the *I*2 statistic25 with a one-sided upper 97·5% confidence interval. Study-specific associations were excluded if the SNP was not in Hardy-Weinberg equilibrium (see Appendix Table 2).

Our approach to SNP selection was designed to prune the number of SNPs at *PCSK9* used in the analysis, without loss of information. We decided a priori to combine the four, approximately independent SNPs in a weighted gene-centric score (GS) using the inverse variance weighted method for fixed and random effects26. The GS provides a more precise estimate of the downstream consequences of variation at *PCSK9* by incorporating maximal biological variation, and the corresponding effect of pharmacological treatment. Furthermore, when the four SNP effects are homogeneous (assessed using the heterogeneity measures Q-test and *I*2) the GS estimates will be more powerful and precise compared to individual SNPs in isolation. If, however, the SNP effects are heterogeneous (meaning that the *PCSK9* effects are different according to which part of the gene is evaluated) the GS method will be less powerful than the individual SNP tests (depending on the degree of heterogeneity). While our aim is to estimate the effect of the PCSK9 locus as a whole, SNP specific estimates are provided both in the main figures as well as the appendix figures. Other important assumptions of the GS approach are (approximate) independence of the included SNPs (assessed by pairwise linkage disequilibrium (r2) and using multivariable regression models, see further down), and the additivity of allele effects (assessed as described below). We also evaluated whether the association of individuals SNPs with diabetes risk was in proportion to the association with LDL-C.

Estimates are presented as mean differences or odds ratios (OR) with 95% confidence intervals (CI), presented either per LDL-C decreasing allele or, in the case of GS to mimic the effect of PCSK9 mAbs, per 1 mmol/L lower in LDL-C (1 mmol/L LDL-C equals 38·67 mg/dL to 2 decimal points). The per 1 mmol/L GS effect estimates were derived by multiplying point estimates and their variances by the multiplicative inverse of the estimated SNP-LDL-C effects. Similar to most genetic studies, missing data were excluded in an available case manner; assuming a missing completely at random mechanism 27,28. Analyses were conducted using the statistical programme R29. To avoid potential bias due to population stratification analyses excluded individuals of non-European ancestry. Ancestry can be a potential source of confounding bias (i.e., population stratification bias) when environment is related to both the genes and the outcome of interest. Given the different genetic constructs between ethnicities the former is known to be true30. Chen et al31 reports different trends in T2DM between ethnic groups, showing that the latter may also be true.

Variants influencing circulating LDL-C have been reported previously to influence the probability of being prescribed a lipid lowering drug32, in the current analysis we did not account for this because: 1) prescription data on these treatments were often not available, and 2) if they were recorded, they were only available for single follow-up point. For lipid lowering treatments a single record of treatment does not properly reflect the time-varying therapy received, and adjusting for only a single record when in fact treatment varies over follow-up may increase bias33. Typically, diabetes therapies are much less-variable over time and correction for this may seem advisable however, due to the strong correlation between history of T2DM and T2DM related drugs, any correction for the latter would essentially correct for prevalent T2DM as well. Finally, any influence of lipid lowering drug therapy would attenuate rather than inflate any associations.

*Sensitivity analyses*

We assumed that the allele effects were additive, which we assessed in available individual participant data by comparing an additive model to a non-additive model (allowing for dominance or recessiveness) using a likelihood ratio test (meta-analysed using Fisher’s method34). Because measurement error may be larger in prevalent cases (ascertained e.g. from hospital records) we conducted a further sensitivity analysis separately analysing incident and prevalent T2DM. This sensitivity analysis was conducted not because we expect the *true* associations of *PCSK9* to be different with regard to prevalent and incident case status but merely reflected a quality control check. While SNPs were selected to be independent there was some degree of residual dependency (Appendix Figure 1, maximum r-squared of 0.26). To explore the impact of this residual correlation between the four study SNPs (Appendix Figure 1) we compared results from a multivariable analysis (including the 4 SNPs in the same model) in studies with individual participant data (correcting for this correlation) to pairwise results (ignoring any between-SNP correlation) based on the same data.

*Role of the funding source*

The funder(s) of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author AFS had full access to all the data in the study and shared final responsibility for the decision to submit for publication with all authors.

