THE LANCET

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: 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; published online Sept 24. http://dx.doi.org/10.1016/S0140-6736(14)61183-1.

HMG-CoA reductase inhibition, type 2 diabetes and body weight: evidence from genetic analysis and randomized trials

Web appendix

Swerdlow DI, Preiss D, Kuchenbaecker KE, Holmes MV *et al*.

Table of Contents

Supplementary methods

Genetic studies - datasets contributing to analysis

Supplementary Table 3 describes the details of studies contributing individual participant-level data (IPD) to the genetic analysis. The participating studies are detailed below.

- Prospective cohort studies

Twenty-five prospective cohort studies (including birth cohorts) contributed genotype and phenotype (either biomarkers, clinical events, or both) data to the genetic analysis. The designs of these studies have been described previously.

From the UCL-LSHTM-Edinburgh-Bristol (UCLEB) Consortium: British Regional Heart Study (BRHS)[1], the British Women's Heart and Health Study (BWHHS)[2], the Caerphilly Prospective Study (CaPS)[3], the Edinburgh Artery Study (EAS)[4], the English Longitudinal Study of Aging (ELSA)[5], the Edinburgh Type 2 Diabetes Study (ET2DS)[6], the Northwick Park Heart Study II (NPHS-II)[7], the Whitehall-II (WHII) study[8], and the MRC National Survey of Health and Development (MRC NSHD, 1946 British Birth Cohort)[9].

From the National Heart, Lung and Blood Institute (NHLBI) Candidate Gene Association Resource (CARe) Consortium: the Atherosclerosis Risk in Communities Study (ARIC)[10], the Coronary Artery Risk Development in Young Adults (CARDIA) Study[11], the Cleveland Family Study (CFS)[12,13], the Cardiovascular Health Study (CHS)[14], the Framingham Heart Study (FHS)[15], and the Multi-Ethnic Study of Atherosclerosis (MESA)[16].

From the Health, Alcohol and Psychosocial factors In Eastern Europe (HAPIEE) study[17]: the Czech Republic (HAPIEE-CZ), Lithuania (HAPIEE-LT), Poland (HAPIEE-PL) and Russia (HAPIEE-RU) arms of the study.

Also, the Copenhagen City Heart Study (CCHS)[18], the MRC Ely Cohort[19], the MRC Fenland Cohort[20], the IMPROVE study[21], the Prevention of Renal and Vascular Endstage Disease Study (PREVEND)[22], and the Second Manifestations of Arterial Disease (SMART) Study[23].

- Case-control and case-cohort studies

Three nested case-control studies were included - the InterAct consortium of studies[24], the Utrecht Cardiovascular Pharmacogenetics study (UCP)[25] and the Women's Health Initiative (WHI)[26].

- Intervention trials

Five RCT studies were included - the Aspirin in Asymptomatic Atherosclerosis (AAA) trial[27], the Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER) trial[28], the PROSPER trial follow-up study[29], the Thrombosis Prevention Trial follow-up study (TPT)[30], and the Women's Genome Health Study (WHGS)[31]. These trials were treated as prospective cohorts for the purposes of these analyses.

- GWA studies

Two GWA study consortia contributed data.

Data from the Meta-Analysis of Glucose and Insulin-related traits Consortium (MAGIC) were included in the form of single-SNP lookups[1]. Estimates of the association between the lead *HMGCR* SNPs and fasting insulin were extracted from analysis using the Illumina Metabochip SNP genotyping platform.

An estimate of the association of the lead *HMGCR* SNPs with body weight were accessed from publicly available data from the Genetic Investigation of Anthropometric Traits (GIANT) Consortium[32] at http://csg.sph.umich.edu/locuszoom/ (Accessed 11 August 2011).

From the Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium[33], we included the following studies: deCODE[34,35], the Diabetes Gene Discovery Group (DGDG)[36], the Diabetes Genetics Initiative (DGI)[36,37], the European Special Population Network (EUROSPAN)[37–40], the Cooperative Health Research in the Region of Augsburg, Southern Germany Study (KORAgen)[41–43], the Rotterdam Study[44], and the Wellcome Trust Case-Control Consortium (WTCCC)[45].

Genetic studies - Biomarker measurement and definitions

Availability of biomarker data in genetic studies is detailed in **Supplementary Table 4.**

All biomarkers were measured using validated protocols and assays reported previously. Biomarker data used in the present analyses were taken from either the baseline phase of data collection in each study, or the next phase soonest after baseline at which data on the greatest number of biomarkers were available. Where certain biomarker variables were unavailable at the principal survey phase, they were included from closest subsequent phase with data available.

In some studies (BRHS, HAPIEE-Russia, HAPIEE-Lithuania, HAPIEE-Poland, HAPIEE-Czech Republic, CaPS and Whitehall II), LDL-cholesterol (LDL-C) concentration was not measured directly but was derived using the Friedewald formula[46]. In studies where data on total, LDL- and non-HDL-cholesterol were available in units of mg/L, these variables were converted to mmol/L using a multiplication factor of 0.02586.

Genetic studies - T2DM clinical outcome

Definitions of T2D cases differed between studies. Common definitions were therefore set in order to accommodate data from as many studies as possible whilst avoiding reporting or measurement bias. Studies used validated biochemical criteria or physician diagnosis, and others self-report to identify T2D cases (details are given in **Supplementary Table 5)**. We included all T2D cases, although, where possible, unvalidated self-reported cases were excluded.

In the majority of studies, the type of diabetes mellitus (DM) - i.e. type 1, type 2 or others - was not recorded. However, given the relative preponderance of type 2 DM in the general population and the mean age of study participants, we assumed that the overwhelming majority of cases were T2D. In ET2DS, where all participants were diabetic, individuals free from T2D from EAS were used as controls. EAS was selected as the most appropriate source of controls since its participants were recruited from the same area of Scotland as those in ET2DS.

Genetic studies - SNP selection

The aim of the SNP selection process was to identify a minimal set of variants with associations with LDL-C concentration that were both statistically strong and of greatest possible magnitude. Moreover, the highest possible minor allele frequency (MAF) was sought to increase the power of the analysis. To identify suitable SNPs for use as instruments in this analysis, 4,678 individuals in the Whitehall II study (an observational longitudinal study of UK civil servants) were genotyped using the IBC HumanCVD BeadChip (Cardiochip), a gene-centric platform including c. 48,000 SNPs in around 2,100 genes implicated in CVD[47].

Associations of the 38 SNPs included on the array at the *HMGCR* locus with LDL-C concentration were estimated using univariate linear regression models in PLINK [48]. We evaluated suitability of SNPs at the *HMGCR* locus for inclusion based on these parameters: i) p-value $<1x10^{-5}$ for association with LDL-C; ii) linkage disequilibrium (LD) between SNPs assessed using data from the 1000 Genomes Project Pilot 1 (using SNAP pairwise LD tool - www.broadinstitute.org/mpg/snap/); and, iii) a minor allele frequency (MAF) ≥20%. Within the 38 SNPs at the HMGCR locus included on the Cardiochip, we retained only one SNP from groups in strong LD ($r^2>0.85$) (**Supplementary Table 1**). Two SNPs, rs17238484 and rs12916 (between-SNP LD $R^2=0.37$) were selected for analysis in the collaborating studies, with the rs17238484 SNP designated the lead variant. Both rs17238484[49] and rs12916[50] have previously been associated with plasma LDL-C at a genome-wide level of statistical significance. Genotype data in collaborating studies were excluded if the allele call rate was <90% or the Hardy-Weinberg equilibrium (HWE) p-value was $<1x10^{-5}$ (**Supplementary Table 2**).

To confirm the influence of *HMGCR* mRNA transcript levels on LDL, we carried out imputation and eQTL analysis of a liver gene expression dataset and genotypes for 966 human liver samples of unrelated European-Americans[51]. Colocalisation analysis was run using the liver eQTL data together with LDL summary data from a publicly available meta-analysis of LDL in more than 100,000 individuals of European ancestry [50], as previously described (http://arxiv.org/abs/1305.4022).

In studies where the lead *HMGCR* SNPs were not directly genotyped, proxy SNPs were used in the analysis. Proxies were defined on the basis of their LD with each lead SNP, with a threshold of $r^2>0.85$.

Genetic studies - genotyping and genotype data quality control

Genotypes were coded as 0 (GG), 1 (GT), and 2 (TT) and included in regression models as the predictor (independent) variable. Since this coding assumes the T allele to be the effect allele, the sign of regression model beta-coefficients was inverted in order to present the effect of G allele carriage. Existing genotype data were used where available, and new genotyping commissioned where necessary. Details of genotyping platforms and data quality control are shown in **Supplementary Table 2**.

Genetic studies - statistical analysis

- Participant inclusion and relatedness

Only individuals with data available on the lead *HMGCR* SNPs (or their proxies) and at least one phenotype (biomarker or clinical endpoint) were included in the analysis. To avoid bias caused by relatedness, in studies where data on family structure were available, only the eldest member of a group of related individuals was included.

- Associations of *HMGCR* genotype with biomarkers

Mean differences in biomarkers between genotype classes were estimated using the 'metan' command in Stata v11.2 (Stata Corp., College Station, Texas), using an inverse-variance fixed effects meta-analysis model. For $\log_e$ biomarker variables, the summary meta-analysis estimate was exponentiated and subtracted from 1 to yield a proportional difference in geometric means between genotype classes. The Stata 'regress' command (or equivalent in other computer packages) was used to estimate the mean difference in biomarker level associated with carriage of each additional copy of the lead SNP effect allele. Here too, the proportional difference in geometric mean was estimated for $\log_e$-transformed variables the same technique. Within-study estimates of the per-allele SNP association with each biomarker were combined using inverse-variance fixed-effects meta-analysis.

- Subgroup analysis of SNP-biomarker associations

In order to assess potential effect modification or confounding in our estimates of genetic associations with biomarkers, we stratified the analysis according to a number of pre-specified subgroups. These were:

i. Within-study tertiles of non-HDL-C (as a surrogate for LDL-C for data were available in a greater number of studies);
ii. Users and non-users of lipid-lowering drugs;
iii. Normal weight (BMI <25), overweight (BMI ≥25-<30) and obese (BMI ≥30) individuals;
iv. Males and females; and,
v. Individuals with and without prevalent or incident T2D

Biomarker associations were estimated as described above, and heterogeneity between subgroup strata assessed with meta-regression modelling of combined meta-analysis effect estimates (using the 'metareg' command in Stata). These models sought differences between binary subgroup strata, and evidence of linear association in subgroups with ≥3 strata.

- Associations of *HMGCR* genotype with T2D outcome

Where data were available, the association of the lead *HMGCR* SNPs with risk of prevalent and incident T2D was estimated under genotypic and additive models, similarly to the associations with biomarkers. Odds ratios (ORs) for rs17238484 GT vs TT and for GG vs TT genotype, and for rs12916 AA vs AG and AA vs GG were estimated using fixed effects Mantel-Haenszel meta-analysis (Stata 'metan' command). The OR per LDL-lowering allele for each SNP was estimated using logistic regression models within each study (Stata 'logit' command), and within-study estimates were combined using inverse-variance fixed effects meta-analysis in Stata, as above. Meta-analyses were repeated using a random effects model.

Statin treatment trials included

Trials of participants with organ transplants or on dialysis were excluded, as were those with differences in participant follow-up between treatment arms, or which investigated dual lipid-lowering therapy. These meta-analyses included data from 13 intervention trials comparing statin treatment with placebo (ten trials) or standard care (three trials)[52], and from five trials comparing high- to moderate-dose statin treatment[53], respectively. At the time of their publication, data were unavailable from three eligible trials (CARE[54], SPARCL[55] and LIPS[56]). Data on new-onset T2D have subsequently been published for SPARCL[57], and data for LIPS were made available by its industrial sponsor for the present analysis, which now includes 20 trials.

Below are the randomised statin trials included in the present analysis, with references to their respective published reports. The trials are listed in chronological order of publication:

Scandinavian Simvastatin Survival Study (4S)[58]

West of Scotland Coronary Protection Study (WOSCOPS)[59]

Air Force/Texas Coronary Atherosclerosis Prevention Study (AFCAPS TexCAPS)[60]

Long-Term Intervention with Pravastatin in Ischaemic Disease Study (LIPID)[61]

GISSI-Prevenzione [62]

Lescol® Intervention Prevention Study (LIPS)[63]

Heart Protection Study (HPS)[64]

Pravastatin in elderly individuals at risk of vascular disease (PROSPER)[65]

Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial - Lipid-lowering Treatment arm (ALLHAT-LLT)[66]

Anglo-Scandinavian Cardiac Outcomes Trial--Lipid Lowering Arm (ASCOT-LLA)[67]

Pravastatin or Atorvastatin Evaluation and Infection Therapy–Thrombolysis in Myocardial Infarction 22 (PROVE-IT TIMI 22)[68]

A to Z Trial[69]

Treating to New Targets Trial (TNT)[70]

Incremental Decrease in End Points Throught Aggressive Lipid Lowering Study (IDEAL)[71]

Stroke Prevention by Aggresive Reduction in Cholesterol Levels Trial (SPARCL)[55]

MEGA Trial[72]

Controlled Rosuvastatin Multinational Trial in Heart Failure (CORONA)[73]

GISSI-Heart Failure Trial (GISSI-HF)[74]

Justification for the Use of Statins in Primary Prevention: An Intervention Trial Evaluating Rosuvastatin (JUPITER)[75]

Study of the Effectiveness of Additional Reductions in Cholesterol and Homocysteine (SEARCH)[76]

Statin treatment trials - T2D outcome definition

Diagnostic criteria for T2D varied among the 20 contributing trials, but included one or more of, (i) physician-reported diagnosis of T2D; (ii) commencement of glucose-lowering medication; or, (iii) fasting glucose >7.0mmol/L (on at least one occasion; see **Table 1**); and (iv) T2D defined according to World Health Organisation 1999 criteria. In a subset of trials, T2D was diagnosed by the presence of elevated fasting glucose. In trials where plasma glucose was measured frequently, T2D was diagnosed after two elevated glucose readings, and after one in trials with less frequent testing.

Supplementary results

HMGCR SNP expression and co-localisation analysis

Co-localisation analysis yields a high probability of a shared signal for the eQTL in *HMGCR* and LDL-C lipid biomarkers, suggesting that *HMGCR* gene expression mediates the LDL signal (**Supplementary Figure 2**). Extensive linkage disequilibrium at the eQTL signal, however, limits the power of co-localisation analysis at this locus. The fact that other genes 1Mb around the *HMGCR* gene have a probability of co-localisation less than 3% increases the probability that *HMGCR* is the gene responsible for the LDL signal.

