Genetic assessment of potential long-term on-target side effects of PCSK9 inhibitors

Christopher P Nelson PhD,[1,2] Florence Y Lai MPhil[1,2], Mintu Nath PhD[1,2], Shu Ye PhD,[1,2] Thomas R Webb PhD[1,2], Heribert Schunkert MD[3,4], Nilesh J Samani FRCP[1,2]

[1]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK

[2]National Institute for Health Research Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK

[3]Deutsches Herzzentrum München, Technische Universität München, Germany

[4]Deutsches Zentrum für Herz- und Kreislaufforschung (DZHK), Munich Heart Alliance.

Nelson: Side effects of genetic PCSK9 inhibition

Correspondence to

Dr Nilesh J Samani, Department of Cardiovascular Sciences, University of Leicester, BHF Cardiovascular Research Centre, Glenfield Hospital, Groby Rod, Leicester, LE3 9QP, UK.

e-mail: njs@le.ac.uk Tel. No: (+44)116 2044756 Fax No. (+44)116 2875792

Total word count: 8193

Abstract

Background: Although short-term trials have suggested that PCSK9 inhibitors are safe and reduce risk of cardiovascular diseases, their long-term safety is unclear. Genetic variants associated with lower activity of a gene can act as proxies to identify potential long-term side effects of drugs as recently exemplified by association of LDL-lowering variants in the *HMGCR* (target for statins) and *PCSK9* genes with increased risk of type 2 diabetes mellitus (T2DM). However, analyses of the full spectrum of potential side effects of PCSK9 inhibition using a genetic approach has not been undertaken.

Methods: We examined the association of a LDL-lowering variant in the *PCSK9* gene (T allele of rs11591147), as well as two LDL-lowering *HGCMR* variants (G allele of rs17238484 and T allele of rs12916) with 80 diseases and traits in up to 479,522 individuals in UK Biobank.

Results: The *PCSK9* T allele was significantly (Bonferroni $P<6.25x10^{-4}$) associated with risk of T2DM, increased body mass index, waist circumference, waist-hip ratio, diastolic blood pressure, Type 1 diabetes mellitus (T1DM) and insulin use. The *HMGCR* variants were also associated with risk of T2DM although their previously reported associations with anthropometric traits were found to be confounded. Mediation analysis suggested that the association of the *PCSK9 T* allele with risk of T2DM but not DBP was largely independent of its association with BMI and central obesity. Nominally significant associations of the *PCSK9 T* allele were also seen with peptic ulcer disease, depression, asthma, chronic kidney disease and venous thromboembolism.

Conclusions: Our findings support previous genetic analyses suggesting that long-term use of PCSK9 inhibitors, like statins, may be associated with increased risk of T2DM. Some other

potential side-effects need to be looked for in future studies of PCSK9 inhibitors although we did not find signals that raise substantial concerns about their long-term safety.

Introduction

Elevated level of plasma low-density lipoprotein (LDL) cholesterol is a major causal risk factor for coronary heart disease (CHD). Statins which reduce LDL-cholesterol by inhibiting HMG-CoA reductase (HMGCR) have been shown to be beneficial for both primary and secondary prevention of CHD. However, the benefits of any drugs need to be balanced against possible side-effects and risks. In this context, the late emergence of evidence that statins increase risk of type 2 diabetes mellitus (T2DM), after almost two decades of increasing statin use and increasingly broader recommendations in guidelines,[1-3] sparked considerable debate and public controversy as to their appropriateness especially for primary prevention in relatively low-risk individuals, and led to a change in statin safety labelling by the US Food and Drug Administration (FDA).[4] Could the effect of statins on risk of T2DM have been predicted earlier? Swerdlow et al.[5] used a genetic approach (Mendelian randomisation, MR) to investigate this issue. They showed that two single nucleotide variants (G allele of SNP rs17238484 and T allele of SNP rs12916) in the *HMGCR* gene, that were associated with lower plasma LDL-cholesterol (rs17238484 G allele: -0.063 mmol/l per allele; rs12916 T allele: -0.073 mmol/l per allele)[6] were also associated with higher body weight, waist circumference, plasma insulin and glucose concentrations and risk of T2DM. In effect, these alleles acted as partial proxies for HMGCR inhibition by statins and their analysis illustrated that if suitable datasets are available then a genetic approach can help to identify potential on-target side-effects of drug treatments.

Proprotein convertase subtilisin/kexin type 9 (PCSK9) blockers have emerged as a new and powerful treatment for lowering LDL-cholesterol[7,8] and in randomised clinical trials have been shown to reduce cardiovascular events without any major side effects emerging[8]

(ODYSSEY Outcomes ACC18: http://www.acc.org/latest-in-cardiology/clinical-trials/2018/03/09/08/02/odyssey-outcomes). The first two of these agents, alirocumab and evolocumab, have been approved by the US FDA and the UK National Institute for Health and Care Excellence for lowering cholesterol where statins and other drugs are not tolerated or insufficient. PCSK9 is a secreted hepatic protein that binds to the LDL receptor and targets it for endocytosis and intracellular degradation, thereby reducing the number of LDL receptors at the cell surface and increasing plasma cholesterol.[9] The development of PCSK9 inhibitors was stimulated by the demonstration that nonsense mutations and other coding variants in the *PCSK9* gene are associated with lower plasma LDL-cholesterol and lower life-long risk of CHD.[9] Specifically a low frequency variant 137G -> T, replacing arginine at position 46 with leucine, is associated with a ~15% lower (-0.497 mmol/l per copy of the allele) LDL-cholesterol.[6,10] This variant, the T allele of the single nucleotide polymorphism (SNP) rs11591147 can therefore act as a proxy for PCSK9 inhibition by PCKS9 inhibitors, in a similar fashion to the HMGCR variants for statins. For PCSK9 inhibitors, while trials to date have not found evidence of an increase in T2DM risk in those on treatment, this has only been assessed on a relatively short follow-up of only three to four years[8] (ODYSSEY Outcomes ACC18: http://www.acc.org/latest-in-cardiology/clinical-trials/2018/03/09/08/02/odyssey-outcomes). Indeed, initial MR studies have shown that genetic risk scores including this variant or other variants at the *PCSK9* locus associated with lower LDL-cholesterol are also associated with increased risk of T2DM raising the possibility that, as with statins, this side effect may emerge with longer use of PCSK9 blockers.[11,12] No study has yet explored the full spectrum of long term potential side effects of PCSK9 inhibitors using a genetic approach.