All studies contributing data to these analyses were approved by their local ethics committees.

**Results**

50 studies shared participant level data from up to 245,942 individuals, which was supplemented by summary effect estimates from data repositories, resulting in a maximum available sample size of 568,448 subjects, including 51,623 cases of incident or prevalent T2DM. Individual studies were similar with respect to the distribution of biochemical measures (assessed using the median and inter-quartile range of study specific means): for LDL-C these were 3·41 mmol/L (0·39), fasting glucose 5·38 mmol/L (0·58), and HbA1c  5·50% (0·38); see Appendix Table 3 for more details. Pooled pairwise linkage disequilibrium estimates for the four *PCSK9* SNPs were all r2 < 0·30 (Appendix Figure 1) confirming the selected SNPs were in low correlation in the collected data.

*Biomarker associations*

The four *PCSK9* SNPs were associated with LDL-C reductions ranging from -0·02 mmol/L (95%CI: -0·03; -0·02) for rs11583680 to -0·34 mmol/L (95%CI: -0·36; -0·32) for rs11591147 per LDL-C decreasing allele (Figure 1).

Figure 2 depicts the associations of the four *PCSK9* SNPs after scaling the SNP effect to 1 mmol/L lower LDL-C. Results of the *PCSK9* GS analysis indicated that a 1 mmol/L lower LDL-C was associated with 1·03 kg higher body weight (95%CI 0·24; 1·82), a 0·11 kg/m2 difference in BMI (95%CI -0·09; 0·30), and a difference 0·006 in WHR (95%CI 0·003; 0·010). Associations of the *PCSK9* GS with glycaemia measures were 0·09 mmol/L higher fasting plasma glucose (95%CI 0·02; 0·15), a 0·03% difference in HbA1c (95%CI -0·01; 0·08), and 0·00% for fasting insulin (95%CI -0·06; 0·07). SNP specific forest plots are presented in Appendix Figures 2-89. The estimates were similar when corrected for linkage disequilibrium (Appendix Figure 90), and no systematic deviations from an additive model were observed (Appendix Table 4). Finally, we noted an unanticipated effect on height: mean difference 0·008 meters (95%CI 0·0008; 0·015) see Appendix Figures 91-95.

*Association with T2DM*

Figure 3 shows the associations of individual *PCSK9* variants and the GS with T2DM. Using the *PCSK9* GS, 1 mmol/L lower LDL-C was associated with an increased risk of T2DM; OR 1·29 (95%CI 1·11; 1·50). Exploring the *PCSK9* associations with incident (Appendix Figure 96) or prevalent (Appendix Figure 97) T2DM separately showed directional concordance: incident T2DM (OR 1·15 95%CI 0·76; 1·72) and prevalent T2DM (OR 1·26; 95%CI: 0·88, 1·80). Associations of individual SNPs with LDL-C and T2DM showed a dose-response relationship (Figure 4).

**Discussion**

Genetic variants in *PCSK9,* used as a proxy for pharmacological inhibition of PCSK9, were associated with lower LDL-C concentration and increased risk of T2DM. The same variants were also associated with higher fasting glucose, body weight and waist-to-hip ratio, and with directionally concordant (but non-significant) associations for BMI and HbA1c and a seemingly neutral association for fasting insulin. These results mirror previous findings for variants in the *HMGCR* gene encoding the target of statin drugs, with statins modestly increasing body weight and the risk of T2DM3.

When scaled to a 1mmol/L lower LDL-C, the risk for T2DM based on the previously studied *HMGCR* variant rs129163 was an OR of 2·02 (95%CI 1·42; 2·87), greater than the corresponding scaled estimate for this *PCSK9* GS (OR 1·29, 95%CI 1·11; 1·50), and for an estimate based on SNPs affecting LDL-C selected from throughout the genome (OR 1·27 (95% CI 1·14; 1·41)5. However, effect estimates obtained from MR studies proxy *lifetime* exposure to natural genetic variation, and may therefore not directly translate to the magnitude of effect of any corresponding pharmacological treatment introduced much later in life for a shorter duration of time35. For example, in a meta-analysis of RCTs of statin treatment36, the OR for T2DM was 1·12 (95%CI 1·06; 1·18).