The T allele of rs12916 is associated with lower expression of *HMGCR* in the liver (effect = 0.054, standard error=0.012; $p=1.30x10^{-5}$; MAF = 0.39, imputation $R^2= 0.99$). Out of all the SNPs used in the co-localisation analysis, rs12916 was the SNP with the highest probability of driving the shared association between eQTL and LDL-C (54%, see **Supplementary Figure 1**). Data were unavailable in the dataset for the *HMGCR* rs17238484 SNP or its proxies.

1 Walker M, Whincup PH, Shaper AG. The British Regional Heart Study 1975-2004. *Int J Epidemiol* 2004; **33**: 1185–92.

2 Lawlor DA, Bedford C, Taylor M, Ebrahim S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J Epidemiol Community Health* 2003; **57**: 134–40.

3 Caerphilly and Speedwell Collaborative Group. Caerphilly and Speedwell collaborative heart disease studies. The Caerphilly and Speedwell Collaborative Group. *J Epidemiol Community Health* 1984; **38**: 259–62.

4 Fowkes FG, Housley E, Cawood EH, Macintyre CC, Ruckley CV, Prescott RJ. Edinburgh Artery Study: prevalence of asymptomatic and symptomatic peripheral arterial disease in the general population. *Int J Epidemiol* 1991; **20**: 384–92.

5 Steptoe A, Breeze E, Banks J, Nazroo J. Cohort Profile: The English Longitudinal Study of Ageing. *Int J Epidemiol* 2012. doi:10.1093/ije/dys168.

6 Price JF, Reynolds RM, Mitchell RJ, *et al.* The Edinburgh Type 2 Diabetes Study: study protocol. *BMC Endocr Disord* 2008; **8**: 18.

7 Miller GJ, Bauer KA, Barzegar S, *et al.* The effects of quality and timing of venepuncture on markers of blood coagulation in healthy middle-aged men. *Thromb Haemost* 1995; **73**: 82–6.

8 Marmot M, Brunner E. Cohort Profile: the Whitehall II study. *Int J Epidemiol* 2005; **34**: 251–6.

9 Wadsworth M, Kuh D, Richards M, Hardy R. Cohort Profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development). *Int J Epidemiol* 2006; **35**: 49–54.

10 ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* 1989; **129**: 687–702.

11 Friedman GD, Cutter GR, Donahue RP, *et al.* CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol* 1988; **41**: 1105–16.

12 Redline S, Tishler PV, Tosteson TD, *et al.* The familial aggregation of obstructive sleep apnea. *Am J Respir Crit Care Med* 1995; **151**: 682–7.

13 Buxbaum SG, Elston RC, Tishler PV, Redline S. Genetics of the apnea hypopnea index in Caucasians and African Americans: I. Segregation analysis. *Genet Epidemiol* 2002; **22**: 243–53.

14 Fried LP, Borhani NO, Enright P, *et al.* The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1991; **1**: 263–76.

15 Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health* 1951; **41**: 279–81.

16 Bild DE, Bluemke DA, Burke GL, *et al.* Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* 2002; **156**: 871–81.

17 Peasey A, Bobak M, Kubinova R, *et al.* Determinants of cardiovascular disease and other non-communicable diseases in Central and Eastern Europe: rationale and design of the HAPIEE study. *BMC Public Health* 2006; **6**: 255.

18 Frikke-Schmidt R, Nordestgaard BG, Stene MCA, *et al.* 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**: 2524–32.

19 Williams DR, Wareham NJ, Brown DC, *et al.* Undiagnosed glucose intolerance in the community: the Isle of Ely Diabetes Project. *Diabet Med* 1995; **12**: 30–5.

20 Medical Research Council. MRC Fenland Study. MRC Fenland Study. 2010.http://www.mrc-epid.cam.ac.uk/Research/Studies/Fenland/index.html (accessed 25 Oct2010).

21 Baldassarre D, Nyyssönen K, Rauramaa R, *et al.* Cross-sectional analysis of baseline data to identify the major determinants of carotid intima-media thickness in a European population: the IMPROVE study. *Eur Heart J* 2010; **31**: 614–22.

22 Diercks GF, Janssen WM, van Boven AJ, *et al.* Rationale, design, and baseline characteristics of a trial of prevention of cardiovascular and renal disease with fosinopril and pravastatin in nonhypertensive, nonhypercholesterolemic subjects with microalbuminuria (the Prevention of REnal and Vascular ENdstage Disease Intervention Trial [PREVEND IT]). *Am J Cardiol* 2000; **86**: 635–8.

23 Simons PC, Algra A, van de Laak MF, Grobbee DE, van der Graaf Y. Second manifestations of ARTerial disease (SMART) study: rationale and design. *Eur J Epidemiol* 1999; **15**: 773–81.

24 InterAct Consortium, Langenberg C, Sharp S, *et al.* Design and cohort description of the InterAct Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type 2 diabetes in the EPIC Study. *Diabetologia* 2011; **54**: 2272–82.

25 Peters BJM, Pett H, Klungel OH, *et al.* Genetic variability within the cholesterol lowering pathway and the effectiveness of statins in reducing the risk of MI. *Atherosclerosis* 2011; **217**: 458–64.

26 Anderson GL, Manson J, Wallace R, *et al.* Implementation of the Women's Health Initiative study design. *Ann Epidemiol* 2003; **13**: S5–17.

27 Fowkes FGR, Price JF, Stewart MCW, *et al.* Aspirin for prevention of cardiovascular events in a general population screened for a low ankle brachial index: a randomized controlled trial. *JAMA* 2010; **303**: 841–8.

28 Ridker PM, JUPITER Study Group. Rosuvastatin in the primary prevention of cardiovascular disease among patients with low levels of low-density lipoprotein cholesterol and elevated high-sensitivity C-reactive protein: rationale and design of the JUPITER trial. *Circulation* 2003; **108**: 2292–7.

29 Shepherd J, Blauw GJ, Murphy MB, *et al.* The design of a prospective study of Pravastatin in the Elderly at Risk (PROSPER). PROSPER Study Group. PROspective Study of Pravastatin in the Elderly at Risk. *Am J Cardiol* 1999; **84**: 1192–7.

30 MRC GPRF. Thrombosis prevention trial: randomised trial of low-intensity oral anticoagulation with warfarin and low-dose aspirin in the primary prevention of ischaemic heart disease in men at increased risk. The Medical Research Council's General Practice Research Framework. *Lancet* 1998; **351**: 233–41.

31 Ridker PM, Chasman DI, Zee RYL, *et al.* Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. *Clin Chem* 2008; **54**: 249–55.

32 Speliotes EK, Willer CJ, Berndt SI, *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 2010; **42**: 937–48.

33 Voight BF, Scott LJ, Steinthorsdottir V, *et al.* Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 2010; **42**: 579–89.

34 Steinthorsdottir V, Thorleifsson G, Reynisdottir I, *et al.* A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* 2007; **39**: 770–5.

35 Thorleifsson G, Walters GB, Gudbjartsson DF, *et al.* Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat Genet* 2009; **41**: 18–24.

36 Sladek R, Rocheleau G, Rung J, *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007; **445**: 881–5.

37 Rudan I, Campbell H, Rudan P. Genetic epidemiological studies of eastern Adriatic Island isolates, Croatia: objective and strategies. *Coll Antropol* 1999; **23**: 531–46.

38 Aulchenko YS, Heutink P, Mackay I, *et al.* Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* 2004; **12**: 527–34.

39 Pattaro C, Marroni F, Riegler A, *et al.* The genetic study of three population microisolates in South Tyrol (MICROS): study design and epidemiological perspectives. *BMC Med Genet* 2007; **8**: 29.

40 Hicks AA, Pramstaller PP, Johansson A, *et al.* Genetic determinants of circulating sphingolipid concentrations in European populations. *PLoS Genet* 2009; **5**: e1000672.

41 Wichmann H-E, Gieger C, Illig T. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* 2005; **67 Suppl 1**: S26–30.

42 Herder C, Illig T, Baumert J, *et al.* RANTES/CCL5 gene polymorphisms, serum concentrations, and incident type 2 diabetes: results from the MONICA/KORA Augsburg case-cohort study, 1984-2002. *Eur J Endocrinol* 2008; **158**: R1–5.

43 Huth C, Heid IM, Vollmert C, *et al.* IL6 gene promoter polymorphisms and type 2 diabetes: joint analysis of individual participants' data from 21 studies. *Diabetes* 2006; **55**: 2915–21.

44 Hofman A, Breteler MMB, van Duijn CM, *et al.* The Rotterdam Study: objectives and design update. *Eur J Epidemiol* 2007; **22**: 819–29.

45 Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007; **447**: 661–78.

46 Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972; **18**: 499–502.

47 Keating BJ, Tischfield S, Murray SS, *et al.* Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS ONE* 2008; **3**: e3583.

48 Purcell S, Neale B, Todd-Brown K, *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–75.

49 Kathiresan S, Melander O, Guiducci C, *et al.* Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* 2008; **40**: 189–97.

50 Teslovich TM, Musunuru K, Smith AV, *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010; **466**: 707–13.

51 Schadt EE, Woo S, Hao K. Bayesian method to predict individual SNP genotypes from gene expression data. *Nature Genetics* 2012; **44**: 603–8.

52 Sattar N, Preiss D, Murray HM, *et al.* Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. *The Lancet* 2010; **375**: 735–42.

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

54 Sacks FM, Pfeffer MA, Moye LA, *et al.* The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events Trial investigators. *N Engl J Med* 1996; **335**: 1001–9.

55 Amarenco P, Bogousslavsky J, Callahan A 3rd, *et al.* High-dose atorvastatin after stroke or transient ischemic attack. *N Engl J Med* 2006; **355**: 549–59.

56 Serruys PWJC, de Feyter P, Macaya C, *et al.* Fluvastatin for prevention of cardiac events following successful first percutaneous coronary intervention: a randomized controlled trial. *JAMA* 2002; **287**: 3215–22.

57 Waters DD, Ho JE, DeMicco DA, *et al.* Predictors of new-onset diabetes in patients treated with atorvastatin: results from 3 large randomized clinical trials. *J Am Coll Cardiol* 2011; **57**: 1535–45.

58 SSSS Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994; **344**: 1383–9.

59 Shepherd J, Cobbe SM, Ford I, *et al.* Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med* 1995; **333**: 1301–7.

60 Downs JR, Clearfield M, Weis S, *et al.* Primary prevention of acute coronary events with lovastatin in men and women with average cholesterol levels: results of AFCAPS/TexCAPS. Air Force/Texas Coronary Atherosclerosis Prevention Study. *JAMA* 1998; **279**: 1615–22.

61 LIPID Study Group. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. The Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. *N Engl J Med* 1998; **339**: 1349–57.

62 GISSI-Prevenzione Investigators. Results of the low-dose (20 mg) pravastatin GISSI Prevenzione trial in 4271 patients with recent myocardial infarction: do stopped trials contribute to overall knowledge? GISSI Prevenzione Investigators (Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico). *Ital Heart J* 2000; **1**: 810–20.

63 Serruys PW, De Feyter PJ, Benghozi R, Hugenholtz PG, Lesaffre E. The Lescol(R) Intervention Prevention Study (LIPS): a double-blind, placebo-controlled, randomized trial of the long-term effects of fluvastatin after

successful transcatheter therapy in patients with coronary heart disease. *Int J Cardiovasc Intervent* 2001; **4**: 165–72.

64 Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* 2002; **360**: 7–22.

65 Shepherd J, Blauw GJ, Murphy MB, *et al.* Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *Lancet* 2002; **360**: 1623–30.

66 ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in moderately hypercholesterolemic, hypertensive patients randomized to pravastatin vs usual care: The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT-LLT). *JAMA* 2002; **288**: 2998–3007.

67 Sever PS, Dahlöf B, Poulter NR, *et al.* Prevention of coronary and stroke events with atorvastatin in hypertensive patients who have average or lower-than-average cholesterol concentrations, in the Anglo-Scandinavian Cardiac Outcomes Trial--Lipid Lowering Arm (ASCOT-LLA): a multicentre randomised controlled trial. *Lancet* 2003; **361**: 1149–58.

68 Cannon CP, Braunwald E, McCabe CH, *et al.* Intensive versus moderate lipid lowering with statins after acute coronary syndromes. *N Engl J Med* 2004; **350**: 1495–504.

69 De Lemos JA, Blazing MA, Wiviott SD, *et al.* Early intensive vs a delayed conservative simvastatin strategy in patients with acute coronary syndromes: phase Z of the A to Z trial. *JAMA* 2004; **292**: 1307–16.

70 LaRosa JC, Grundy SM, Waters DD, *et al.* Intensive lipid lowering with atorvastatin in patients with stable coronary disease. *N Engl J Med* 2005; **352**: 1425–35.

71 Pedersen TR, Faergeman O, Kastelein JJP, *et al.* High-dose atorvastatin vs usual-dose simvastatin for secondary prevention after myocardial infarction: the IDEAL study: a randomized controlled trial. *JAMA* 2005; **294**: 2437–45.

72 Nakamura H, Arakawa K, Itakura H, *et al.* Primary prevention of cardiovascular disease with pravastatin in Japan (MEGA Study): a prospective randomised controlled trial. *Lancet* 2006; **368**: 1155–63.

73 Kjekshus J, Apetrei E, Barrios V, *et al.* Rosuvastatin in older patients with systolic heart failure. *N Engl J Med* 2007; **357**: 2248–61.

74 Gissi-HF Investigators, Tavazzi L, Maggioni AP, *et al.* Effect of rosuvastatin in patients with chronic heart failure (the GISSI-HF trial): a randomised, double-blind, placebo-controlled trial. *Lancet* 2008; **372**: 1231–9.

75 Ridker PM, Danielson E, Fonseca FAH, *et al.* Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N Engl J Med* 2008; **359**: 2195–207.

76 Study of the Effectiveness of Additional Reductions in Cholesterol and Homocysteine (SEARCH) Collaborative Group, Armitage J, Bowman L, *et al.* Intensive lowering of LDL cholesterol with 80 mg versus 20 mg simvastatin daily in 12,064 survivors of myocardial infarction: a double-blind randomised trial. *Lancet* 2010; **376**: 1658–69.

Acknowledgements

D Swerdlow acknowledges the support of the UCL MBPhD Programme.

The **EPIC-NL** investigators thank GBA, Statistics Netherlands, Dutch Cancer Registry and the Institute PHARMO for providing data on vital status, cause of death and occurrence of cancer and other chronic diseases.