UK Biobank is a large (~ 500,000 individuals) national health research resource established to understand the genetic and environmental determinants of adult diseases.[13] DNAs of UK Biobank participants have been assayed using a specially designed array providing, after imputation, genome-wide genotypes on over 70 million SNPs. Genotype-data on around 480,000 of the participants have been released. The availability of this data together with the extensive phenotypic data on the participants allows a detailed exploration to see if the *PCSK9* rs11591147 T variant is associated with any of a wide range of diseases or traits which may indicate other potential on-target effects of long-term inhibition PCSK9 inhibitors and inform their use. Here we have undertaken such an analysis and at the same time reexamined the associations of the two *HMGCR* SNPs studied by Swerdlow et al.[5] with BMI and T2DM as well as a broader range of diseases.

Methods

Study population and design: Details of the design of the UK Biobank study have been reported previously.[13] Participants were members of the UK general population aged between 40-69 years at recruitment, identified through primary care lists who accepted an invitation to attend one of the 22 assessment centers that were serially established across the UK between 2006 and 2010. The UK Biobank study was approved by the North West Multi-centre Research Ethics Committee (MREC) and all participants provided written informed consent.

At recruitment, detailed information was collected via a standardised questionnaire on socio-demographic characteristics, health status and physician-diagnosed medical conditions, family history and lifestyle factors. Selected physical and functional measurements were obtained including height, weight, waist-hip ratio (WHR) and systolic (SBP) and diastolic (DBP) blood pressures. The UK Biobank data were subsequently linked to Hospital Episode Statistics (HES) data, as well as national death and cancer registries. The HES data available for the current analysis covers all hospital admissions to NHS hospitals in England and Scotland from April 1997 to March 2015, with the Scottish data dated back as early as 1981. HES uses International Classification of Diseases ICD 9 and 10 to record diagnosis information, and OPCS-4 (Office of Population, Censuses and Surveys: Classification of Interventions and Procedures, version 4) to code operative procedures. Death registries include all deaths in the UK up to December 2015, with both primary and contributory causes of death coded in ICD-10. Cancer registries cover registrations across the UK from 1970s to the end of 2014 with the diagnoses (coded in ICD9 and 10).

Disease definition: We integrated self-reported physician-diagnosed medical conditions at recruitment with HES data and death and cancer registries where possible to define the number of individuals with different disorders, for the ~480,000 participants for which genetic data is available. The case definitions for all disorders studied are shown in **Supplementary Table S1**. For ease of navigating thorough the data, we grouped disorders/traits into cardiovascular, endocrine, neurological, digestive, genito-urinary, musculoskeletal, respiratory, eye, cancer and others. Given the allele frequency of the T allele of rs11591147 in UK Biobank (1.67%) participants, we calculated that we had 80% power to detect a nominal association with a categorical disease phenotype at an odds ratio of 1.5 with 1,150 cases for this variant. We therefore excluded diseases that had accrued less than 1,150 cases at the date of data download in June 2017, where the HES data covers all NHS hospital admissions up to March 2015, and death registries up to January 2016.

Genotypes: Genotyping of UK Biobank participants was undertaken using a custom-build genome-wide array (the UK Biobank Axiom array: http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/UK-Biobank-Axiom-Array-Datasheet-2014.pdf). Genotyping was done in two phases. 50,000 subjects were initially typed as part of the UK BiLEVE project.[14] Rest of the participants were genotyped using a slightly modified array. Imputation and initial quality checks of the genetic data were undertaken by UK Biobank. We extracted genotypes for rs11591147, the strongest variant associated with LDL-C reduction in PCSK9, as well as the two *HMGCR* SNPs (rs17238484 and rs12916) from the full UK Biobank imputed genetic dataset release on 487,406 participants. Of these 7,884 subjects were excluded due to sex mismatches, ambiguous sex, high heterozygosity, excess relatives, a call rate <98% or patient withdrawal. Before undertaking any analysis, we confirmed the quality of the genotypes for the *PCSK9* and *HMGCR* variants by confirming that the allele

frequencies were similar to those reported in the literature and that the genotype frequencies satisfied Hardy-Weinberg equilibrium in the UK Biobank subjects.

Statistical analysis: We compared the proportion of individuals with disease or trait in groups partitioned by genotype and derived odds ratios per copy of the allele associated with lower LDL-cholesterol for binary phenotypes and beta coefficients for quantitative phenotypes. Due to the level of relatedness within UK Biobank a family identifier was derived based on kinship scores, where any individuals found to be related (kinship>0.044) were joined into a family creating 395,232 unique families. All regression analyses were undertaken using mixed models in Stata adjusting for age, sex, array type (BiLEVE array versus main UK Biobank array), the first five principal components and lipid-lowering medication use as fixed effects, with family fitted as a random effect using an independent covariance structure. A Bonferroni corrected statistical significance was set at $p<6.25x10^{-4}$ to account for the 80 phenotypes tested.

To examine the possibility that any observed associations with our genetic variants were due to one or more linked variants in the region we examined the associations of other variants in the region around both the *PCSK9 and HMGCR* genes with selected traits. We excluded genetic variants that had low imputation quality (info<0.4) or with a rare allele frequency (<1%), or were not in Hardy-Weinberg equilibrium ($p<1x10^{-8}$). Where another signal was found to have a stronger p-value we performed a joint model to test the effect of our lead SNP conditional on all other signals after excluding those in high linkage disequilibrium (LD) with $r^2>0.8$ to avoid collinearity.

Given the relationship between these variants and the reduced risk of lipid-lowering medication use (see Results) an adjustment for lipid-lowering medication was included as a

fixed effect due the potential confounding effect of treatment with disease risk. In traits that were shown to be associated with our lead variants after adjustment for other SNPs in the region we performed stratified analyses by lipid-lowering medication use in order to assess for potential collider bias[15].

To relate findings from the genetic analysis to potential therapeutic inhibition of LDL-cholesterol we used externally available estimates of LDL-cholesterol reduction for the T allele of rs11591147[6] and the absolute reduction in LDL-cholesterol using the PCSK9 inhibitor evolocumab[8,16]. With these we can estimate the potential effect size of blocking PCSK9 on disease risk if the effect is driven entirely through reductions in LDL-cholesterol.

To examine whether the associations of rs11591147 with T2DM and DBP could be explained by its association with BMI or WHR we examined the associations after adjustment for these variables. Subsequently a formal mediation analysis was performed for T2DM using BMI and waist-hip-ratio (WHR) as separate mediators of the effect. For these analyses genetically correlated samples were excluded leaving only one sample from each genetically similar family group. We estimated direct and indirect effects using `paramed` (Emsley R, Liu H. PARAMED: Stata module to perform causal mediation analysis using parametric regression models. Statistical software components. 2013) and estimated the proportion of the effect that is mediated using bootstrapping[17] in Stata.