In the case of statins, the treatment benefit in terms of CHD risk reduction greatly outweighs any potential adverse effect on T2DM risk, partly because the magnitude of the risk reduction in CHD is greater than the risk increase in T2DM, and partly because the absolute risk of CHD in primary prevention populations eligible for statin treatment is greater than the absolute risk of T2DM37. A similarly precise risk assessment for PCSK9 inhibitors awaits results from larger and longer term RCTs. A recent meta-analysis38 reported that treatment with alirocumab was associated with an OR for T2D of 0·89 (95%CI 0·62; 1·28) compared to placebo based on 133 T2DM events.

We have previously reported examples of common variants in genes encoding a protein drug target mimicking the on-target effects of pharmacological interventions on biomarkers and disease outcomes in type, direction, and relative magnitude3,39,40. However, such analyses cannot predict off-target effects of therapies. It should be noted that we refer to on-target effects that are due to a drug effect on the intended target (in this case PCSK9) and off-target effects that might occur due to the drug also binding to an unintended target (in this case, any target other than PCSK9). Although mAb therapeutics are often highly-specific, perhaps more so than small molecule therapeutics, they retain the potential for off-target effects. Hence, in the presence of off-target effects, results from ongoing RCTs could differ from the genetic associations reported here.

Our main findings are based on four *PCSK9* SNPs in combination and scaled to 1 mmol/L lower LDL-C. This approach assumes additive effects across the SNPs, an assumption that held well on sensitivity analyses. A potentially unobserved non-additive effect may explain why we observed a genetic association with fasting glucose and a concordant (although non-signification) association with HbA1c, while fasting insulin seemed unaffected. It is also important to note that there is conflicting evidence on a possible role of *PCSK9* and PCSK9mAbs in disruption of the pancreatic islet function41,42. The lack of association with HbA1c may be related to the large amount of heterogeneity between the 4 SNPs (upper bound *I2* 72%). Interestingly, we found that the association of the *PCSK9* GS with BMI was smaller than that with body weight, which might be explained by a slightly greater average height among subjects with *PCSK9* variants associated with lower LDL-C concentrations. A further potential reason for the slight discrepancy between the BMI and body weight associations could be the greater heterogeneity in the associations of *PCSK9* SNPs with BMI than with weight. Of note, the GS effect estimates were often driven by a large effect of SNP rs11591147, as our dose response analysis shows (Figure 4), the larger influence of this SNP appropriately reflects the proportionally larger LDL-C effect of this SNP. Finally, we did not have access to measures of PCSK9 in this analysis but others43 have shown associations between common and rare *PCSK9* alleles (including the same SNPs used here) and circulating PCSK9 concentrations.

Setting aside associations with glycaemia and weight, T2DM risk could also be partially increased because lifelong exposure to genetic variation in *PCSK9* may reduce mortality, making it conceivable that subjects survive longer and, hence have more time to develop T2DM. Whether *PCSK9* genotype reduces mortality has not be conclusively shown8,44 and irrespective of the nature of the *PCSK9* association with T2DM, large RCTs should determine whether this relation also holds for PCSK9 mAbs.

Using a single SNP in *PCSK9*, a recent study13 provided weak evidence of an association between *PCSK9* variants and risk of T2D (OR= 1.19, 95%CI 1.02-1.38; P=0.03, per 1 mmol/l reduction in LDL-C). In the current study, we incorporated data from 4 SNPs in a *PCSK9* gene score with participant data from 50 studies supplemented by large genetic consortia and are able to provide far more robust, precise and definitive estimates (OR 1.29; 95%CI: 1.11, 1.50; P=9x10-4 per 1 mmol/l lower LDL-C). Previous studies on LDL-C lowering *HMGCR*3 and *NPC1L1*13 variants (encoding pharmacological targets of statins and ezetimibe, respectively) and more widely on LDL-C lowering variants from multiple GWAS-associated loci5, as well as analyses of patients with monogenic hypercholesterolaemia6 have provided evidence of a link between LDL-C and T2DM, compatible with the findings from the present study. However, it is far from certain that all LDL-C lowering interventions will increase risk of T2DM, as not all share the same mechanism of action. The major site of both statins and PCSK9 inhibitors is thought to be the liver, through increased cellular membrane expression of the LDL-C receptor. The liver is also the site of action of the developmental apolipoprotein-B antisense oligonucleotide, mipomersen, whereas ezetemibe, the other licensed LDL-C lowering drug, acts in the intestine to limit LDL-C absorption. A potential unifying mechanism might be pancreatic beta-cell LDL-R upregulation, increased lipid accumulation and beta cell dysfunction6, but this will need to be tested experimentally.