Authors representing the **Fenland Study** thank the Fenland Study Investigators, Study Co-ordination team and the Epidemiology Field, Data and Technical teams as well as the study participants and general practice staff for help with recruitment. Biochemical assays were performed by the National Institute for Health Research, Cambridge Biomedical Research Centre, Core Biochemistry Assay Laboratory, and the Cambridge University Hospitals NHS Foundation Trust, Department of Clinical Biochemistry.

The **Ely Study** investigators are grateful to all those who participated and to the staff of the St Mary's Street surgery, Ely, and all those who worked on the study.

The **HAPIEE** study investigator would like to thank researchers, interviewers and participants in Novosibirsk, Krakow, Kaunas, Havířov/Karviná, Jihlava, Ústí nad Labem, Liberec, Hradec Králové, and Kroměříz without whose help the HAPIEE study would not have been possible.

Authors representing the **IMPROVE Study** express their deep and sincere appreciation to all members of the IMPROVE group for their time and extraordinary commitment. The IMPROVE study group: Dipartimento di Scienze Farmacologiche e Biomolecolari, Università di Milano,, Milan, Italy: C.R. Sirtori, S. Castelnuovo; Centro Cardiologico Monzino, IRCCS, Milan Italy: M. Amato, B. Frigerio, A. Ravani, D. Sansaro, C. Tedesco; Atherosclerosis Research Unit, Departments of Medicine and Cardiology, Karolinska University Hospital Solna, & Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden: M. Ahl, G. Blomgren, M.J. Eriksson, P. Fahlstadius, M. Heinonen, L. Nilson; University College of London, Department of Medicine, Rayne Institute, London, United Kingdom: J. Cooper, J. Acharya; Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland: K. Huttunen, E. Rauramaa, H Pekkarinen, I.M. Penttila, J. Törrönen; Department of Medicine, University Medical Center Groningen, Groningen & Isala Clinics Zwolle, Department of Medicine; the Netherlands: A.I. van Gessel, A.M van Roon, G.C. Teune, W.D. Kuipers, M. Bruin, A. Nicolai, P. Haarsma-Jorritsma, D.J. Mulder, H.J.G. Bilo, G.H. Smeets; Assistance Publique - Hôpitaux de Paris; Service Endocrinologie-Metabolisme, Groupe Hôpitalier Pitié-Salpetriere, Unités de Prévention Cardiovasculaire, Paris, France: J.L. Beaudeux, J.F. Kahn, V. Carreau, A. Kontush; Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio Campus: J. Karppi, T. Nurmi, K. Nyyssönen, R. Salonen, T.P. Tuomainen, J.Tuomainen, J. Kauhanen; Internal Medicine, Angiology and Arteriosclerosis Diseases, Department of Clinical and Experimental Medicine, University of Perugia, Perugia, Italy: G. Vaudo, A. Alaeddin, D. Siepi, G. Lupattelli, G. Schillaci.

The authors thank the **WHI** investigators and staff for their dedication, and the study participants for making the program possible. A listing of WHI investigators can be found at http://www.whiscience.org/publications/WHI_investigators_shortlist_2010-2015.pdf.

The **CARe Consortium** wishes to acknowledge the support of the National Heart, Lung, and Blood Institute and the contributions of the research institutions, study investigators, field staff and study participants in creating this resource for biomedical research (NHLBI contract number HHSN268200960009C). The following nine parent studies have contributed parent study data, ancillary study data, and DNA samples through the Massachusetts Institute of Technology - Broad Institute (N01-HC-65226) to create this genotype/phenotype database for wide dissemination to the biomedical research community: the Atherosclerosis Risk in Communities (ARIC) study, the Cardiovascular Health Study (CHS), the Cleveland Family Study (CFS), the Cooperative Study of Sickle Cell Disease (CSSCD), the Coronary Artery Risk Development in Young Adults (CARDIA) study, the

Framingham Heart Study (FHS), the Jackson Heart Study (JHS), the Multi-Ethnic Study of Atherosclerosis (MESA), and the Sleep Heart Health Study (SHHS).

The **Caerphilly Prospective study** was undertaken by the former MRC Epidemiology Unit (South Wales) and was funded by the Medical Research Council of the United Kingdom.

Authors representing the **ARIC study** thank the staff and participants of the ARIC study for their important contributions.

The authors representing the **AAA**, **EAS** and **ET2DS** studies thank the co-investigators, staff and participants for their invaluable contributions.

Funding information

The **UCLEB (UCL-LSHTM-Edinburgh-Bristol) Consortium** (including BRHS, BWHHS, CaPS, EAAS, ELSA, ET2DS, NPHS-II and WHII) was supported by a programme grant from the British Heart Foundation (RG/10/12/28456).

The **Utrecht Cardiovascular Pharmacogenetics Study (UCP)** was funded by Veni grant Organization for Scientific Research (NWO), Grant no. 2001.064 Netherlands Heart Foundation (NHS), and TI Pharma Grant T6-101 Mondriaan. The Department of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, has received unrestricted research funding from the Netherlands Organisation for Health Research and Development (ZonMW), the Dutch Health Care Insurance Board (CVZ), the Royal Dutch Pharmacists Association (KNMP), the private-public funded Top Institute Pharma (www.tipharma.nl, includes co-funding from universities, government, and industry), the EU Innovative Medicines Initiative (IMI), EU 7th Framework Program (FP7), the Dutch Medicines Evaluation Board, the Dutch Ministry of Health and industry (including GlaxoSmithKline, Pfizer, and others)."

The Candidate Gene Association Resource (CARe) was supported by contract no. HHSN268200625226C and from the National Institutes of Health/NHLBI, and subcontract no. 5215810-55000000041 to C.L.W. A full listing of the grants and contracts that have supported CARe is provided at http://public.nhlbi.nih.gov/GeneticsGenomics/home/care.aspx. MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support is provided by grants and contracts N01 HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169 and RR-024156.

The **Cardiovascular Health Study (CHS)** was supported by contracts HHSN268201200036C, HHSN268200800007C, N01 HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and grant HL080295 from the National Heart, Lung, and Blood Institute (NHLBI), with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at www.CHS-NHLBI.org.

The **EPIC-Netherlands (EPIC-NL)** study was funded by 'Europe against Cancer' Programme of the European Commission (SANCO); the Dutch Ministry of Public Health, Welfare and Sports (formerly Ministry of Welfare, Public Health and Culture); the Dutch Cancer Society; ZonMW the Netherlands Organisation for Health Research and Development; and World Cancer Research Fund (WCRF). Genotyping of the IBC-chip was funded by IOP Genomics grant IGE05012 from NL Agency.

The **Fenland Study** is funded by the Wellcome Trust and the MRC.

The **Ely Study** was funded by the MRC, Diabetes UK, and Eastern Region NHS R&D.

The **IMPROVE study** was supported by the European Commission (Contract number: QLG1- CT- 2002-00896), the Swedish Heart-Lung Foundation, the Swedish Research Council (projects 8691 and 0593), the Foundation for Strategic Research, the Stockholm County Council (project 562183), the Foundation for Strategic Research, the Academy of Finland (Grant #110413), the Knut and Alice Wallenberg Foundation, the Torsten and Ragnar Söderberg Foundation, the Strategic Cardiovascular Programme of the Karolinska Institutet and the British Heart Foundation (RG2008/014).

The **Prospective Study of Pravastatin in the Elderly at Risk (PROSPER) study** was supported by a European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° HEALTH-F2-2009-223004 PHASE, grants from the Interuniversity Cardiology Institute of the Netherlands (ICIN) and the Durrer Center for Cardiogenetic Research both Institutes of the

Netherlands Royal Academy of Arts and Sciences (KNAW) the Netherlands Consortium for Healthy Ageing (NCHA).
The **Women's Health Initiative (WHI)** is funded by the National Heart, Lung and Blood Institute (grant RFP-NHLBI-WH-11-10 - Women's Health Initiative Extension 2010-2015), US Department of Health and Human Services. Clinical trial registration NCT00000611.

Funding for the **AFCAPS/TexCAPS** trial was provided by Merck & Co.

The **ASCOTT-LLA** trial was funded by Pfizer Inc.

GISSI-PREVENTION was funded by Bristol-Myers Squibb, Pharmacia-Upjohn, Societa Prodotti, Antibiotici, and Pfizer.

GISSI-HF was funded by SPA, Pfizer, Sigma Tau, and AstraZeneca

DNA extraction in the **Caerphilly Prospective Study (CaPS)** study was funded by a grant from the MRC.

The **British Women's Heart and Health Study (BWHHS)** is supported by funding from the British Heart Foundation (BHF) and the Department of Health Policy Research Programme (England). HumanCVD genotyping of the BWHHHS was funded by the BHF [PG/07/131/24254].

The **British Regional Heart Study (BRHS)** is a BHF research group and is supported by a BHF programme grant RG/08/013/25942.

The **MRC National Survey of Health and Development (MRC NSHD)** is supported by the MRC (MC_UU_12019/1).

The **Prevention of Renal and Vascular Endstage Disease Study (PREVEND)** genetics is supported by the Dutch Kidney Foundation (Grant E033), the EU project grant GENECURE (FP-6 LSHM CT 2006 037697), the National Institutes of Health (grant LM010098), the Netherlands organisation for health research and development (NWO VENI grant 916.761.70), and the Dutch Inter University Cardiology Institute Netherlands (ICIN).

The **Copenhagen City Heart Study (CCHS)** is supported by the Danish MRC (grant no. 10-083788), the Research Fund at Rigshospitalet, Copenhagen University Hospital, Chief Physician Johan Boserup and Lise Boserup's Fund, Ingeborg and Leo Dannin's Grant, Henry Hansen and Wife's grant, and a grant from the Odd Fellow Order.

The **Whitehall II Study (WHII)** is supported by the National Institute on Aging, the MRC (K013351), the British Heart Foundation and the National Heart, Lung and Blood Institute (HL036310).

The **Atherosclerosis Risk in Communities Study (ARIC)** is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

The **Health, Alcohol and Psychosocial factors In Eastern Europe studies (HAPIEE)** are supported by Wellcome Trust (064947/Z/01/Z and 081081/Z/06/Z); a grant from the National Institute on Aging (1R01 AG23522-01); and a grant from MacArthur Foundation.

The **Edinburgh Artery Study (EAS)** was funded by the British Heart Foundation.

The **Aspirin in Asymptomatic Atherosclerosis Trial (AAA)** was funded by the British Heart Foundation and the Chief Scientist Office, Scotland.

The **Edinburgh Type 2 Diabetes Study (ET2DS)** was funded by the MRC of the United Kingdom (project grant G0500877).

The **Second Manifestations of Arterial Disease study (SMART)** was financially supported by the Biobanking and Biomolecular Research Infrastructure BBMRI-NL, financed by the Dutch government (NWO 184.021.007).

The **Women's Genome Health Study (WGHS)** is supported by HL043851 and HL080467 from the National Heart, Lung, and Blood Institute and CA047988 from the National Cancer Institute, with collaborative scientific support and funding for genotyping provided by Amgen.

The Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin trial (JUPITER) and genetic analysis in the JUPITER population were funded by AstraZeneca.

The **Coronary Artery Risk Development in Young Adults (CARDIA)** Study is conducted and supported by the National Heart, Lung, and Blood Institute in collaboration with the University of Alabama at Birmingham (HHSN268201300025C & HHSN268201300026C), Northwestern University (HHSN268201300027C), University of Minnesota (HHSN268201300028C), Kaiser Foundation Research Institute (HHSN268201300029C), and Johns Hopkins University School of Medicine (HHSN268200900041C). CARDIA is also partially supported by the Intramural Research Program of the National Institute on Aging. This manuscript has been reviewed and approved by CARDIA for scientific content.

Supplementary tables and figures

Supplementary Table 1 - SNPs within 55kb of *HMGCR* on the Cardiochip platform and their associations with LDL-C in the Whitehall II study

					Association with LDL-C					
Selected for analysis	SNP	Chr.	Base pair position	No. of participants	LDL-C β-coefficient	95% Confidence interval	Regression model p-value	LDL-C R^2	F-statistic	MAF
✓	rs12916	5	74692295	4,666	0.12	0.08 to 0.17	6.0e-09	7.2e-03	33.96	0.40
	rs10038095	5	74673467	4,667	0.12	0.07 to 0.16	5.8e-08	6.3e-03	29.51	0.38
	rs12654264	5	74684359	4,665	0.12	0.07 to 0.16	8.3e-08	6.1e-03	28.83	0.38
	rs3846662	5	74686840	4,661	0.11	0.07 to 0.15	9.8e-08	6.1e-03	28.51	0.43
✓	rs17238484	5	74684252	4,658	0.12	0.07 to 0.17	1.0e-06	5.1e-03	23.96	0.23
	rs10474434	5	74680437	4,671	0.12	0.07 to 0.17	1.5e-06	4.9e-03	23.17	0.23
	rs3804231	5	74732535	4,672	0.14	0.07 to 0.20	1.5e-05	4.0e-03	18.75	0.13
	rs2303152	5	74677463	4,677	0.07	0.00 to 0.14	.038	9.2e-04	4.31	0.10
	rs17238540	5	74691254	4,673	-0.12	-0.25 to 0.02	.098	5.9e-04	2.74	0.02
	rs3761739	5	74667257	4,672	0.04	-0.02 to 0.10	.152	4.4e-04	2.05	0.15
	rs3761740	5	74667889	4,676	0.04	-0.03 to 0.11	.229	3.1e-04	1.45	0.09
	rs17238372	5	74672629	4,677	0.13	-0.08 to 0.34	.231	3.1e-04	1.44	0.01
	rs17648288	5	74732394	4,674	0.04	-0.03 to 0.11	.258	2.7e-04	1.28	0.09
	rs16872536	5	74714241	4,673	0.04	-0.03 to 0.11	.274	2.6e-04	1.20	0.08
	rs10474435	5	74693036	4,674	0.10	-0.09 to 0.29	.285	2.5e-04	1.15	0.01
	rs16872526	5	74711473	4,664	0.04	-0.03 to 0.11	.29	2.4e-04	1.12	0.08
	rs10038689	5	74718367	4,677	0.10	-0.09 to 0.29	.302	2.3e-04	1.07	0.01
	rs17244848	5	74679426	4,632	0.03	-0.04 to 0.11	.336	2.0e-04	0.92	0.09
	rs17244939	5	74666852	4,671	-0.07	-0.22 to 0.08	.369	1.7e-04	0.81	0.02
	rs4704209	5	74670229	4,670	0.04	-0.05 to 0.13	.42	1.4e-04	0.65	0.06
	rs3889900	5	74741803	4,678	-	-	-	-	-	<0.01
	rs17238477	5	74680879	4,678	-	-	-	-	-	<0.01
	rs17244715	5	74667984	4,678	-	-	-	-	-	<0.01
	rs17238519	5	74689134	4,671	-	-	-	-	-	<0.01
	rs17244946	5	74667566	4,676	-	-	-	-	-	<0.01
	rs17238435	5	74677582	4,678	-	-	-	-	-	<0.01
	rs17244708	5	74667603	4,675	-	-	-	-	-	<0.01
	rs17244890	5	74688485	4,671	-	-	-	-	-	<0.01
	rs17238337	5	74667055	4,675	-	-	-	-	-	<0.01
	rs16872532	5	74713739	4,678	-	-	-	-	-	<0.01
	rs17244722	5	74668837	4,676	-	-	-	-	-	<0.01
	rs17244960	5	74680009	4,676	-	-	-	-	-	<0.01
	rs17244883	5	74688082	4,676	-	-	-	-	-	<0.01
	rs17238449	5	74678164	4,678	-	-	-	-	-	<0.01
	rs17244757	5	74672460	4,676	-	-	-	-	-	<0.01

Selected for analysis	SNP	Chr.	Base pair position	No. of participants	Association with LDL-C: LDL-C β-coefficient	95% Confidence interval	Regression model p-value	LDL-C R^2	F-statistic	MAF
	rs17244729	5	74669229	4,676	-	-	-	-	-	<0.01
	rs17238554	5	74693408	4,674	-	-	-	-	-	<0.01
	rs17238575	5	74680466	4,671	-	-	-	-	-	<0.01

SNPs are ranked on p-value and β-coefficient; β-coefficients are reported on the log scale. For SNPs where β-coefficient is absent, no variation in genotype was observed in the Whitehall II study sample owing to very low minor allele frequency and consequently a regression model could not be fitted.