Results

In total we analyzed phenotype and genotype information on up to 479,522 individuals in UK Biobank. The demographic characteristics of the subjects is shown in **Table 1**. The mean age at recruitment was 56.5 years (range 37-73), 54.2% were female and 5.3% reported a non-white ethnic background. 17.6% and 21.0% of the participants were taking lipid-lowering medication and blood pressure-lowering medication, respectively.

The number of individuals with the various disorders/traits that we analyzed are shown in **Table 2**. These include both prevalent cases (i.e. those reported at the time of enrolment) and subsequent cases as recorded in HES or in disease registries. The proportion of each disorder based on reported history only (i.e. not recorded in the available HES data or in registries) are also shown in **Table 2**.

The allele frequency of the *PCSK9* rs11591147 T allele was 1.67% (n= 15,862 carriers) while the frequencies for the *HMGCR* alleles at rs17238484 (G allele) and rs12916 (T allele) that are associated with lower LDL-cholesterol were 77.0% (n = 453,891 carriers) and 59.9% (401,891 carriers), respectively. All variants were well imputed (Info score 1 for rs11591147 and rs17238484 and 0.997 for rs12916) and all SNPs were in Hardy-Weinberg equilibrium. The demographic characteristics of individuals carrying different number of copies of these variants are shown in **Supplementary Table S2**.

Consistent with their known association with lower LDL-cholesterol levels, all variants were associated with a lower odds of the carriers being on lipid-lowering medication (**Tables 3** and **4**). The effect size was greater for the *PCSK9* rs11591147 T allele (**Table 3**), than for either of the *HMGCR* alleles (**Table 4**), reflecting its greater effect on LDL-cholesterol. Significant associations with other diseases and traits, after adjustment for lipid-lowering

medication and other variables described in Methods and allowing for multiple testing, for the *PCSK9* rs11591147 T allele and the lead *HMGCR* variant (either rs17238484 G allele or rs12916 T allele) are shown in **Tables 3** and **4,** respectively. Full results for all diseases and traits analysed are shown in **Supplementary Tables S3-5** for models with and without adjustment for lipid-lowering medication.

The *PCSK9* rs11591147 T allele was associated with a higher risk of T2DM as well as higher BMI, waist circumference and waist-hip ratio (**Table 3**). There was a consistent 7.2% (3.0%-11.6%) higher odds of obesity (BMI > 30 kg/m^2). The *PCSK9* rs11591147 T allele was also associated with an average 0.33 (95% CI 0.16, 0.50) mmHg higher diastolic BP and a greater use of blood pressure lowering medication. Interestingly, there was also a significant association between the *PCSK9* rs11591147 T allele and Type 1 diabetes mellitus and use of insulin (**Table 3**), p=3.07x10^{-4} and p=3.29x10^{-4} respectively. Finally, the *PCSK9* rs11591147 T allele was associated with lower risk of CAD (**Table 3**). This association was stronger (OR 0.676 (95% CI 0.619, 0.738), p=2.625x10^{-18}) without adjustment for lipid-lowering medication (**Supplementary Table S3**).

Regional association plots around the *PCSK9* gene for T2D, CAD, BMI, WHR and DBP are shown in **Supplementary Figure S1**. From these analyses it is clear that rs11591147 is the lead associated variant in the region for all the above traits except BMI where a second stronger signal is present (rs144932160). Adjusting for rs144932160 did not remove the association with rs11591147 (Beta=0.153 (95% CI 0.076-0.260); p=9.23 x 10^{-5}).

The associations that were significant for the two *HMGCR* variants were largely very similar to those for the *PCSK9* rs11591147 T allele except that the effect sizes were generally smaller. The result for the variant showing the strongest association is shown in **Table 4** with

full results in **Supplementary Table S4 and S5**. Both *HMGCR* variants were associated with increased risk of T2DM as well as higher BMI, waist circumference, waist-hip ratio, obesity, as well as increased systolic BP, risk of hypertension, risk of stroke and use of blood-pressure lowering medication. There were nominal associations (p=0.04) of both HMGCR variants with risk of Type 1 diabetes mellitus (**Supplementary Tables S4** and **S5**). The HMGCR rs12916 T allele was also nominally associated with risk of CAD (OR 0.970 (0.951, 0.990); p=0.003) which was attenuated after adjustment for lipid-lowering medication (**Supplementary Table S5**). However for all the traits/diseases, except for T2DM and stroke, there were other SNPs in the region not in high LD with our lead variants which showed a stronger association (**Supplementary Figure S2**). Adjusting for these variants led to an attenuation of the associations with the lead variants (**Table 4**, joint estimates). Notably, the associations with BMI, waist circumference, WHR and systolic BP all became non-significant, although the association with history of hypertension and BP medication persisted.

For traits significantly associated with *PCSK9* rs11591147 T allele, analyses stratified by lipid-lowering medication use showed consistent effects in both groups (**Supplementary Table S6**) with only CAD risk, as expected, being attenuated in the lipid-lowering medication strata. This suggests our results are not due to collider bias. Similarly for the *HMGCR* variants we observe significant associations in both strata for the four traits - T2DM, stroke, hypertension and blood pressure medication - that were still associated with our *HMGCR* variants after adjusting for other variants in the region (**Supplementary Table S7**).

To investigate if the observed associations between rs11591147 and T2DM and DBP are mediated through effects on adiposity, we considered both BMI and WHR as potential mediators. We found that the association of the variant on DBP became non-significant

after adjusting for BMI (p=0.480) or WHR (p=0.543). On the other hand the association with T2DM risk remained largely unchanged after further adjusting for BMI (OR=1.26 (1.16, 1.37)) or WHR (OR of 1.24 (1.15, 1.35)). In formal mediation analysis we found that the direct association of rs11591147 with T2DM had an OR=1.24 (1.14, 1.34) and the indirect effect through BMI to have an OR=1.02 (1.01, 1.03). The proportion of the association of rs11591147 with T2DM that was mediated was estimated to be just 0.09% with BMI and 3.14% with WHR.

The *PCSK9* rs11591147 T allele is associated with a 0.497 mmol/l lower LDL-cholesterol on average.[6] The PCSK9 antibody blocker evolocumab has been reported to reduce LDL-cholesterol by an average of 1.81 mmol/l (70mg/dl) over standard therapy, which may or may not have included lipid-lowering medications.[8] Therefore, if the association of PCSK9 inhibition with risk of T2DM was driven entirely through and proportionate to its LDL-lowering effect then we would estimate the increased risk of T2DM with evolocumab to be 2.47 (1.89 , 3.35) fold higher. When given on top of lipid-lowering medication, as in the FOURIER trial[16] evolocumab further reduced cholesterol on average by 50mg/dl (1.295 mmol/l). This would result in an increased odds ratio of risk of T2DM of 1.90 (1.57, 2.36).