In conclusion, genetic variants in *PCSK9* that associate with lower concentrations of LDL-C are also associated with a modestly higher risk of T2DM and with associated differences in measures of glycaemia and body weight. Ongoing RCTs of PCSK9 inhibitor agents should carefully monitor changes in metabolic markers, including weight and glycaemia, and the incidence of T2DM. Genetic studies of the type used here could be more widely used to interrogate the safety and efficacy of novel drug targets.

**Author contributions**

Amand F Schmidt, Daniel I Swerdlow, Michael V Holmes, Riyaz S. Patel, Folkert W Asselbergs, Juan-Pablo Casas, Brendan J Keating, Aroon D Hingorani, David Preiss, Naveed Sattar contributed to the idea and design of the study. Amand F Schmidt, Daniel I Swerdlow, Michael V Holmes, designed the analysis scripts shared with individual centres. Amand F Schmidt performed the meta-analysis and had access to all the data. Amand F Schmidt, Daniel I Swerdlow, Michael V Holmes drafted the initial manuscript. Riyaz S. Patel, Zammy Fairhurst-Hunter, Donald Lyall, Fernando Pires Hartwig, Bernardo Lessa Horta ,Elina Hyppönen, Christine Power, Max Moldovan , Erik van Iperen, Kees Hovingh, Ilja Demuth, Kristina Norman, Elisabeth Steinhagen-Thiessen, Juri Demuth, Lars Bertram, Tian Liu, Stefan Coassin, Johann Willeit, Stefan Kiechl, Karin Willeit, Dan Mason, John Wright, Richard Morris, Goya Wanamethee, Peter Whincup, Yoav Ben-Shlomo, Stela McLachlan, Jackie F. Price, Mika Kivimaki, Catherine Welch, Adelaida Sanchez-Galvez, Pedro Marques-Vidal, Andrew Nicolaides, Andrie G. Panayiotou, N. Charlotte Onland-Moret, Yvonne T. van der Schouw, Giuseppe Matullo, Giovanni Fiorito, Simonetta Guarrera, Carlotta Sacerdote, Nicholas J Wareham, Claudia Langenberg, Robert Scott, Jian'an Luan, Martin Bobak, Soa Malyutina, Andrzej Paj¡k , Ruzena Kubinova , Abdonas Tamosiunas , Hynek Pikhart, Lise Lotte Nystrup Husemoen, Niels Grarup, Oluf Pedersen, Torben Hansen, Allan Linneberg, Kenneth Starup Simonsen, Jackie Cooper, Steve E Humphries, Murray Brilliant, Terrie Kitchner, Hakon Hakonarson, David S. Carrell, Catherine A. McCarty, Kirchner, H Lester, Eric B. Larson, David R. Crosslin, Mariza de Andrade, Dan M Roden, Joshua C Denny, Cara Carty, Stephen Hancock, John Attia, Elizabeth Holliday, Martin O'Donnell, Salim Yusuf, Michael Chong, Guillaume Pare, Pim van der Harst, Abdullah M. Said, Ruben N. Eppinga, Niek Verweij, Harold Snieder for the LifeLines Cohort study, Tim Christen, D.O. Mook-Kanamori, Stefan Gustafsson, Lars Lind, Erik Ingelsson, Raha Pazoki, Oscar Franco, Albert Hofman, Andre Uitterlinden, Abbas Dehghan, Alexander Teumer, Sebastian Baumeister, Marcus Dörr, Markus M. Lerch, Uwe Välker, Henry Välzke, Joey Ward, Jill P Pell, Daniel J Smith, Tom Meade, Anke H. Maitland-van der Zee, Ekaterina V. Baranova, Robin Young, Ian Ford, Archie Campbell, Sandosh Padmanabhan, Michiel L Bots, Diederick E. Grobbee, Philippe Froguel, Dorothée Thuillier, Beverley Balkau, Amélie Bonnefond, Bertrand Cariou, Melissa Smart, Yanchun Bao, Meena Kumari, Anubha Mahajan, Paul M Ridker, Daniel I. Chasman, Alex P. Reiner, Leslie A Lange, Maryllyn D Ritchie, Folkert W Asselbergs, Brendan J Keating, Juan-Pablo Casas, Aroon D Hingorani, David Preiss, and Naveed Sattar were responsible for study specific analyses and critically revised the manuscript.