Abbreviations: Chr – chromosome; R^2 – proportion of variance explained; MAF – minor allele frequency

Supplementary Table 2 – Genotype information for studies contributing individual participant-level data

		rs17238484								rs12916							
Study	Genotyping platform	SNP call rate (%)	Proxy SNP used (LD, r^2)*	GG	GT	TT	MAF (%)	HWE χ^2	HWE p	SNP call rate (%)	Proxy SNP used (LD, r^2)*	TT	TC	CC	MAF (%)	HWE χ^2	HWE p
AAA	KASPAR		-	1,358	792	100	22.0	1.31	0.25	95.9	-	999	943	321	35.0	16.12	6.0×10^{-5}
ARIC	IBC Cardiochip		-	5,635	3,386	565	23.6	3.53	0.06	100.0	-	3,432	4,528	1,628	40.6	4.16	0.04
BRHS	KASPar	98.3	-	2,328	1,325	223	27.4	3.56	0.06	97.3	-	1,620	1,562	657	37.5	66.5	3.5×10^{-16}
BWHHS	IBC Cardiochip		-	2,010	1,245	185	23.5	0.19	0.67	99.9	-	1,218	1,668	557	40.4	0.12	0.72
CaPS	KASPar		-	837	498	70	22.7	0.14	0.71	98.5	-	563	627	216	37.7	3.55	0.06
CARDIA	IBC Cardiochip		-	836	527	80	23.8	0.07	0.80	100.0	-	472	736	235	41.8	3.38	0.07
CCHS	ABI TaqMan	90.5	-	5,532	3,385	488	23.2	1.03	0.31	-	-	-	-	-	-	-	-
CFS	IBC Cardiochip		-	79	48	8	23.7	0.04	0.84	100.0	-	45	58	32	45.2	2.38	0.12
CHS	IBC Cardiochip		-	2,344	1,381	225	23.2	1.31	0.25	100.0	-	1,413	1,911	628	40.1	0.18	0.67
EAS	ABI TaqMan		-	518	287	46	22.3	0.57	0.45	93.9	-	323	402	131	38.8	0.10	0.75
ELSA	KASPar		-	3,352	1,852	277	22.0	1.04	0.31	-	-	-	-	-	-	-	-
Ely	Metabochip		rs3761742 (1.00)	989	544	68	21.2	0.39	0.53	-	-	-	-	-	-	-	-
EPIC-NL	IBC Cardiochip		-	3,009	1,837	251	22.6	1.02	0.31	-	-	-	-	-	-	-	-
ET2DS[1]	KASPar		-	690	311	52	19.71	4.68	0.03	97.8	-	501	390	155	33.5	27.7	0.14
MRC Fenland	Metabochip		rs3761742 (1.00)	1,953	1,061	172	22.1	3.11	0.08	-	-	-	-	-	-	-	-
FHS	IBC Cardiochip		-	807	456	69	22.3	0.19	0.66	100.0	-	491	647	198	39.0	0.41	0.52
HAPIEE-CZ	KASPar		-	3,996	2,384	341	22.8	0.26	0.55	-	-	-	-	-	-	-	-
HAPIEE-LT	KASPar		-	4,049	2,424	407	23.5	3.04	0.08	-	-	-	-	-	-	-	-
HAPIEE-PL	KASPar		-	3,456	1,988	282	22.3	0.03	0.86	-	-	-	-	-	-	-	-
HAPIEE-RU	KASPar		-	4,264	2,423	377	22.5	1.82	0.18	-	-	-	-	-	-	-	-
IMPROVE	Metabochip		-	1,937	1,310	215	25.1	0.11	0.74	100.0	-	1,081	1,748	636	43.6	2.32	0.13
InterAct	Illumina 660W/ Sequenom/ ABI TaqMan	100.0	rs3761742 (1.00)	11,843	6,209	838	20.9	0.45	0.50	100.0	-	3,122	9,138	6,634	40.7	0.07	0.79
JUPITER	Illumina Omni 1M Quad	100.0	-	5,363	2,983	403	21.7	0.21	0.65	100.0	-	3,249	4,157	1,340	39.1	0.03	0.86
MESA	IBC Cardiochip	100.0	-	1,333	829	135	23.9	0.17	0.68	100.0	-	792	1,138	368	40.77	1.47	0.22
MRC NSHD	KASPar	95.8	-	1,432	830	140	23.1	1.82	0.18	100.0	-	1,048	967	387	36.2	39.9	2.7×10^{-10}
NPHS-II	ABI TaqMan		-	1,646	923	131	21.9	0.01	0.91	100.0	-	960	1,286	456	41.1	0.10	0.75
PROSPER	Illumina 660K	>97.5	rs3761742 (1.00)	3,307	1,710	227	20.6	0.10	0.75	-	-	-	-	-	-	-	-
SMART	KASPar	99.4	-	4,915	3,007	449	23.3	0.15	0.70	-	-	-	-	-	-	-	-

TPT	ABI TaqMan		-	2,349	1,403	197	22.8	0.45	0.50	-	-	-	-	-	-	-	-	-
UCP	IBC Cardiochip		-	973	556	102	23.3	3.48	0.06	-	-	-	-	-	-	-	-	-
WGHS	Illumin HumanHap Duo+	100.0	rs1051795 (0.95)	14,559	7,708	1,020	20.9	2.84×10^{-5}	0.99	99.99	-	8,260	11,213	3,818	40.5	0.02	0.90
WHI	IBC Cardiochip		-	3,442	1,779	316	21.8	17.85	2×10^{-5}	100.0	-	2,042	2,692	981	40.7	3.37	0.07
Whitehall II	IBC Cardiochip		-	3,020	1,765	252	22.5	0.08	0.78	99.8	-	1,794	2,421	832	40.5	0.10	0.75

Abbreviations: HWE - Hardy-Weinberg equilibrium; MAF - minor allele frequency; LD – linkage disequilibrium

*Linkage disequilibrium data derived from 1000 Genomes Pilot 1

[1] Minor allele frequency in diabetes cases

Supplementary Table 3 - Collaborating genetic study characteristics

Study name	Study design	Geographical location	Sampling frame	Participants included	Baseline year(s)	% Male	Age (Mean ,SD)	Fasting blood samples?
AAA	RCT	UK	General practice	2,250	1998-2001	28.5	61.9 (6.6)	No
ARIC	Cohort	USA	Community	9,586	1986	53.5	54.3 (5.7)	
BRHS	Cohort	UK	General practice	3,876	1978-1980	100	68.7 (5.5)	
BWHHS	Cohort	UK	General practice	3,440	1999-2001	0	69.3 (5.5)	
CaPS	Cohort	UK	Community	1,405	1984-1988	100	56.7 (4.4)	
CARDIA	Cohort	USA	Community	1,443	1985-1986	53.4	25.6 (3.4)	
CCHS	Cohort	Denmark	Population	9,405	1990-2003	44.2	56 .7 (16)	No
CFS	Cohort	USA	Community (family)	135	1995	41.5	53.9 (14.7)	
CHS	Cohort	USA	Community	3,950	1989	45.0	70.4 (5.3)	Yes
deCODE	Cohort	Iceland	Population, clinic	24,650	-	-	-	-
DGDG	C-C	France	Clinic, population	1,376	-	-	-	-
DGI	C-C	Finland, Sweden	Population	2,097	-	47.7	-	-
EAS	Cohort	UK	General practices	851	1987	50.8	64.3 (5.6)	Yes
ELSA	Cohort	UK	HSE respondents	5,481	1998, 1999, 2001	45.6	61.0 (9.6)	
Ely	Cohort	UK	General practice	1,601	1990-1992	47.0	61.1 (9.2)	
EPIC-NL	Nested C-Ct	Netherlands	Population	5,194	1993-1997	21.9	54.1 (10.1)	No
ET2DS	Cohort	UK	Community (cases)	1,053	2006-2007	50.3	67.9 (4.2)	Yes
EUROSPAN	C-C	Europe	Population	3,978	-	-	-	-
MRC Fenland	Cohort	UK	General practices	3,186	2004 (ongoing)	46.0	46.9 (7.1)	
FHS	Cohort	USA	Community	1,332	1948/1971/2002	52.7	45.7 (10.1)	
FUSION	C-C	Finland	Community	2,335	-	-	-	-
HAPIEE-Czech Republic	Cohort	Czech Republic	Population	6,721	2002-2005	46.6	58.3 (7.1)	
HAPIEE-Lithuania	Cohort	Lithuania	Population	6,880	2006-2008	45.4	61.0 (7.6)	
HAPIEE-Poland	Cohort	Poland	Population	5,726	2002-2005	45.4	57.6 (6.9)	
HAPIEE-Russia	Cohort	Russia	Population	7,064	2002-2005	45.6	58.9 (7.1)	
IMPROVE	Cohort	Europe	Clinic	3,462	2004	48.3	64.2 (5.4)	Yes
InterAct	Cohort	Europe	Population	18,894	1991	49.7	55.6	
JUPITER	RCT	America, Europe, Africa	Clinic	8,749	2003-2005	67.8	66.1	Yes
KORAgen	Nested C-C	Germany	Population	1,871	-	-	-	-
MESA	Cohort	USA	Population	2,297	2000-2002	47.8	62.7 (10.2)	Yes
MRC NSHD	Birth cohort	UK	Birth register	2,402	1946	49.8	53.0 (0.0)	No
NPHS-II	Cohort	UK	General practices	2,700	1989-1994	100	56.6 (3.4)	

PREVEND	Prospective	Netherlands	Population	3,649	1997	51		
PROSPER	RCT	Europe	Clinic	5,804	1997-1999	48.3	75.3 (3.3)	Yes
Rotterdam	Nested C-C	Netherlands	Population	5,939	-	-	-	-
SMART	Cohort	Netherlands	Clinic	8,371	1996-2011	65.1	56.5 (12.4)	Yes
TPT	RCT	UK	General practice	3,949	1984-1989	100	57.4 (6.7)	
UCP	Nested C-C	Netherlands	Clinic	1,631	2007	74.5	62.8 (9.6)	
WGHS	RCT	North America	Professional	23,294	1992-1994	0	54.7	Yes
WHI	Nested C-C	USA	Community	5,537	1994-1998	0	68.1 (6.6)	Yes
Whitehall II	Cohort	UK	Workplace	5,037	1985-1988	73.6	49.2 (6.0)	
WTCCC	C-C	UK	Population	4,862	-	-	-	-

Abbreviations: C-C : case-control; C-Ct: case-cohort; HSE - Health Survey for England
- : Data unavailable

Supplementary Table 4 - Data availability for biomarkers and events in studies contributing to the primary genetic analysis

	Biomarkers										Clinical events (prevalent & incident)
	Total chol.	LDL-C	Non-HDL-C	Weight	BMI	Waist circ.	Hip circ.	Waist:hip ratio	Plasma glucose	Plasma insulin	Diabetes mellitus
AAA	2,242	-	-	-	-	-	-	-	-	-	48
ARIC	9,567	9,412	9,567	9,578	9,578	9,576	9,576	9,575	7,740	9,578	1,184
BRHS	3,855	3,746	3,740	3,867	3,873	3,850	3,859	3,848	3,856	-	294
BWHHS	3,257	3,177	3,253	3,249	3,243	3,233	3,236	3,239	3,252	3,274	439
CaPS	1,338	1,299	-	1,376	1,360	1,354	1,353	1,352	1,339	696	117
CARDIA	1,436	1,436	1,436	1,438	1,437	1,436	1,437	1,436	1,331	1,321	99
CCHS	9,394	9,379	9,388	9,385	9,382	9,364	9,365	9,365	-	-	843
CFS	39	38	39	134	134	45	45	45	39	39	17
CHS	3,941	-	3,936	-	3,938	-	-	-	3,936	-	-
deCODE	-	-	-	-	-	-	-	-	-	-	1,456
DGDG	-	-	-	-	-	-	-	-	-	-	679
DGI	-	-	-	-	-	-	-	-	-	-	1,022
EAS	850	846	846	851	851	-	-	-	849	-	30
ELSA	5,449	5,304	5,447	5,323	5,235	5,362	5,351	5,346	-	-	-
Ely	1,600	1,597	-	1,599	1,598	1,588	1,587	1,586	1,598	1,596	89
EPIC-NL	4,337	4,261	3,060	5,186	5,185	5,181	5,181	5,179	3,455	-	1,110
ET2DS	1,025	-	1,025	1,032	1,032	1,028	1,028	1,028	1,021	-	1,053
EUROSPAN	-	-	-	-	-	-	-	-	-	-	268
MRC Fenland	3,180	3,155	-	3,184	3,184	3,183	3,183	3,182	3,178	3,148	38
FHS	845	844	844	858	316	785	784	783	825	749	161
FUSION	-	-	-	-	-	-	-	-	-	-	1,161
HAPIEE-CZ	6,568	6,301	6,543	6,719	6,716	-	-	6,708	914	-	355
HAPIEE-LT	6,838	6,655	6,720	6,874	6,874	-	-	6,873	6,708	-	389
HAPIEE-PL	5,720	5,647	5,720	5,720	5,716	-	-	5,719	5,713	-	423
HAPIEE-RU	7,062	7,058	7,060	7,064	7,063	-	-	7,061	1,042	-	254
IMPROVE	3,456	3,395	-	3,461	3,461	3,462	3,462	3,462	3,455	-	858
InterAct	-	-	-	-	-	-	-	-	-	-	8,247
JUPITER	-	-	-	8,729	8,723	-	-	-	-	-	279
KORAgen	-	-	-	-	-	-	-	-	-	-	433
MESA	2,292	2,260	2,290	2,297	2,297	2,297	2,297	2,297	2,292	-	220