Discussion

Major side-effects of drugs are sometimes identified during their initial clinical evaluation and more systematically during randomised double-blind clinical trials (RCTs) when they are compared with placebo. However many RCTs are of relatively short duration and may not pick up side effects that are weak and/or only emerge during longer-term therapy. The evidence that statin therapy is associated with an increased risk of T2DM only emerged late

in the context of its widespread clinical use and raised considerable professional and public concern and debate.[1-3]

Use of genetic variants as un-confounded instruments (MR) to clarify whether observed epidemiological associations between biomarkers and disease are causal or not, particularly as a means of is identifying reliable therapeutic targets, is gaining increasing traction.[18] In an extension of this approach Swerdlow et al[5] demonstrated that such genetic instruments could also help to flag up potential longer term on target side-effects of drugs by showing that variants in the *HMGCR* gene associated with lower LDL-cholesterol were also associated with T2DM.

Because the genetic instruments used in such analysis usually have a subtle effect on the target gene (unlike the drug), large datasets are required to discern any association with a downstream side-effect. Unless a specific side-effect is being evaluated, such datasets should also ideally information on a broad variety of outcomes. In this regard, the UK Biobank cohort[13] offers an excellent resource because of its scale and comprehensiveness as well as the availability of genetic data. Using UK Biobank, we confirmed the association of LDL-lowering variants in the *HMGCR* gene with T2DM and also found, as reported recently[10,11] that genetic variation at the *PCSK9* locus associated with increased risk of T2DM. Therefore, despite no such signal emerging in the trials of PCSK9 blockers to date, which have been of a duration of 48 months or less,[8,16] one should anticipate that with longer-term use of such drugs T2DM is likely to emerge as a side effect. If the effect is related to the level of LDL-lowering we may except to see an approximately two-fold increase in risk of T2DM.

An important consideration, especially when considering the joint association of a variant with multiple traits, is to exclude the possibility that one (or more) of the associations is actually due to a stronger effect of a linked variant with the association found with the selected variant representing a "spillover" association. In their study of the two *HMGCR* variants, Swedlow et al.[5] also observed an association with BMI and waist circumference and proposed that the association of the variants (and by inference lipid-lowering medication) with T2DM may be mediated via a primary association with insulin resistance and central adiposity. Here we clearly show that while the G allele of rs17238484 and the T allele of rs12916 are independently associated with T2DM, their association with anthropometric traits is confounded.

In contrast to the *HMGCR* variants, *PCSK9* rs11591147 has the strongest association in the region for not only T2DM but also for WHR and DBP. There was one stronger signal for BMI but adjusting for this did not substantially attenuate the association of rs11591147 with BMI. This suggests that unlike the *HMGCR* variants, rs11591147 has primary associations with anthropometric traits. We therefore explored to what extent these associations mediate the association with DBP and T2DM. This analysis showed that while the association of rs11591147 T allele with DBP may be mediated via central adiposity, the association with T2DM is largely independent.

The lead variants were all highly significantly associated with lower use of lipid-lowering medication (**Tables 3 and 4**) validating their use as proxies for such treatment. Because of the potential confounding effect of such treatment on disease risk mediated by the genotype, we adjusted for use of lipid-lowering medication in our primary analysis. Indeed, the association of *PCSK9* rs11591147 T allele with risk of CAD was partly attenuated by such

adjustment (**Supplementary Table S3**). On the other hand several of the observed associations were substantially stronger or only emerged when adjustment was made for lipid-lowering medication (**Supplementary Tables S3-S5**). Although likely to reflect the true association of genotype with disease/trait, it raises the possibility of collider bias whereby conditioning on a common effect (LDL-lowering medication) of a pair of variables (genotype and trait) may cause a spurious association between them[15]. To investigate this possibility we undertook association analyses stratified by lipid-lowering medication use for those traits/diseases showing an association in the primary analysis (**Supplementary Tables S6 and S7**). Although the proportion of disease cases or mean values of the quantitative traits differed in individuals taking or not taking lipid-lowering medication we found associations in both groups indicating that the observed associations are not due to collider bias (**Supplementary Tables S6 and S7**). Recently, large scale GWAS analyses across a wide range of phenotypes in UK Biobank, including T2D, CAD, BMI, WHR and DBP have been reported (https://biobankengine.stanford.edu/ and https://data.broadinstitute.org/alkesgroup/UKBB/). These analyses include the three variants studied here and report associations similar to our analyses unadjusted for lipid-lowering medication. Our findings emphasise the potential importance of considering confounders when interpreting the results of such GWAS analyses.

The *PCSK9* rs11591147 T allele was associated with a higher risk of not only T2DM but also T1DM. The association of this allele also with increased use of insulin (**Table 3**) suggests that this disease association is genuine. The mechanism for the association is unclear. It should be noted that while the main site of *PCSK9* expression is in the liver it is also expressed in several other tissues including the small intestine, colon, kidney and cerebellum. [19]

Furthermore, *PCSK9* null mice display hypoinsulinemia, hyperglycemia, glucose intolerance and have pancreatic islet abnormalities.[20] PCSK9 inhibitors have been shown to be particularly efficacious in lowering LDL-cholesterol in patients with familial hypercholesterolaemia.[21] Childhood initiation of lipid-lowering medication is recommended by guidelines in patients with FH (https://www.nice.org.uk/guidance/cg71/resources/familial-hypercholesterolaemia-identification-and-management-pdf-975623384005) and the possible risk of precipitating T1DM in this group by treatment with PCSK9 blockers needs to be born in mind.

A potentially reassuring finding from our analysis is that apart from the associations discussed above, we did not observe any other major safety signals across a broad range of systems, including cancers and dementia, for either the *PCSK9* rs11591147 T allele or indeed the *HMGCR* variants. The findings with respect to the latter are consistent with the established safety profile of statins. For the *PCSK9* rs11591147 T allele the only other conditions that achieved association of nominal significance were peptic ulcer disease (OR 1.145 (1.043, 1.256); p=0.005), depression (OR 1.089 (1.026, 1.157); p=0.005); asthma (OR1.063 (1.013, 1.115); p=0.012), chronic kidney disease (OR 1.176 (1.015, 1.361); p=0.031); multiple sclerosis (OR 0.741 (0.560, 0.981); p=0.036) and venous thromboembolism (OR 1.092 (1.005, 1.186), p=0.038) (**Supplementary Table S3**). Because of multiple testing, these associations should be viewed with caution but nonetheless, the incidence the incidence of these conditions should be particularly monitored in future clinical trials and longer-term safety monitoring of anti-PCSK9 therapy.

Although not the major objective of our study, our analysis also allows us to comment on whether anti-PCSK9 therapy may have cardiovascular benefits beyond reducing risk of CAD.