**Declaration of interests**

**Kees Hovingh or his institution (AMC) received honoraria for consultancy, ad boards, and/or conduct of clinical trials from: AMGEN, Aegerion, Pfizer, Astra Zeneca, Sanofi, Regeneron, KOWA, Ionis pharmaceuticals and Cerenis. GKH received research support from Aegerion, AMGEN, Sanofi, Astra Zeneca, and Synageva. Lise Lotte Nystrup Husemoen is Employed by Novo Nordisk. Bertrand Cariou has received research funding from Pfizer and Sanofi, received honoraria from AstraZeneca, Pierre Fabre, Janssen, Eli-Lilly, MSD Merck & Co., Novo-Nordisk, Sanofi, and Takeda, and has acted as a consultant/advisory panel member for Amgen,, Eli Lilly, Novo-Nordisk, Sanofi, and Regeneron. David Preiss consulted for Sanofi on two occasions in previous employment, related to PCSK9 inhibitors; and was an investigator on clinical trials of PCSK9 inhibition funded by Amgen. Naveed Sattar consulted for Amgen and Sanofi related to PCSK9 inhibitors; and was an investigator on clinical trials of PCSK9 inhibition funded by Amgen. Naveed Satter has also consulted for MSD, Boehringer Ingelheim, Janssen, and NovoNordisk. Daniel Swerdlow has consulted to Pfizer for work unrelated to this paper. Folkert W. Asselbergs is supported by a Dekker scholarship-Junior Staff Member 2014T001 – Netherlands Heart Foundation and UCL Hospitals NIHR Biomedical Research Centre. Andrzej Pająk acted as a consultant/advisory pannel member for Amgen. Erik Ingelsson is a scientific advisor and consultant for Precision Wellness, Inc. and scientific advisor for Cellink for work unrelated to this paper. All other authors declare no competing interests..**

**References**

1. 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**(9716): 735-42.

2. Preiss D, Seshasai SR, Welsh P, et al. Risk of incident diabetes with intensive-dose compared with moderate-dose statin therapy: a meta-analysis. *JAMA* 2011; **305**(24): 2556-64.

3. Swerdlow DI, Preiss D, Kuchenbaecker KB, et al. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* 2014.

4. Fall T, Xie W, Poon W, et al. Using Genetic Variants to Assess the Relationship Between Circulating Lipids and Type 2 Diabetes. *Diabetes* 2015; **64**(7): 2676-84.

5. White J, Swerdlow DI, Preiss D, et al. Association of Lipid Fractions With Risks for Coronary Artery Disease and Diabetes. *JAMA Cardiol* 2016; **1**(6): 692-9.

6. Besseling J, Kastelein JJ, Defesche JC, Hutten BA, Hovingh GK. Association between familial hypercholesterolemia and prevalence of type 2 diabetes mellitus. *JAMA* 2015; **313**(10): 1029-36.

7. Abifadel M, Varret M, Rabes JP, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* 2003; **34**(2): 154-6.

8. Cohen JC, Boerwinkle E, Mosley TH, Jr., Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* 2006; **354**(12): 1264-72.

9. Navarese EP, Kolodziejczak M, Schulze V, et al. Effects of Proprotein Convertase Subtilisin/Kexin Type 9 Antibodies in Adults With Hypercholesterolemia: A Systematic Review and Meta-analysis. *Ann Intern Med* 2015; **163**(1): 40-51.