MRC NSHD	2,236	2,072	2,082	2,382	2,379	2,961	2,388	2,957	-	-	-
NPHS-II	2,680	1,703	1,789	2,698	2,697	-	-	-	-	-	221
PREVEND	3,626	3,496	3,554	3,623	3,622	3,622	3,622	3,622	3,465	3,402	140
PROSPER	5,244	5,242	-	5,243	5,242	-	-	-	5,060	5,032	269
Rotterdam	-	-	-	-	-	-	-	-	-	-	1,178
SMART	7,897	7,448	7,879	7,925	7,924	6,715	6,713	6,716	7,879	4,373	-
TPT	3,912	-	-	3,932	3,932	-	-	-	-	-	-
UCP	-	-	-	1,592	1,487	-	-	-	-	-	-
WGHS	-	-	-	23,043	23,291	-	-	-	-	-	1,444
WHI	2,920	1,472	2,861	-	5,489	-	-	-	-	-	272
Whitehall II	4,739	4,658	4,724	4,752	4,746	4,693	4,692	4,688	4,543	4,245	334
WTCCC	-	-	-	-	-	-	-	-	-	-	1,924
Total	117,545	101,901	84,415	143,114	152,005	69,735	69,159	96,067	73,490	37,453	27,348

Abbreviations: TC - total cholesterol; HDL-C - high density lipoprotein cholesterol; LDL-C - low density lipoprotein cholesterol

Supplementary Table 5 - Clinical outcome definitions

	Diabetes (all aetiology, prevalent & incident)		
	Self report	Medical records	Clinical/lab. measures
AAA	●	●	
ARIC		●	●
BRHS		●	
BWHHS	●	●	
CaPS	●	●	●
CARDIA	●	●	●
CCHS		●	
CFS		●	
CHS		●	●
deCODE	-	-	-
EAS	●		
ELSA	n/a	n/a	n/a
Ely	n/a	n/a	n/a
EPIC-NL	●	●	
ET2DS		●	●
EUROSPAN	-	-	-
MRC Fenland	n/a	n/a	n/a
FHS		●	
FUSION	-	-	-
HAPIEE-CZ	●		
HAPIEE-LT	●		
HAPIEE-PL	●		
HAPIEE-RU	●		
IMPROVE	n/a	n/a	n/a
InterAct	●	●	●
JUPITER		●	●
KORAgen			
MESA	●		●
MRC NSHD	n/a	n/a	n/a
NPHS-II	●	●	
PROSPER		●	●
Rotterdam	-	-	-
SMART	n/a	n/a	n/a
TPT		●	
WGHS	●	●	-
WHI	●		
Whitehall II	●	●	●
WTCCC	-	-	-

Abbreviations: n/a - T2D data not included in analysis; - - T2D definition information unavailable

Supplementary Table 6 – Associations of HMGCR rs17238484 with biomarkers and T2D risk in pre-specified subgroups.

P-values are for meta-regression between subgroup strata.

		Tertiles of non-HDL-C		
Biomarker	Low	Mid	High	
		Estimate (95% CI)		*Meta-regression p-value*
LDL-C (mmol/L)	-0.013 (-0.022 to -0.004)	-0.013 (-0.021 to -0.006)	-0.034 (-0.047 to -0.020)	0.055
Plasma glucose (%)	1.658 (-0.386 to 3.744)	2.364 (0.281 to 4.490)	1.923 (-0.201 to 4.092)	0.855
Plasma insulin (%)	4.628 (1.451 to 7.905)	3.761 (0.625 to 6.995)	2.884 (-0.089 to 5.946)	0.445
BMI (kg/m^2)	0.035 (0.018 to 0.052)	0.035 (0.018 to 0.052)	0.029 (0.012 to 0.045)	0.596
Body weight (kg)	0.032 (0.013 to 0.050)	0.042 (0.024 to 0.060)	0.020 (0.002 to 0.037)	0.354

		Lipid-lowering drug use	
Biomarker	Users	Non-users	
	Estimate (95% CI)		*Meta-regression p-value*
LDL-C (mmol/L)	-0.044 (-0.072 to -0.016)	-0.075 (-0.088 to -0.062)	0.122
Plasma glucose (%)	1.531 (-1.460 to 4.613)	0.773 (-0.909 to 2.483)	0.671
Plasma insulin (%)	5.265 (0.562 to 10.189)	2.987 (0.424 to 5.615)	0.423
BMI (kg/m^2)	0.039 (0.013 to 0.065)	0.025 (0.013 to 0.038)	0.349
Body weight (kg)	0.035 (0.008 to 0.062)	0.023 (0.010 to 0.036)	0.448
T2D (odds ratio)	0.981 (0.890 to 1.082)	1.038 (0.985 to 1.094)	0.316

		BMI classes		
Biomarker	Normal (<25)	Overweight (25-30)	Obese (>30)	
		Estimate (95% CI)		*Meta-regression p-value*
LDL-C (mmol/L)	-0.073 (-0.090 to -0.056)	-0.059 (-0.076 to -0.043)	-0.070 (-0.092 to -0.048)	0.247
Plasma glucose (%)	1.631 (0.033 to 3.254)	0.199 (-1.565 to 1.994)	1.126 (-1.935 to 4.282)	0.249
Plasma insulin (%)	2.558 (0.136 to 5.039)	-0.130 (-2.622 to 2.426)	2.488 (-1.422 to 6.552)	0.150
BMI (kg/m^2)	-	-	-	-
Body weight (kg)	0.007 (-0.003 to 0.017)	0.011 (0.002 to 0.021)	0.007 (-0.012 to 0.026)	0.538

		Sex	
Biomarker	Male	Female	
	Estimate (95% CI)		*Meta-regression p-value*
LDL-C (mmol/L)	-0.065 (-0.079 to -0.051)	-0.061 (-0.076 to -0.046)	0.909
Plasma glucose (%)	0.881 (-0.823 to 2.613)	2.097 (0.326 to 3.900)	0.346
Plasma insulin (%)	2.054 (-0.495 to 4.668)	2.868 (0.350 to 5.450)	0.665
BMI (kg/m^2)	0.019 (0.007 to 0.031)	0.032 (0.017 to 0.046)	0.197
Body weight (kg)	0.022 (0.009 to 0.034)	0.026 (0.012 to 0.039)	0.650

		T2D cases/controls	
Biomarker	Cases	Controls	
	Estimate (95% CI)		*Meta-regression p-value*
LDL-C (mmol/L)	-0.051 (-0.085 to -0.016)	-0.065 (-0.077 to -0.053)	0.570
Plasma glucose (%)	-2.704 (-7.494 to 2.335)	0.698 (-0.306 to 1.712)	0.199
Plasma insulin (%)	0.245 (-6.159 to 7.086)	1.945 (0.012 to 3.915)	0.637
BMI (kg/m^2)	0.022 (-0.011 to 0.055)	0.024 (0.013 to 0.035)	0.906
Body weight (kg)	-0.003 (-0.035 to 0.030)	0.024 (0.012 to 0.035)	0.136

Supplementary Table 7 – Associations of *HMGCR* rs12916 with plasma lipids, T2D-related biomarkers and T2D risk – fixed effects meta-analysis

		rs12916			Heterogeneity	
Biomarker	**Total n (studies)**	**Per-allele effect**	**95% CI**	**Effect p-value**	**I^2**	**Heterogeneity p-value**
Total cholesterol (mmol/L)	40,712 (15)	-0.077	-0.091 to -0.062	4.74×10^{-25}	27.5	0.15
LDL-C (mmol/L)	32,674 (13)	-0.083	-0.098 to -0.068	4.40×10^{-28}	33.3	0.12
Non-HDL-C (mmol/L)	39,591 (15)	-0.074	-0.089 to -0.059	9.50×10^{-22}	28.2	0.15
Plasma insulin (%)	19,918 (7)	0.659	-0.631 to 1.967	0.318	44.4	010
Plasma glucose (%)	30,641 (12)	0.125	-0.135 to 0.387	0.35	0.0	0.54
Body weight (kg)	65,871 (15)	0.203	0.043 to 0.362	0.01	0.0	0.84
BMI (kg/m^2)	75,140 (17)	0.083	0.031 to 0.135	1.62×10^{-3}	0.0	0.90
Height (m)	39,771, (14)	0.0001	-0.001 to 0.001	0.77	0	0.67
Waist circumference (cm)	30,313 (11)	0.303	0.102 to 0.504	0.003	0	0.55
Hip circumference (cm)	30,313 (11)	0.193	0.050 to 0.337	8.07×10^{-3}	28.6	0.17
Outcome	**No. of cases/controls (studies)**	**Per-allele odds ratio**	**95% CI**	**Effect p-value**	**I^2**	**Heterogeneity p-value**
T2D	14,976/73,574 (16)	1.06	1.03 to 1.09	9.58×10^{-5}	36.7	0.07

Supplementary Table 8 - Associations of rs17238484 and rs12916 with T2D-related biomarkers and T2D risk - random effects meta-analyses

rs17238484	Biomarker	No. of individuals (studies)	Per-allele beta coefficient	95% confidence interval	Effect p-value
	LDL-C (mmol/L)	101,919 (26)	-0.065	-0.077 to -0.052	6.20×10^{-25}
	Total cholesterol (mmol/L)	117,545 (30)	-0.067	-0.079 to -0.055	4.60×10^{-30}
	Non-HDL-C (mmol/L)	103,375 (27)	-0.067	-0.081 to -0.053	1.20×10^{-21}
	Body weight (kg)	143,113 (30)	0.301	0.175 to 0.428	3.20×10^{-6}
	BMI (kg/m^2)	152,004 (32)	0.105	0.066 to 0.145	1.80×10^{-7}
	Height (m)	77,291 (23)	0.001	-0.000 to 0.002	0.23
	Waist circumference (cm)	69,163 (19)	0.315	0.158 to 0.472	8.30×10^{-5}
	Hip circumference (cm)	69,159 (19)	0.211	0.101 to 0.321	1.70×10^{-4}
	Plasma glucose (%)	73,490 (23)	0.230	0.020 to 0.441	0.03
	Plasma insulin (%)	37,453 (12)	0.016	0.005 to 0.027	3.60×10^{-3}
	Outcome	**No. of cases/controls (studies)**	**Odds ratio**	**95% confidence interval**	**p-value**
	T2DM	26,236/164,021 (35)	1.02	1.00 to 1.04	0.10

rs12916	Biomarker	No. of individuals (studies)	Per-allele beta coefficient	95% confidence interval	Effect p-value
	LDL-C (mmol/L)	32,674 (13)	-0.081	-0.101 to -0.062	5.50×10^{-16}
	Total cholesterol (mmol/L)	40,712 (15)	-0.075	-0.093 to -0.057	7.20×10^{-16}
	Non-HDL-C (mmol/L)	39,591 (15)	-0.072	-0.091 to -0.053	8.10×10^{-14}
	Body weight (kg)	65,871 (15)	0.203	0.043 to 0.362	0.01
	BMI (kg/m^2)	75,140 (17)	0.083	0.031 to 0.135	1.60×10^{-3}
	Waist circumference (cm)	39,771 (14)	0.000	-0.001 to 0.001	0.77
	Hip circumference (cm)	30,313 (11)	0.303	0.102 to 0.504	3.10×10^{-3}
	Plasma glucose (%)	30,313 (11)	0.195	0.008 to 0.382	0.04
	Plasma insulin (%)	30,641 (12)	0.130	-0.140 to 0.391	0.35
	Height (m)	19,918 (7)	0.004	-0.016 to 0.023	0.72
	Outcome	**No. of cases/controls (studies)**	**Odds ratio**	**95% confidence interval**	**p-value**
	T2DM	14,976/73,574 (16)	1.07	1.02 to 1.12	1.10×10^{-3}

Supplementary Table 9 – Effect of statin therapy on body weight and new-onset T2D risk– fixed effects meta-analysis

Trials	Weight change (kg)	95 % CI
Placebo- or standard care- controlled	0.33	0.26 to 0.41
Intensive vs moderate dose	-0.08	-0.16 to 0.01
All trials combined	0.16	0.11 to 0.22

Trials	T2D Odds Ratio	95 % CI
Placebo- or standard care- controlled	1.12	1.05 to 1.19
Intensive vs moderate dose	1.12	1.04 to 1.22
All trials combined	1.12	1.07 to 1.17

Supplementary Table 10 – Associations of *HMGCR* rs17238484 and rs12916 with T2D risk in 35 studies contributing to this analysis, and in combination with data from GWA study data from the DIAGRAM consortium