In this regard we found no evidence of that the *PCSK9* rs11591147 T was protective towards risk of stroke (OR 0.923 (0.833, 1.024) p= 0.130 before adjustment for lipid-lowering medication and OR 1.056 (0.949, 1.176) p= 0.314 after adjustment). For aortic valve stenosis (AVS) there was a trend towards a protective association (OR 0.732 (0.555, 0.965), p= 0.027 before adjustment for lipid-lowering medication and OR 0.837 (0.634, 1.106); p=0.210 after adjustment). A protective association of this variant on risk of AVS has also been reported by others.[22] LDL-lowering through statins have not been found to reduce the rate of progression of AVS.[23] However, anti-PCSK9 therapy also lowers lipoprotein a (Lp(a)) levels[8] and both genetic and biomarker studies support a causal role of Lp(a) in the aetiology of AVS.[24,25] Therefore, our genetic findings support the possibility that anti-PCSK9 therapy may be beneficial in preventing the development or progression of AVS.

Several limitations and caveats of our analysis need to be highlighted. First, in trying to identify potential side effects (or additional benefits) we are extrapolating any impact of a life-long (including developmental) exposure to a relatively modest genetic reduction in *PCSK9* gene activity to that which might occur from a much more powerful drug inhibition of the protein products of the genes over a shorter-term time frame in adults. As an extension to this, if the side effects are dependent on the degree of inhibition, then major side-effects may not be detectable in genetic analysis despite large datasets. Second, although UK Biobank currently provides the largest and most comprehensive dataset to investigate potential side effects using a genetic approach, the number of cases for several disorders is still small reducing the power to detect (or exclude) associations that may be clinically relevant. This will improve with time as UK Biobank accrues more cases. Third, UK Biobank is made up predominantly of white Caucasian individuals (~ 95%) and we were unable to

analyse for any ethnic-specific associations. Finally, genetic analysis of the type undertaken here does not allow off-target effects of the drugs (i.e. those not related to inhibition of the primary target) to be identified and which may be molecule specific.

In summary, we show that despite their favourable short-term profile and clear benefit in terms of reducing cardiovascular events, longer term use of anti-PCSK9 therapy is likely to be associated with higher risk of T2DM and possibly insulin-requiring diabetes as well as development of several features of the metabolic syndrome. Beyond this we did not find any evidence for any other major longer term on-target side effects although our power to identify potentially clinically meaningful associations for some of the diseases was limited.

Acknowledgments

We are grateful to UK Biobank for access to their data. This analysis was performed under UK Biobank approved project 9922.

Sources of Funding

CPN, SY, TRW and NJS are funded by the British Heart Foundation. FL is funded by the UK National Institute for Health Research.

Disclosures

HS has received honoraria for consulting for AstraZeneca, MSD/Merck, Amgen and Sanofi.

References

1. Sattar N, Preiss D, Murray HM, Welsh P, Buckley BM, de Craen AJ, Seshasai SR, McMurray JJ, Freeman DJ, Jukema JW, Macfarlane PW, Packard CJ, Stott DJ, Westendorp RG, Shepherd J, Davis BR, Pressel SL, Marchioli R, Marfisi RM, Maggioni AP, Tavazzi L, Tognoni G, Kjekshus J, Pedersen TR, Cook TJ, Gotto AM, Clearfield MB, Downs JR, Nakamura H, Ohashi Y, Mizuno K, Ray KK, Ford I. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. *Lancet* 2010; 375:735–42. doi: 10.1016/S0140-6736(09)61965-6.

2. Preiss D, Seshasai SRK, Welsh P, Murphy SA, Ho JE, Waters DD, DeMicco DA, Barter P, Cannon CP, Sabatine MS, Braunwald E, Kastelein JJ, de Lemos JA, Blazing MA, Pedersen TR, Tikkanen MJ, Sattar N, Ray KK. Risk of incident diabetes with intensive-dose compared with moderate-dose statin therapy: a meta-analysis. *JAMA* 2011; 305: 2556–64. doi: 10.1001/jama.2011.860.

3. Carter AA, Gomes T, Camacho X, Juurlink DN, Shah BR, Mamdani MM. Risk of incident diabetes among patients treated with statins: population based study. *BMJ* 2013; 346:f2610. doi: 10.1136/bmj.f2610.

4. US Food and Drug Administration. FDA Drug Safety Communication: important safety label changes to cholesterol-lowering statin drugs.2012. http://www.fda.gov/ Drugs/ DrugSafety/ucm293101.htm

5. Swerdlow DI, Preiss D, Kuchenbaecker KB, Holmes MV, Engmann JE, Shah T, Sofat R, Stender S, Johnson PC, Scott RA, Leusink M, Verweij N, Sharp SJ8, Guo Y, Giambartolomei C, Chung C, Peasey A, Amuzu A, Li K, Palmen J, Howard P, Cooper JA, Drenos F, Li YR, Lowe G, Gallacher J, Stewart MC, Tzoulaki I, Buxbaum SG, van der A DL, Forouhi NG, Onland-Moret NC, van der Schouw YT, Schnabel RB, Hubacek JA, Kubinova R, Baceviciene M, Tamosiunas A, Pajak A, Topor-Madry R, Stepaniak U, Malyutina S, Baldassarre D, Sennblad B, Tremoli E,

de Faire U, Veglia F, Ford I, Jukema JW, Westendorp RG, de Borst GJ, de Jong PA, Algra A, Spiering W, Maitland-van der Zee AH, Klungel OH, de Boer A, Doevendans PA, Eaton CB, Robinson JG, Duggan D; DIAGRAM Consortium; MAGIC Consortium; InterAct Consortium, Kjekshus J, Downs JR, Gotto AM, Keech AC, Marchioli R, Tognoni G, Sever PS, Poulter NR, Waters DD, Pedersen TR, Amarenco P, Nakamura H, McMurray JJ, Lewsey JD56, Chasman DI, Ridker PM, Maggioni AP, Tavazzi L, Ray KK, Seshasai SR, Manson JE, Price JF, Whincup PH, Morris RW, Lawlor DA, Smith GD, Ben-Shlomo Y, Schreiner PJ, Fornage M, Siscovick DS, Cushman M, Kumari M, Wareham NJ, Verschuren WM, Redline S, Patel SR, Whittaker JC, Hamsten A, Delaney JA, Dale C, Gaunt TR, Wong A, Kuh D, Hardy R, Kathiresan S, Castillo BA, van der Harst P, Brunner EJ, Tybjaerg-Hansen A, Marmot MG, Krauss RM, Tsai M, Coresh J, Hoogeveen RC, Psaty BM, Lange LA, Hakonarson H, Dudbridge F, Humphries SE, Talmud PJ, Kivimäki M, Timpson NJ, Langenberg C, Asselbergs FW, Voevoda M, Bobak M, Pikhart H, Wilson JG, Reiner AP, Keating BJ, Hingorani AD, Sattar N. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* 2015; 385:351-61. doi: 10.1016/S0140-6736(14)61183-1.

6. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Do R, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkilä K, Hyppönen E, Isaacs A, Jackson AU, Johansson Å, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikäinen LP, Magnusson PKE, Mangino M, Mihailov E, Montasser ME, Müller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D,

Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney ASF, Döring A, Elliott P, Epstein SE, Ingi Eyjolfsson G, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P, Kastelein JJP, Khaw KT, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimäki T, Lin SY, Lindström J, Loos RJF, Mach F, McArdle WL, Meisinger C, Mitchell BD, Müller G, Nagaraja R, Narisu N, Nieminen TVM, Nsubuga RN, Olafsson I, Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruokonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stančáková A, Stirrups K, Swift AJ, Tiret L, Uitterlinden AG, van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen YI, Collins FS, Cooper RS, Danesh J, Dedoussis G, de Faire U, Feranil AB, Ferrières J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllensten U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Järvelin MR, Jula A, Kähönen M, Kaprio J, Kesäniemi A, Kivimaki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, März W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njølstad I, Pedersen NL, Power C, Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PEH, Sheu WH, Shuldiner AR, Siegbahn A, Spector TD, Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BHR, Ordovas JM, Boerwinkle E, Palmer CNA, Thorsteinsdottir U, Chasman DI,

Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Kathiresan S, Mohlke KL, Ingelsson E, Abecasis GR; Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013; 45:1274-83. doi: 10.1038/ng.2797.

7. Robinson JG, Farnier M, Krempf M, Bergeron J, Luc G, Averna M, Stroes ES, Langslet G, Raal FJ, El Shahawy M, Koren MJ, Lepor NE, Lorenzato C, Pordy R, Chaudhari U, Kastelein JJ; ODYSSEY LONG TERM Investigators. Efficacy and Safety of Alirocumab in Reducing Lipids and Cardiovascular Events. *N Engl J Med* 2015; 372:1489-99. doi: 10.1056/NEJMoa1501031.

8. Sabatine MS, Giugliano RP, Wiviott SD, Raal FJ, Blom DJ, Robinson J, Ballantyne CM, Somaratne R, Legg J, Wasserman SM, Scott R, Koren MJ, Stein EA; Open-Label Study of Long-Term Evaluation against LDL Cholesterol (OSLER) Investigators.. Efficacy and Safety of Evolocumab in Reducing Lipids and Cardiovascular Events. *N Engl J Med* 2015; 372:1500-09. doi: 10.1056/NEJMoa1500858.

9. Ajufo E, Rader DJ. Recent advances in the pharmacological management of hypercholesterolaemia. *Lancet Diabetes Endocrinol* 2016; 4:436-46. doi: 10.1016/S2213-8587(16)00074-7.

10. Cohen JD, Boerwinkle E, Mosley TH, Hobbs H. Sequence Variations in *PCSK9*, Low LDL, and Protection against Coronary Heart Disease. *N Engl J Med* 2006; 354:1264-72. doi: 10.1056/NEJMoa054013

11. Schmidt AF, Swerdlow DI, Holmes MV, Patel RS, Fairhurst-Hunter Z, Lyall DM, Hartwig FP, Horta BL, Hyppönen E, Power C, Moldovan M, van Iperen E, Hovingh GK, Demuth I, Norman K, Steinhagen-Thiessen E, Demuth J, Bertram L, Liu T, Coassin S, Willeit J, Kiechl S, Willeit K, Mason D, Wright J, Morris R, Wanamethee G, Whincup P, Ben-Shlomo Y, McLachlan S, Price JF, Kivimaki M, Welch C, Sanchez-Galvez A, Marques-Vidal P, Nicolaides

A, Panayiotou AG, Onland-Moret NC, van der Schouw YT, Matullo G, Fiorito G, Guarrera S, Sacerdote C, Wareham NJ, Langenberg C, Scott R, Luan J, Bobak M, Malyutina S, Pająk A, Kubinova R, Tamosiunas A, Pikhart H, Husemoen LL, Grarup N, Pedersen O, Hansen T, Linneberg A, Simonsen KS, Cooper J, Humphries SE, Brilliant M, Kitchner T, Hakonarson H, Carrell DS, McCarty CA, Kirchner HL, Larson EB, Crosslin DR, de Andrade M, Roden DM, Denny JC, Carty C, Hancock S, Attia J, Holliday E, O'Donnell M, Yusuf S, Chong M, Pare G, van der Harst P, Said MA, Eppinga RN, Verweij N, Snieder H; LifeLines Cohort study group, Christen T, Mook-Kanamori DO, Gustafsson S, Lind L, Ingelsson E, Pazoki R, Franco O, Hofman A, Uitterlinden A, Dehghan A, Teumer A, Baumeister S, Dörr M, Lerch MM, Völker U, Völzke H, Ward J, Pell JP, Smith DJ, Meade T, Maitland-van der Zee AH, Baranova EV, Young R, Ford I, Campbell A, Padmanabhan S, Bots ML, Grobbee DE, Froguel P, Thuillier D, Balkau B, Bonnefond A, Cariou B, Smart M, Bao Y, Kumari M, Mahajan A, Ridker PM, Chasman DI, Reiner AP, Lange LA, Ritchie MD, Asselbergs FW, Casas JP, Keating BJ, Preiss D, Hingorani AD; UCLEB consortium, Sattar N. PCSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol* 2017; 5:97-105

12. Ference BA, Robinson JG, Brook RD, Catapano AL, Chapman MJ, Neff DR, Voros S, Giugliano RP, Davey Smith G, Fazio S, Sabatine MS. Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. *N Engl J Med* 2016; 375:2144-2153. doi: 10.1016/S2213-8587(16)30396-5

13. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R. 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: e1001779. doi: 10.1371/journal.pmed.1001779.

14. Wain LV, Shrine N, Miller S, Jackson VE, Ntalla I, Soler Artigas M, Billington CK, Kheirallah AK, Allen R, Cook JP, Probert K, Obeidat M, Bossé Y, Hao K, Postma DS, Paré PD, Ramasamy A; UK Brain Expression Consortium (UKBEC), Mägi R, Mihailov E, Reinmaa E, Melén E, O'Connell J, Frangou E, Delaneau O; OxGSK Consortium, Freeman C, Petkova D, McCarthy M, Sayers I, Deloukas P, Hubbard R, Pavord I, Hansell AL, Thomson NC, Zeggini E, Morris AP, Marchini J, Strachan DP, Tobin MD, Hall IP. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respiratory Medicine* 2015; 3:769-81. doi: 10.1016/S2213-2600(15)00283-0.

15. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, Poole C. Illustrating bias due to conditioning on a collider. *Int J Epidemiol.* 2010;39:417-420

16. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, Kuder JF, Wang H, Liu T, Wasserman SM, Sever PS, Pedersen TR; FOURIER Steering Committee and Investigators. Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease *N Engl J Med* 2017; 376:1713-1722. doi: 10.1056/NEJMoa1615664.

17. Buis ML. Direct and indirect effects in a logit model. *The Stata Journal* 2010; 10:11-29.

18. Burgess S, Timpson NJ, Ebrahim S, Davey Smith G. Mendelian randomization: where are we now and where are we going? *Int J Epidemiol* 2015; 44:379-88. doi: 10.1093/ije/dyv108.

19. Seidah NG, Benjannet S, Wickham L, Marcinkiewicz J, Jasmin SB, Stifani S, Basak A, Prat A, Chretien M. The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation. *Proc Natl Acad Sci USA* 2003; 100:928–33. doi: 10.1073/pnas.0335507100.

20. Mbikay M, Sirois F, Mayne J, Wang GS, Chen A, Dewpura T, Prat A, Seidah NG, Chretien M, Scott FW. PCSK9-deficient mice exhibit impaired glucose tolerance and pancreatic islets abnormalities. *FEBS Lett* 2010; 584:701-6. doi: 10.1016/j.febslet.2009.12.018.

21. Raal F, Scott R, Somaratne R, Bridges I, Li G, Wasserman SM, Stein EA. Low-density lipoprotein cholesterol-lowering effects of AMG145, a monoclonal antibody to proprotein convertase subtilsin/kexin 9 serine protease in patients with heterozygous familial hypercholesterolaemia. *Circulation* 2012; 126:2408-17. doi: 10.1161/CIRCULATIONAHA.112.144055.

22. Langsted A, Nordestgaard BG, Benn M, Tybjaerg-Hansen A and Kamstrup PR. PCSK9 R46L Loss-of-Function Mutation Reduces Lipoprotein(a), LDL Cholesterol, and Risk of Aortic Valve Stenosis. *J Clin Endocrinol Metab*. 2016;101:3281-7. doi: 10.1210/jc.2016-1206.

23. Thiago L, Tsuji S, Nyong J, Puga MES, Gois AFT, Macedo CR, Valente O, Atallah ÁN. Statins for aortic valve stenosis. Cochrane Database of Systematic Reviews 2016, Issue 9. Art. No.: CD009571. doi: 10.1002/14651858.CD009571.pub2

24. Cairns BJ, Coffey S, Travis RC, Prendergast B, Green J, Engert JC, Lathrop M, Thanassoulis G, Clarke R. A Replicated, Genome-Wide Significant Association of Aortic Stenosis With a Genetic Variant for Lipoprotein(a): Meta-Analysis of Published and Novel Data. *Circulation*. 2017;135:1181-1183. doi: 10.1161/CIRCULATIONAHA.116.026103.

25. Afshar M, Kamstrup PR, Williams K, Sniderman AD, Nordestgaard BG, Thanassoulis G. Estimating the Population Impact of Lp(a) Lowering on the Incidence of Myocardial Infarction and Aortic Stenosis-Brief Report. *Arterioscler Thromb Vasc Biol*. 2016;36:2421-2423. doi: 10.1161/ATVBAHA.116.308271

26. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med.* 2005;24:2911-35. doi: 10.1002/sim.2165

Table 1. Characteristics of UK Biobank participants analysed in this study

Characteristics		N	Mean(SD) or N(%)
Demographic	Age (years)	479,522	56.5 (8.1)
	Sex - female	479,522	260,046 (54.2%)
	Ethnic background		
	White	477,260	451,928 (94.7%)
	Asian		10,736 (2.2%)
	Black		7,523 (1.6%)
	Mixed and others		7,073 (1.5%)
Cardiovascular	Body mass index (BMI) kg/m2	477,580	27.4 (4.8)
risk factors	Waist circumference (cm)	478,465	90.3 (13.5)
	Hip circumference (cm)	479,522	103.4 (9.2)
	Waist -hip-ratio (WHR)	478,367	0.9 (0.1)
	Systolic Blood Pressure (SBP) mmHg	478,985	141.0 (20.7)
	Diastolic Blood Pressure (DBP) mmHg	478,987	84.3 (11.3)
	Ever smoked	477,070	215,906 (45.3%)
Medication	Lipid-lowering	474,902	83,442 (17.6%)
	Blood pressure lowering	474,902	99,848 (21.0%)

Quantitative data are shown as mean with standard deviation and categorical data as number and percentage.

Table 2. Frequency of selected disorders in UK Biobank

Disease group	Phenotype	N Case	N Control	% Self-reported
Cardiovascular	Coronary artery diseases (CAD)	24,293	438,479	14.90%
diseases	Atrial fibrillation (AF)	17,147	462,375	3.50%
	Heart failure (HF)	6,645	472,877	1.60%
	Peripheral vascular disease (PVD)	4,460	475,062	20.40%
	Venous thromboembolism	17,100	462,422	49.80%
	Aortic valve stenosis	2,023	477,499	1.90%
	Hypertension	259,535	219,987	64.60%
	Stroke	12,210	467,312	44.80%
Endocrine	Diabetes type I	3,555	475,967	3.40%
disorders	Diabetes type II	30,299	447,401	21.40%
	Hyperthyroid	5,239	474,283	57.40%
	Hypothyroid	27,908	451,614	41.20%
Neurological	Anxiety	12,372	467,150	44.50%
	Depression	34,890	444,632	59.80%
	Biopolar	1,939	477,583	33.70%
	Schizophrenia	1,468	478,054	14.40%
	Multiple sclerosis	2,001	477,521	23.40%
	Epilepsy	5,764	473,758	24.30%
	Dementia	1,790	477,732	2.90%
	Parkinsons' disease	1,513	478,009	11.30%