10. Stein EA, Mellis S, Yancopoulos GD, et al. Effect of a monoclonal antibody to PCSK9 on LDL cholesterol. *N Engl J Med* 2012; **366**(12): 1108-18.

11. Hingorani A, Humphries S. Nature's randomised trials. *Lancet* 2005; **366**(9501): 1906-8.

12. Sofat R, Hingorani AD, Smeeth L, et al. Separating the mechanism-based and off-target actions of cholesteryl ester transfer protein inhibitors with CETP gene polymorphisms. *Circulation* 2010; **121**(1): 52-62.

13. Lotta LA, Sharp SJ, Burgess S, et al. Association Between Low-Density Lipoprotein Cholesterol-Lowering Genetic Variants and Risk of Type 2 Diabetes: A Meta-analysis. *JAMA* 2016; **316**(13): 1383-91.

14. Global Lipids Genetics Consortium, Willer CJ, Schmidt EM, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013; **45**(11): 1274-83.

15. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014; **46**(3): 310-5.

16. CARDIoGRAMplusC4D Consortium, Deloukas P, Kanoni S, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet* 2013; **45**(1): 25-33.

17. Scott RA, Lagou V, Welch RP, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet* 2012; **44**(9): 991-1005.

18. Soranzo N, Sanna S, Wheeler E, et al. Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways. *Diabetes* 2010; **59**(12): 3229-39.

19. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 2010; **42**(2): 105-16.

20. Randall JC, Winkler TW, Kutalik Z, et al. Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet* 2013; **9**(6): e1003500.

21. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; **518**(7538): 197-206.

22. Morris AP, Voight BF, Teslovich TM, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012; **44**(9): 981-90.

23. Fuchsberger C, Flannick J, Teslovich TM, et al. The genetic architecture of type 2 diabetes. *Nature* 2016; **536**(7614): 41-7.

24. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015; **12**(3): e1001779.

25. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**(7414): 557-60.

26. Johnson T. Efficient calculation for multi-SNP genetic risk scores. Technical report. *The Comprehensive R Archive Network* 2012: 1–2.

27. Groenwold RH, Donders AR, Roes KC, Harrell FE, Jr., Moons KG. Dealing with missing outcome data in randomized trials and observational studies. *Am J Epidemiol* 2012; **175**(3): 210-7.

28. Little RJA, Rubin DB. Statistical analysis with missing data. 2nd ed. Hoboken, New Jersey: John Wiley & Sons; 2002.

29. R Core Team. R: A Language and Environment for Statistical Computing. [https://www.r-project.org/](https://www.R-project.org). *R Foundation for Statistical Computing* 2016.

30. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015; **47**(10): 1121-30.

31. Chen L, Magliano DJ, Zimmet PZ. The worldwide epidemiology of type 2 diabetes mellitus--present and future perspectives. *Nat Rev Endocrinol* 2011; **8**(4): 228-36.

32. Shah S, Casas JP, Gaunt TR, et al. Influence of common genetic variation on blood lipid levels, cardiovascular risk, and coronary events in two British prospective cohort studies. *Eur Heart J* 2013; **34**(13): 972-81.

33. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; **11**(5): 550-60.

34. Fisher R. Statistical Methods for Research Workers. *Edinburgh: Oliver and Boyd* 1934.

35. Burgess S, Butterworth A, Malarstig A, Thompson SG. Use of Mendelian randomisation to assess potential benefit of clinical intervention. *BMJ* 2012; **345**: e7325.

36. Taylor F, Huffman MD, Macedo AF, et al. Statins for the primary prevention of cardiovascular disease. *Cochrane Database Syst Rev* 2013; (1): CD004816.

37. Ridker PM, Pradhan A, MacFadyen JG, Libby P, Glynn RJ. Cardiovascular benefits and diabetes risks of statin therapy in primary prevention: an analysis from the JUPITER trial. *Lancet* 2012; **380**(9841): 565-71.

38. Colhoun HM, Ginsberg HN, Robinson JG, et al. No effect of PCSK9 inhibitor alirocumab on the incidence of diabetes in a pooled analysis from 10 ODYSSEY Phase 3 studies. *Eur Heart J* 2016; **37**(39): 2981-9.