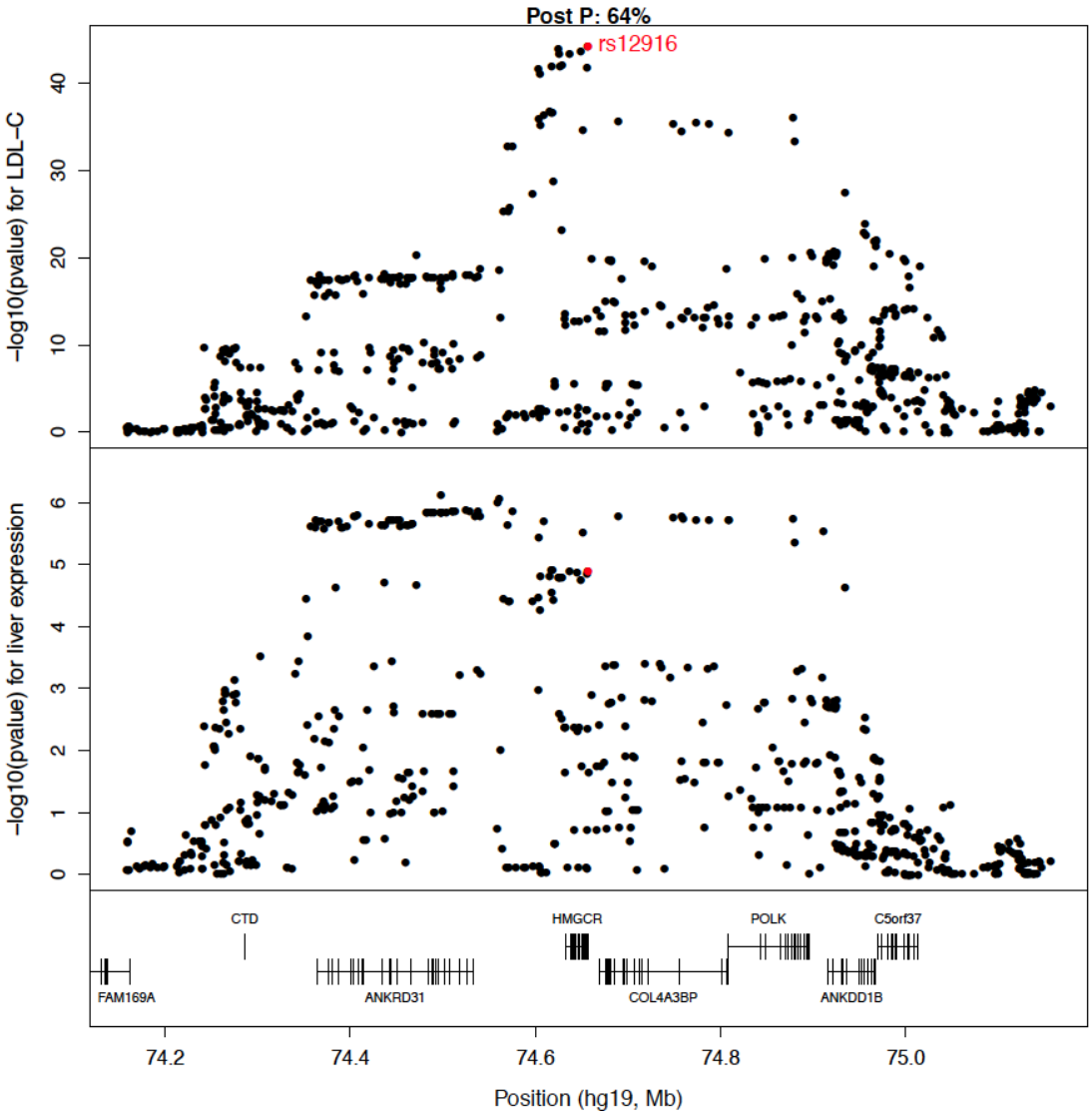
	rs17238484-G *OR (95% CI)*	*Cases/controls or total n (studies)*	**rs12916-T** *OR (95% CI)*	*Cases/controls or total n (studies)*
Risk of T2D: original studies	1.02 (1.00 to 1.05)	26,236/164,021 (35)	1.06 (1.03 to 1.09)	14,976/73,574 (16)
Risk of T2D: Morris GWAS+Metabochip analysis	1.03 (0.99 to 1.06)	51,943 (19)*	1.01 (0.99 to 1.04)	76,774 (19)*
Risk of T2D: original studies + unique Morris data	1.03 (1.01 to 1.06)	181,111 (38)*	1.02 (1.00 to 1.04)	151,329 (38)*

Estimates shown are derived from i) 35 observational population studies contributing to the current analysis; ii) Morris et al. GWAS+Metabochip studies consortium; iii) meta-analysis of original studies and studies contributing to the current analysis combined with the remaining studies unique to the Morris et al. *consortium.*

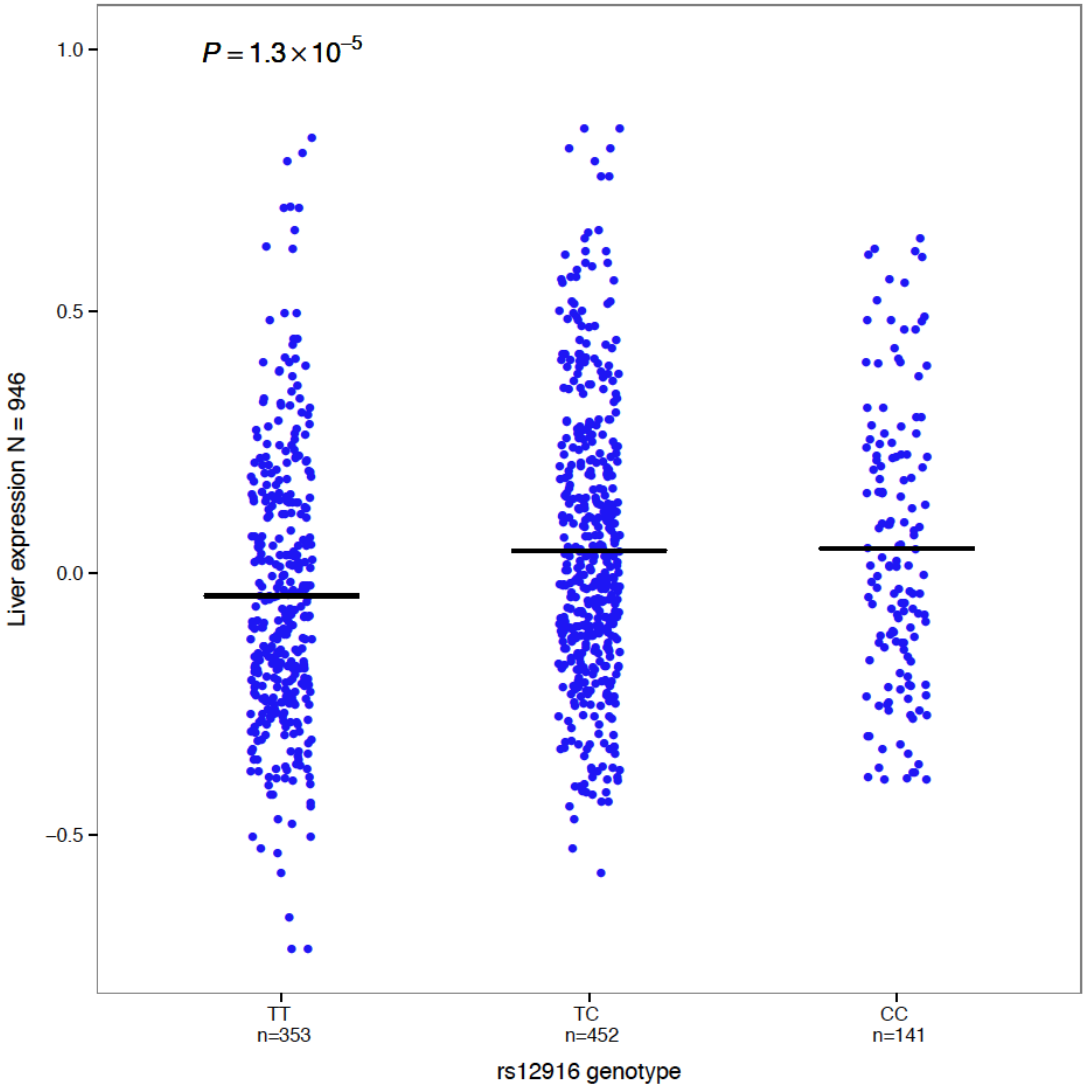
Effective sample size as reported by the DIAGRAM consortium at [*http://diagram-consortium.org/downloads.htm*](http://diagram-consortium.org/downloads.htm)
**Numbers of cases and controls unavailable from Morris* et al*. data*

Supplementary Figure 1 - Associations of SNPs at the *HMGCR* locus with LDL-C concentration and liver gene expression

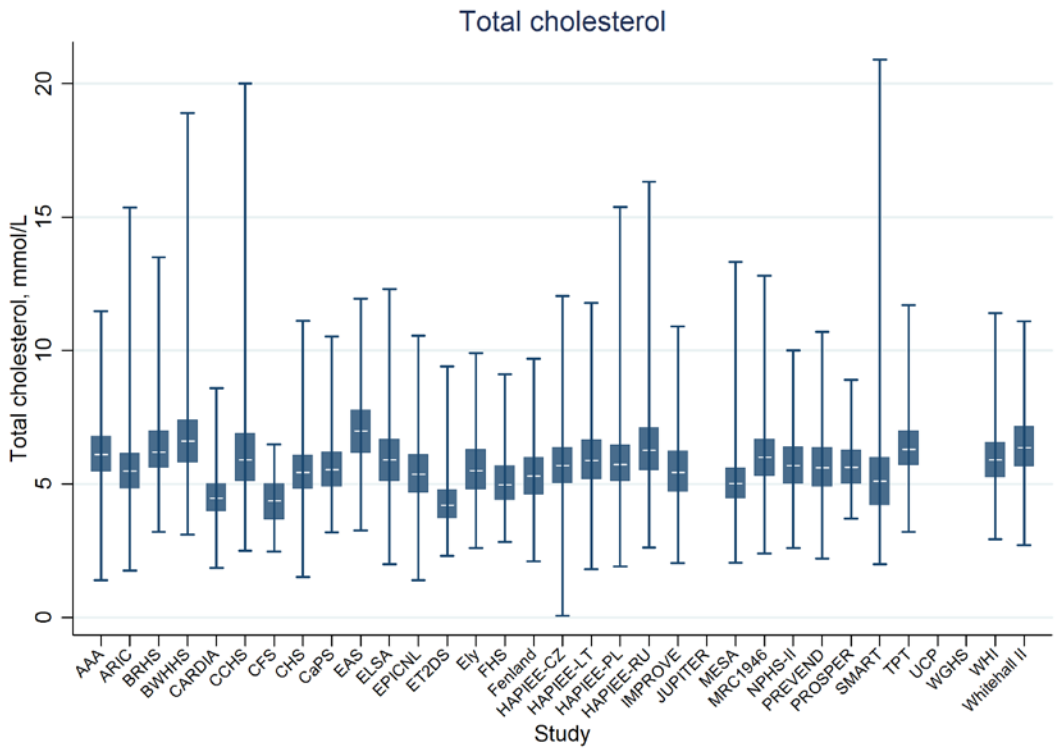
The x-axis shows the physical position on chromosome 5 (Mb). Each dot represents one variant (SNP or indel - imputed or directly typed). Upper panel graph y-axis uses p-values from a published meta-analysis of LDL-C levels in >100,000 individuals[50]. Lower panel y-axis shows the -$\log_{10}$p-value for association with *HMGCR* gene expression in liver. The posterior probability ("Post P") that LDL-C co-localises with *HMGCR* liver expression is shown on top of the graph. There is a high probability that *HMGCR* expression is mediating the LDL-C signal.

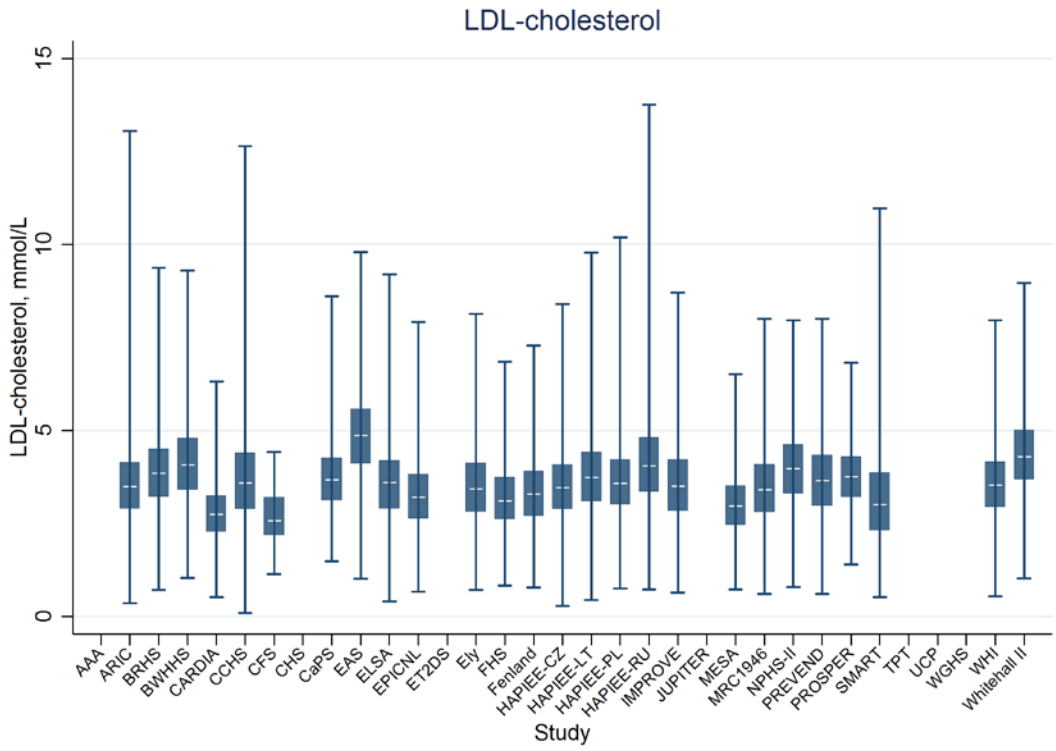


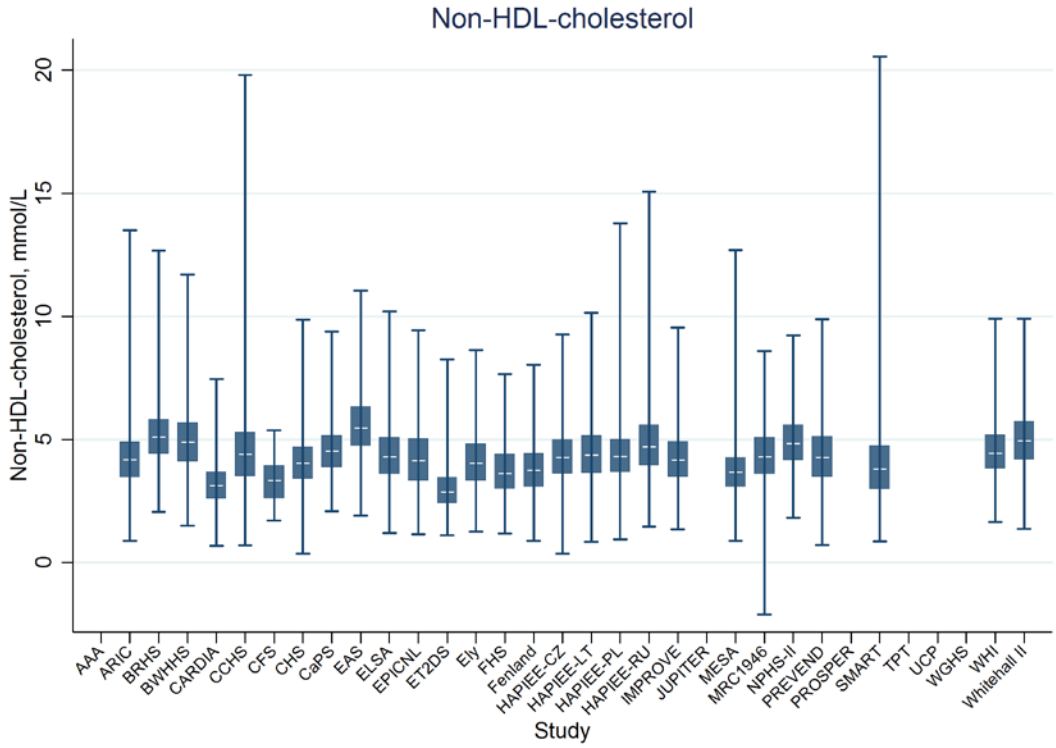
Supplementary Figure 2 - Effect of HMGCR rs12916 genotype on liver expression of HMGCR
946 individuals had expression measurements and genotypes that were directly typed or imputed (uncertain genotypes were removed). The allele T of the SNP rs12916 is associated with a decreased effect of HMGCR expression in liver (1-d.f. $p=1.3x10^{-5}$).