	Migraine	16,055	463,467	79.30%
	Alcohol dependency	6,403	473,119	5.30%
Digestive diseases	Gastro-oesophageal reflux disease (GORD)	41,649	437,873	35.60%
	Irritable bowel syndrome (IBS)	15,480	464,042	58.30%
	Inflammatory bowel disease (IBD)	6,430	473,092	17.50%
	Gallstone	21,006	458,516	21.00%
	Peptic ulcer	13,193	466,329	33.20%
	Liver cirrhosis	2,028	477,494	8.40%
	Appendicitis	8,075	471,447	47.50%
Genito-urinary diseases	Chronic kidney diseases	5,789	473,733	6.00%
	Benign prostatic hyperplasia (BPH)	16,874	202,602	23.20%
	Uterine fibroid	18,719	241,327	31.00%
Musculoskeletal diseases	Gout	8,661	470,861	56.20%
	Rheumatoid arthritis	7,981	471,541	34.60%
	Osteoarthritis	73,466	406,056	30.90%
	Osteoporosis	12,926	466,596	39.20%
	Sciatica	6,999	472,523	59.80%
	Intervertebral disc disorder - prolapsed disc / degenerative disc	17,010	462,512	40.70%
	Hip fracture	2,743	476,779	6.70%
Immune diseases	Sarcoidosis	1,273	478,249	44.70%
	Vasculitis	6,562	472,960	15.50%

Respiratory diseases	Chronic obstructive pulmonary disease (COPD)	16,549	462,973	32.10%
	Asthma	62,486	417,036	50.50%
	Lower respiratory infection / pneumonia	24,279	455,243	21.00%
	Otitis media	2,052	477,470	0.00%
	Allergy	9,780	469,742	100%
	Hayfever /eczema	116,566	362,956	100%
Eye diseases	Glaucoma	8,135	471,387	34.70%
	Cataract	27,288	452,234	11.30%
Cancer	Cancer overall	74,454	405,068	14.80%
	Lung cancer	2,493	477,029	5.40%
	Colorectal cancer	5,635	473,887	7.30%
	Female breast cancer	15,102	244,944	10.40%
	Prostate cancer	8,346	211,130	3.20%
	Melanoma	5,225	474,297	32.50%
	Skin cancer (including melanoma)	24,689	454,833	12.70%
	Cervical cancer	2,392	257,654	58.40%
	Uterus cancer	2,155	257,891	19.50%
	Ovary cancer	1,565	258,481	18.10%
	Kidney cancer	1,459	478,063	8.10%
	Bladder cancer	1,830	477,692	31.20%
	Non-Hodgkin lymphoma	2,243	477,279	5.50%

	Lymphomas and multiple myeloma	3,408	476,114	6.20%
	Leukaemia	1,313	478,209	12.80%
Risk factors	Obese BMI>30	116,336	361,244	NA

%Self-reported are cases whose diagnosis is based on reported history at recruitment without supporting information in hospital records, surgical records or registry.

Table 3. Phenotypes showing significant* associations with the *PCSK9* T allele at rs11591147 in UK Biobank.

Data type	Phenotype	Estimate (95% CI)	P value
Binary	Lipid-lowering medication†	0.616 (0.584 , 0.650)	$5.12x10^{-72}$
	Blood pressure medication	1.179 (1.125 , 1.235)	$5.49x10^{-12}$
	Type 2 Diabetes Mellitus	1.283 (1.187 , 1.386)	$3.26x10^{-10}$
	Diabetes type I	1.407 (1.169 , 1.693)	$3.07x10^{-4}$
	Insulin	1.336 (1.141 , 1.566)	$3.29x10^{-4}$
	Coronary artery diseases (CAD)	0.845 (0.770 , 0.928)	$4.23x10^{-4}$
	Obese BMI>30	1.072 (1.030 , 1.116)	$6.22x10^{-4}$
Continuous	Waist circumference	0.506 (0.322 , 0.691)	$7.39x10^{-8}$
	Waist hip ratio	0.003 (0.002 , 0.004)	$8.12x10^{-9}$
	Diastolic blood pressure	0.330 (0.157 , 0.503)	$1.80x10^{-4}$
	Body Mass Index	0.139 (0.065 , 0.212)	$2.34x10^{-4}$

*Significance based on a Bonferroni adjusted p value of $<6.25x10^{-4}$. Model results are from a mixed model adjusting for age, sex, BiLEVE chip, principle components 1-5 and lipid-lowering medication as fixed effects and family structure as a random effect. †This model did not further adjust for lipid-lowering medication. Diastolic blood pressure was adjusted for antihypertensive medication as described by Tobin et al.[26] The associations for quantitative traits are shown as beta coefficients and for categorical disorders as odds ratios. The full analysis with all traits and disorders is shown in **Supplementary Table S3.**

Table 4. Phenotypes showing significant* associations with at least one of the *HMGCR* variants associated with lower LDL-cholesterol in UK Biobank

Data Type	Trait	Lead HMGCR variant	Estimate	P value	Joint Estimate	Joint P value
Binary	Lipid-lowering medication†	rs12916	0.923 (0.912 , 0.934)	$1.34x10^{-37}$		
	Obese BMI>30	rs12916	1.045 (1.034 , 1.057)	$1.62x10^{-16}$	1.013 (0.999 , 1.027)	0.068
	Diabetes type II	rs12916	1.080 (1.059 , 1.102)	$3.60x10^{-14}$	*lead*	
	Blood pressure medication	rs12916	1.033 (1.021 , 1.046)	$2.12x10^{-7}$	2.306 (1.246 , 4.266)	$7.78x10^{-3}$
	Stroke	rs12916	1.057 (1.029 , 1.086)	$6.75x10^{-5}$	*lead*	
	Hypertension	rs17238484	1.023 (1.011 , 1.034)	$7.30x10^{-5}$	1.016 (1.004 , 1.028)	0.011
Continuous	Body Mass Index	rs12916	0.115 (0.096 , 0.135)	$1.41x10^{-31}$	0.010 (-0.028 , 0.048)	0.604
	Waist circumference	rs17238484	0.303 (0.247 , 0.360)	$3.46x10^{-26}$	-0.632 (-2.640 , 1.376)	0.537
	Waist hip ratio	rs17238484	0.001 (0.001 , 0.001)	$2.85x10^{-11}$	0.000 (0.000 , 0.001)	0.754
	Systolic blood pressure	rs17238484	0.170 (0.079 , 0.260)	$2.33x10^{-4}$	0.037 (-0.097 , 0.171)	0.588

Results for the two HMGCR variants analysed were very similar (**Supplementary Tables S4 and S5**) and only result for the SNP showing the stronger association is shown. *Significance based on a Bonferroni adjusted p value of $<6.25 \times 10^{-4}$. Model results are from a mixed model adjusting for age, sex, BiLEVE chip, principle components 1-5 and lipid-lowering medication as fixed effects and family structure as a random effect. †This model did not further adjust for lipid-lowering medication. Systolic blood pressure was adjusted for antihypertensive medication as described by Tobin et al.[26] The associations for quantitative traits are shown as beta coefficients and for categorical disorders as odds ratios. The full analyses with all traits and disorders are shown in **Supplementary Tables S4 & S5**. Joint estimate is the association of the lead variant after adjustment for other common SNPs in the region not in high LD with the lead SNP showing a stronger association with the trait.