39. Holmes MV, Simon T, Exeter HJ, et al. Secretory phospholipase A(2)-IIA and cardiovascular disease: a mendelian randomization study. *J Am Coll Cardiol* 2013; **62**(21): 1966-76.

40. Interleukin-6 Receptor Mendelian Randomisation Analysis Consortium, Swerdlow DI, Holmes MV, et al. The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. *Lancet* 2012; **379**(9822): 1214-24.

41. Mbikay M, Sirois F, Mayne J, et al. PCSK9-deficient mice exhibit impaired glucose tolerance and pancreatic islet abnormalities. *FEBS Lett* 2010; **584**(4): 701-6.

42. Langhi C, Le May C, Gmyr V, et al. PCSK9 is expressed in pancreatic delta-cells and does not alter insulin secretion. *Biochem Biophys Res Commun* 2009; **390**(4): 1288-93.

43. Chernogubova E, Strawbridge R, Mahdessian H, et al. Common and low-frequency genetic variants in the PCSK9 locus influence circulating PCSK9 levels. *Arterioscler Thromb Vasc Biol* 2012; **32**(6): 1526-34.

44. Benn M, Nordestgaard BG, Grande P, Schnohr P, Tybjaerg-Hansen A. PCSK9 R46L, low-density lipoprotein cholesterol levels, and risk of ischemic heart disease: 3 independent studies and meta-analyses. *J Am Coll Cardiol* 2010; **55**(25): 2833-42.

**Figure legends**

**Figure 1**. Association of genetic variants in *PCSK9* with circulating LDL-cholesterol concentration.

[Figure 1]

Footnote: Effect estimates are presented as mean difference in LDL-C (mmol/L) per LDL-C lowering allele, with 95% confidence interval (CI). Results are pooled using a fixed effect model. The size of the black dots representing the point estimates is proportional to the inverse of the variance. Note that results from individual participant data are supplemented by repository data from GLGC.

**Figure 2**. Association of genetic variants in *PCSK9* with glycaemic and anthropometric biomarkers.

[Figure 2]

Footnote: Effect estimates are presented as mean difference with 95% confidence interval (CI). Associations were scaled to a 1 mmol/L reduction in LDL-C. SNP specific results are pooled using a fixed effect model and weighted gene centric score (GS) models combining all four SNP specific estimates are presented as a fixed and random effects estimates. The size of the black dots representing the point estimates is proportional to the inverse of the variance. Between SNP heterogeneity was measured as a two-sided Q-test () and an with one-sided 97·5% CI. Note that results from individual participant data are supplemented by repository data from GLGC, MAGIC, and GAINT.

**Figure 3**. Association of genetic variants in *PCSK9* with T2DM; individually and as weighted gene centric score.

[Figure 3]

Footnote: Effect estimates are presented as odds ratios (OR) for the incidence or prevalence of T2DM with 95% confidence interval (CI). Associations were scaled to a 1 mmol/L reduction in LDL-C. SNP specific results are pooled using a fixed effect model and weighted gene centric score (GS) models combining all four SNP specific estimates are presented as a fixed and random effects estimates. The size of the black dots representing the point estimates is proportional to the inverse of the variance. Between SNP heterogeneity was measured as a two-sided Q-test () and an with one-sided 97·5% CI. Results from individual participant data are supplemented by repository data from DIAGRAM.

**Figure 4**. Correlation between *PCSK9* associations with LDL-C and T2DM.

[Figure 4]

Footnote: Effect estimates are presented as mean difference in LDL-C (mmol/L) and odds ratios (OR) for the incidence or prevalence of T2DM with 95% confidence interval (CI). Associations are presented per LDL-C decreasing allele. The Pearson correlation coefficient, regression line (in grey) and its 95% confidence interval (in red) were calculated by weighting the SNPs for the inverse of the variance in the T2DM association. Excluding the SNP with the largest effect on LDL-C (rs11591147) resulted in a correlation coefficient of 0·993 and p-value of 0·437. Note the y-axis and all statistics refer to the logOR of T2DM, for presentation purposes the y-axis labels are presented as OR.