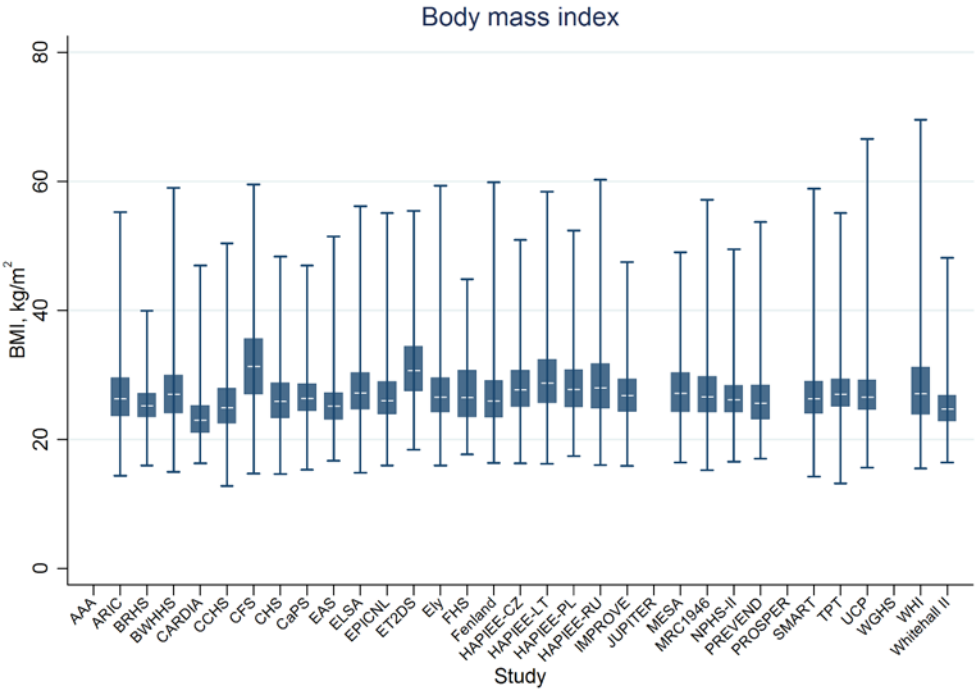


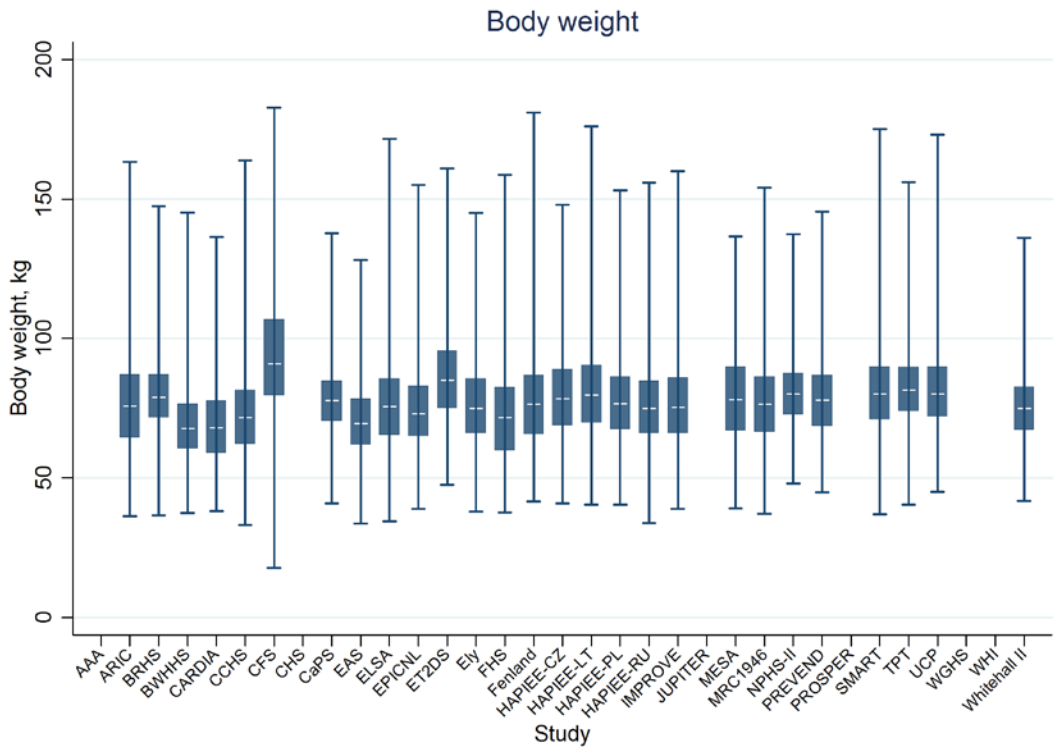
Supplementary Figure 3 - Genetic analysis biomarker data - box and whisker plots

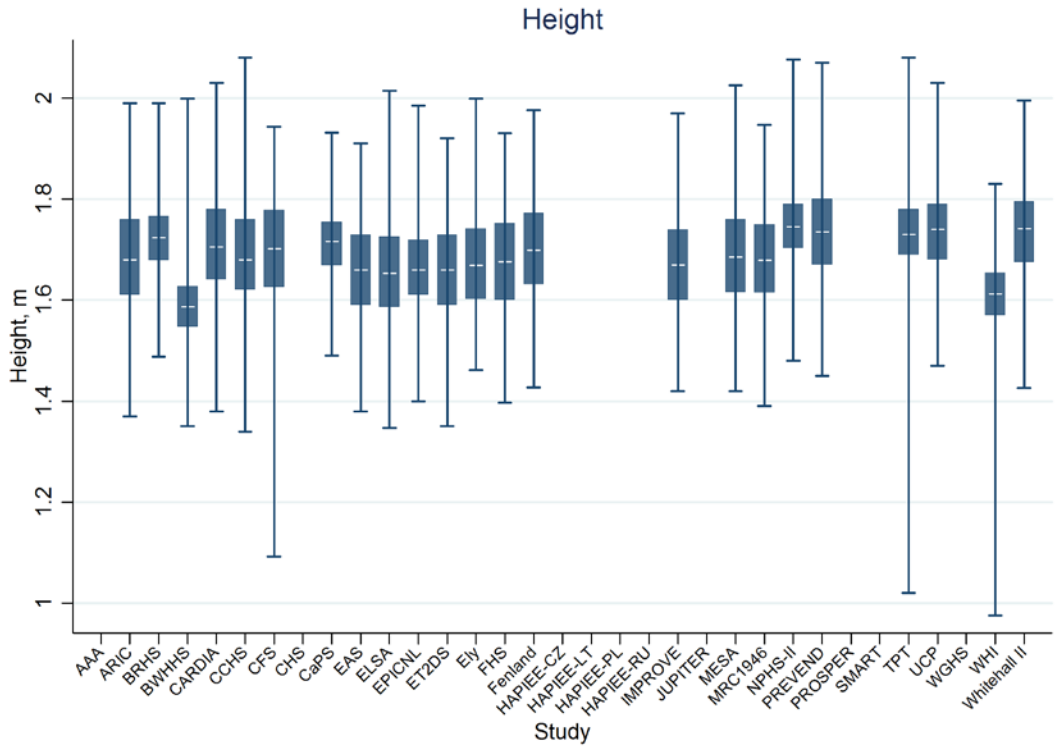


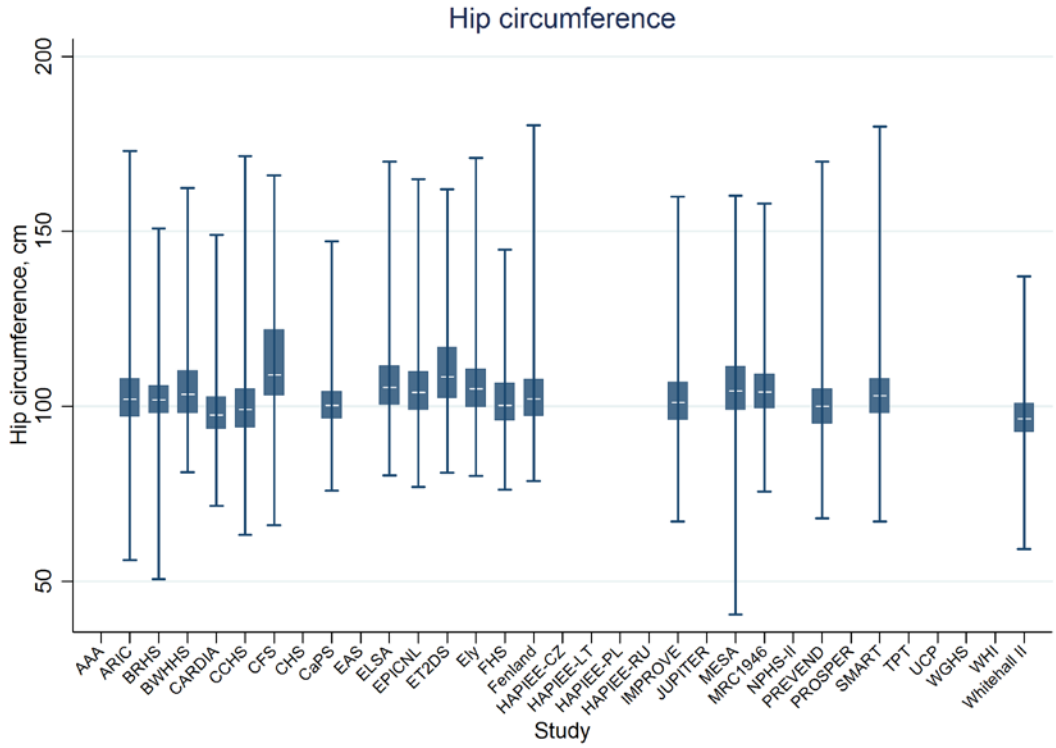


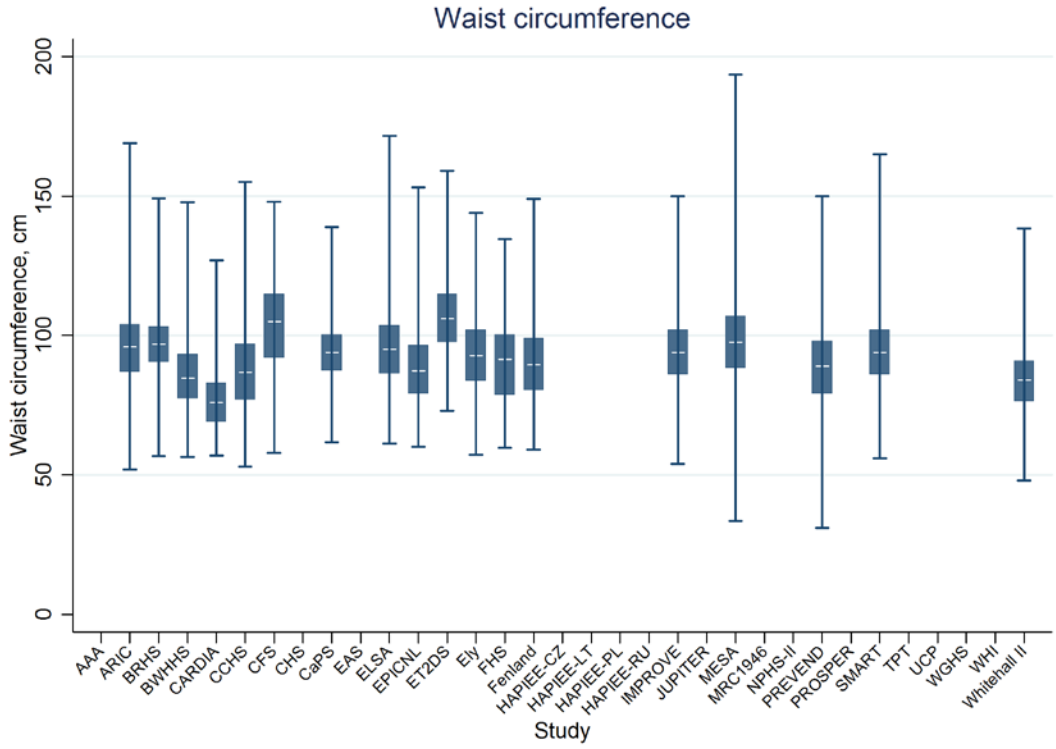

Non-HDL-cholesterol
Non-HDL-cholesterol, mmol/L
0
5
10
15
20
AAA
ARIC
BRHS
BWHHS
CARDIA
CCHS
CFS
CHS
CaPS
EAS
ELSA
EPICNL
ET2DS
Ely
FHS
Fenland
HAPIEE-CZ
HAPIEE-LT
HAPIEE-PL
HAPIEE-RU
IMPROVE
JUPITER
MESA
MRC1946
NPHS-II
PREVEND
PROSPER
SMART
TPT
UCP
WGHS
WHI
Whitehall II
Study

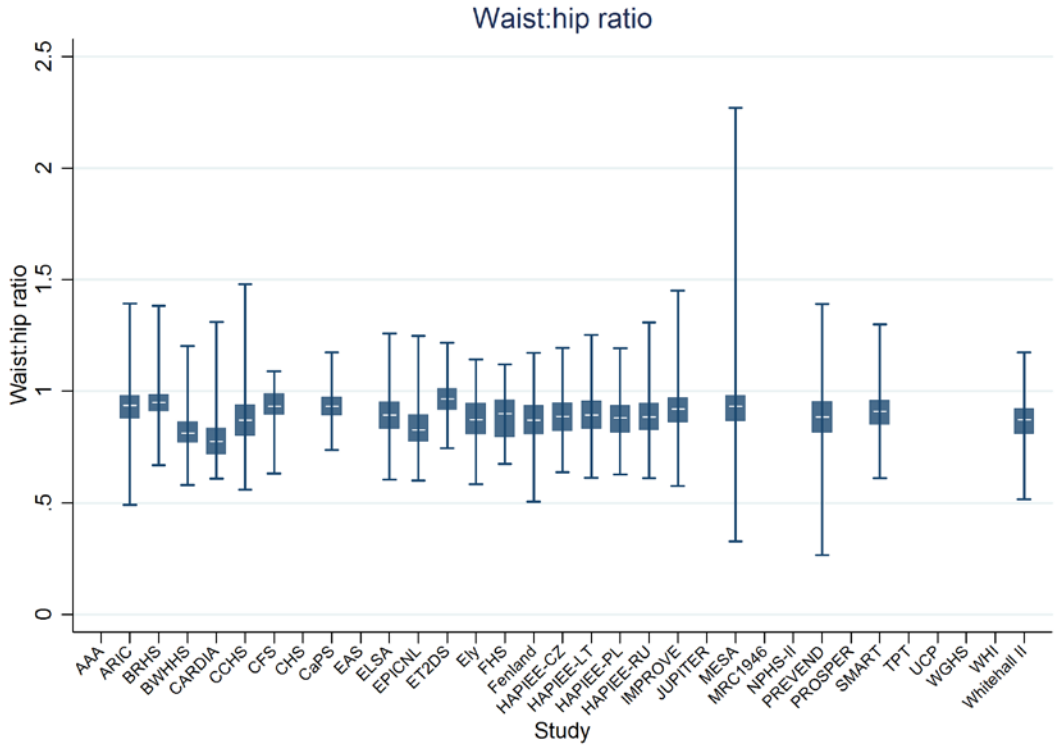

Body mass index
BMI, kg/m^2
0
20
40
60
80
AAA
ARIC
BRHS
BWHHS
CARDIA
CCHS
CFS
CHS
CaPS
EAS
ELSA
EPICNL
ET2DS
Ely
FHS
Fenland
HAPIEE-CZ
HAPIEE-LT
HAPIEE-PL
HAPIEE-RU
IMPROVE
JUPITER
MESA
MRC1946
NPHS-II
PREVEND
PROSPER
SMART
TPT
UCP
WGHS
WHI
Whitehall II
Study

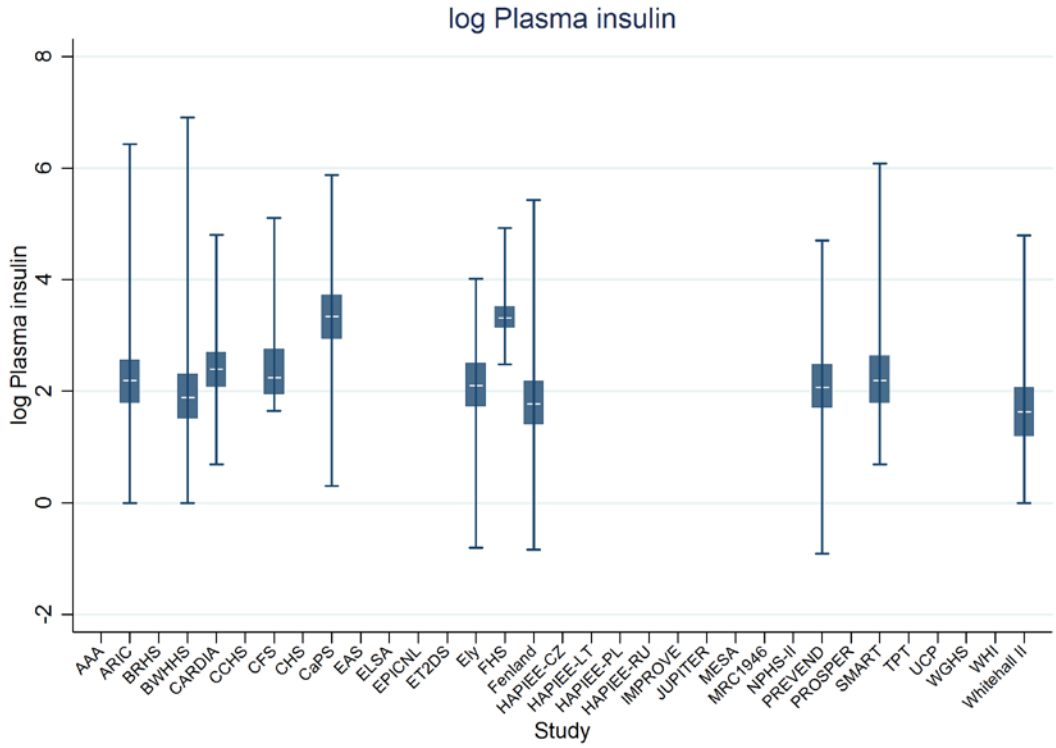

Body weight
Body weight, kg
0
50
100
150
200
AAA
ARIC
BRHS
BWHHS
CARDIA
CCHS
CFS
CHS
CaPS
EAS
ELSA
EPICNL
ET2DS
Ely
FHS
Fenland
HAPIEE-CZ
HAPIEE-LT
HAPIEE-PL
HAPIEE-RU
IMPROVE
JUPITER
MESA
MRC1946
NPHS-II
PREVEND
PROSPER
SMART
TPT
UCP
WGHS
WHI
Whitehall II
Study

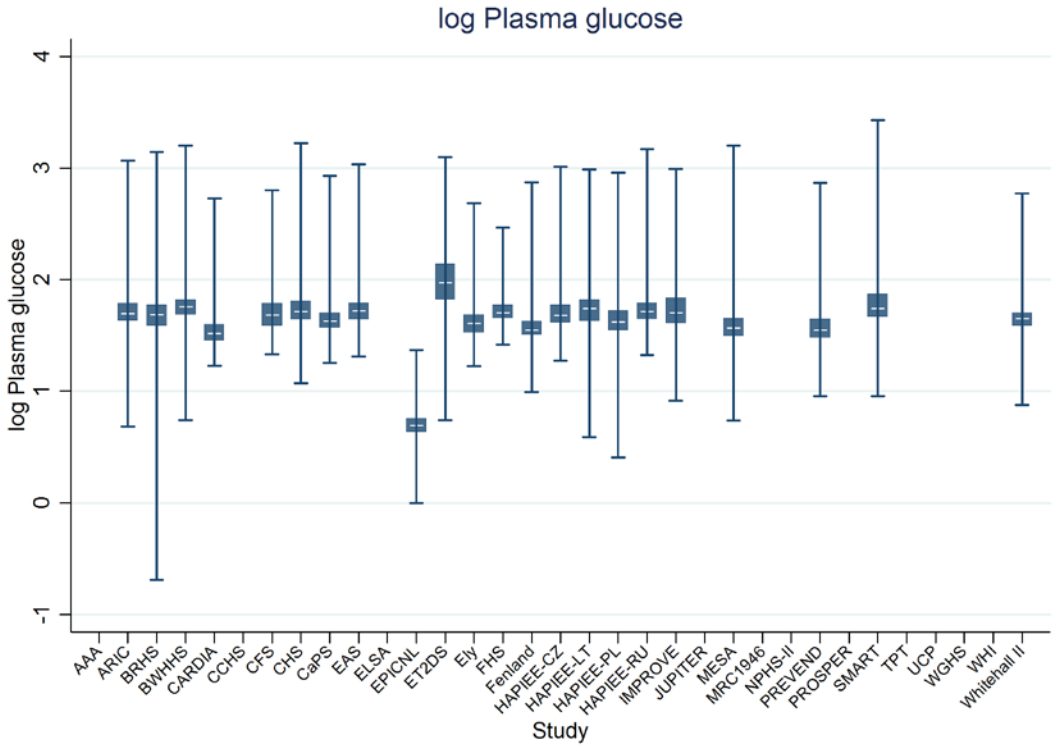

Height
Height, m
1
1.2
1.4
1.6
1.8
2
AAA
ARIC
BRHS
BWHHS
CARDIA
CCHS
CFS
CHS
CaPS
EAS
ELSA
EPICNL
ET2DS
Ely
FHS
Fenland
HAPIEE-CZ
HAPIEE-LT
HAPIEE-PL
HAPIEE-RU
IMPROVE
JUPITER
MESA
MRC1946
NPHS-II
PREVEND
PROSPER
SMART
TPT
UCP
WGHS
WHI
Whitehall II
Study


Hip circumference
Hip circumference, cm
50
100
150
200
AAA
ARIC
BRHS
BWHHS
CARDIA
CCHS
CFS
CHS
CaPS
EAS
ELSA
EPICNL
ET2DS
Ely
FHS
Fenland
HAPIEE-CZ
HAPIEE-LT
HAPIEE-PL
HAPIEE-RU
IMPROVE
JUPITER
MESA
MRC1946
NPHS-II
PREVEND
PROSPER
SMART
TPT
UCP
WGHS
WHI
Whitehall II
Study


Waist circumference
Waist circumference, cm
0
50
100
150
200
AAA
ARIC
BRHS
BWHHS
CARDIA
CCHS
CFS
CHS
CaPS
EAS
ELSA
EPICNL
ET2DS
Ely
FHS
Fenland
HAPIEE-CZ
HAPIEE-LT
HAPIEE-PL
HAPIEE-RU
IMPROVE
JUPITER
MESA
MRC1946
NPHS-II
PREVEND
PROSPER
SMART
TPT
UCP
WGHS
WHI
Whitehall II
Study

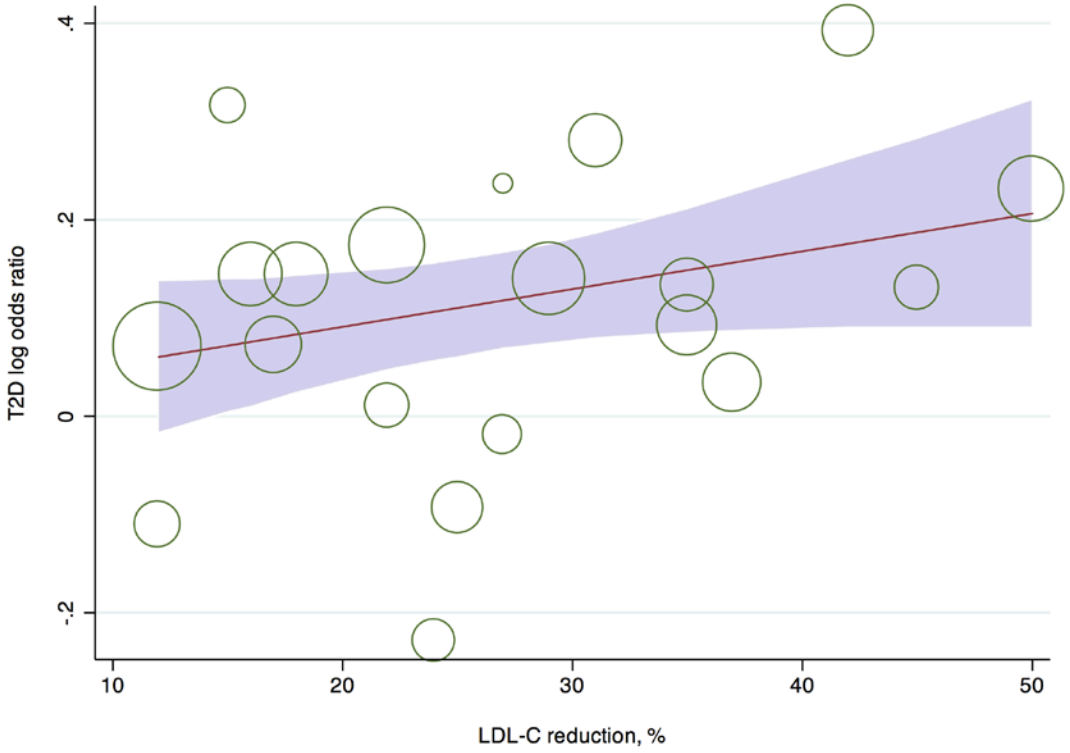

Waist:hip ratio
Waist:hip ratio
0
.5
1
1.5
2
2.5
AAA
ARIC
BRHS
BWHHS
CARDIA
CCHS
CFS
CHS
CaPS
EAS
ELSA
EPICNL
ET2DS
Ely
FHS
Fenland
HAPIEE-CZ
HAPIEE-LT
HAPIEE-PL
HAPIEE-RU
IMPROVE
JUPITER
MESA
MRC1946
NPHS-II
PREVEND
PROSPER
SMART
TPT
UCP
WGHS
WHI
Whitehall II
Study


log Plasma insulin
log Plasma insulin
-2
0
2
4
6
8
AAA
ARIC
BRHS
BWHHS
CARDIA
CCHS
CFS
CHS
CaPS
EAS
ELSA
EPICNL
ET2DS
Ely
FHS
Fenland
HAPIEE-CZ
HAPIEE-LT
HAPIEE-PL
HAPIEE-RU
IMPROVE
JUPITER
MESA
MRC1946
NPHS-II
PREVEND
PROSPER
SMART
TPT
UCP
WGHS
WHI
Whitehall II
Study


log Plasma glucose
log Plasma glucose
-1
0
1
2
3
4
AAA
ARIC
BRHS
BWHHS
CARDIA
CCHS
CFS
CHS
CaPS
EAS
ELSA
EPICNL
ET2DS
Ely
FHS
Fenland
HAPIEE-CZ
HAPIEE-LT
HAPIEE-PL
HAPIEE-RU
IMPROVE
JUPITER
MESA
MRC1946
NPHS-II
PREVEND
PROSPER
SMART
TPT
UCP
WGHS
WHI
Whitehall II
Study

Supplementary Figure 4 – Meta-regression model of the relationship between new-onset diabetes and percentage LDL-C reduction in statin trials at 1 year

20 trials, meta-regression p=0.10. Shaded area represents 95% confidence interval. The size of markers is determined by the standard error of the within-trial estimate of T2D odds ratio.



Supplementary Figure 5 –Meta-regression model of the relationship between change in weight and relative change in LDL-C in statin trials at 1 year

15 trials, meta-regression p=0.58. Shaded area represents 95% confidence interval. The size of markers is determined by the standard error of within-trial estimates of weight change.