

# **PCSK9 genetic variants and risk of vascular and non-vascular diseases in Chinese and UK populations**

Michael V Holmes<sup>1,2,\*</sup>, Christiana Kartsonaki<sup>1,2,\*</sup>, Ruth Boxall<sup>1,2,\*</sup>, Kuang Lin<sup>2</sup>, Nicola Reeve<sup>3</sup>, Canqing Yu<sup>4,5</sup>, Jun Lv<sup>4,5</sup>, Derrick A Bennett<sup>1,2</sup>, Michael R Hill<sup>1,2</sup>, Ling Yang<sup>1,2</sup>, Yiping Chen<sup>1,2</sup>, Huaidong Du<sup>1,2</sup>, Iain Turnbull<sup>2</sup>, Rory Collins<sup>2</sup>, Robert J Clarke<sup>2</sup>, Martin D Tobin<sup>3,6</sup>, Liming Li<sup>4,5</sup>, Iona Y Millwood<sup>1,2,#</sup>, Zhengming Chen<sup>1,2,#</sup>, Robin G Walters<sup>1,2,#</sup>, on behalf of the China Kadoorie Biobank Collaborative Group

\* *Contributed equally*

# *Joint senior authors*

<sup>1</sup> Medical Research Council Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>2</sup> Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, Oxford, UK

<sup>3</sup> Department of Population Health Sciences, University of Leicester, Leicester, UK

<sup>4</sup> Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, Beijing, China

<sup>5</sup> Peking University Center for Public Health and Epidemic Preparedness and Response, Beijing, China

<sup>6</sup> National Institute for Health and Care Research, Leicester Biomedical Research Centre, University of Leicester, Leicester, UK

## **Address for correspondence:**

Associate Professor Robin Walters  
Nuffield Department of Population Health  
University of Oxford  
Old Road Campus  
Roosevelt Drive  
Oxford OX3 7LF

robin.walters@ndph.ox.ac.uk

Professor Zhengming Chen  
Nuffield Department of Population Health  
University of Oxford  
Old Road Campus  
Roosevelt Drive  
Oxford OX3 7LF

zhengming.chen@ndph.ox.ac.uk

## Abstract

**Background:** Lowering low-density lipoprotein cholesterol (LDL-C) through PCSK9 inhibition represents a new therapeutic approach to preventing and treating cardiovascular disease (CVD). Phenome-wide analyses of *PCSK9* genetic variants in large prospective biobanks can help to identify unexpected effects of PCSK9 inhibition.

**Methods:** The prospective China Kadoorie Biobank genotyped >100,000 participants with follow-up for fatal and non-fatal disease events. A *PCSK9* genetic score (*PCSK9*-GS) was constructed using three single nucleotide polymorphisms at the *PCSK9* locus associated with directly-measured LDL-C, including a loss-of-function variant (rs151193009). Logistic regression gave estimated odds ratios (ORs) for *PCSK9*-GS associations with CVD and non-CVD outcomes, scaled to 1-standard deviation (SD) lower LDL-C.

**Results:** *PCSK9*-GS was associated with lower apolipoprotein B, and with lower risks of carotid plaque (n=8340 cases; OR=0.61 [95%CI: 0.45-0.83]; P=0.0015), major occlusive vascular events (n=15,752; 0.80 [0.67-0.95]; P=0.011), and ischaemic stroke (n=11,467; 0.80 [0.66-0.98]; P=0.029). However, *PCSK9*-GS was also associated with higher risk of hospitalisation with chronic obstructive pulmonary disease (COPD: n=6836; 1.38 [1.08-1.76]; P=0.0089), and with even higher risk of fatal exacerbations among individuals with pre-existing COPD (n=730; 3.61 [1.71-7.60]; P=7.3x10<sup>-4</sup>). Risk of acute upper respiratory tract infection (URTI) was also increased (n=1,095; 2.18 [1.34-3.53]; P=0.0016), as previously reported in UK Biobank, with a pooled OR after meta-analysis of 1.87 ([1.38-2.54]; P=5.4x10<sup>-5</sup>). We also replicated a previously-reported association with self-reported asthma (pooled OR 1.17 ([1.04-1.30]; P=0.0071). There was heterogeneity between the associations of *PCSK9*-GS and a polygenic LDL-C score for each of COPD hospitalisation (P-het=0.015), fatal COPD exacerbation (P-het=0.0012), and URTI (P-het=0.013).

**Conclusions:** LDL-C-lowering *PCSK9* genetic variants are associated with lower risk of subclinical and clinical atherosclerotic vascular disease, but higher risks of respiratory diseases which appear unrelated to LDL-C. Pharmacovigilance studies may be required to monitor patients treated with

therapeutic PCSK9 inhibitors for exacerbations of respiratory diseases or respiratory tract infections.

### **Keywords**

PCSK9 inhibition, upper respiratory tract infection, chronic obstructive pulmonary disease, exacerbation, Mendelian randomisation, target-mediated adverse effects

## Background

Lowering of low-density lipoprotein cholesterol (LDL-C) is an established, efficacious approach for treatment and prevention of occlusive cardiovascular disease (CVD)<sup>1</sup>. While statins represent the main class of LDL-C lowering drug used in routine clinical practice, new drugs targeting different components of LDL-C metabolism have gained clinical traction, including inhibition of Niemann-Pick C1-Like 1 (i.e. ezetimibe<sup>2</sup>) and of proprotein convertase subtilisin kinase 9 (PCSK9)<sup>3, 4</sup>.

Compared with statins, PCSK9 inhibitors (e.g. evolocumab, inclisiran) are relatively new, with much less comprehensive long-term data concerning their benefits and potential side-effects. Two large randomised trials have reported that inhibition of PCSK9 over a period of around 2 to 3 years significantly reduced the risk of major CVD events in high-risk patients with prior history of vascular disease<sup>3, 4</sup>. Moreover, these trials have demonstrated a reduction in atheroma volume<sup>5</sup> and that the reduction in CVD events was proportional to the reduction in LDL-C achieved<sup>6</sup>. However, there was suggestive evidence of excess risk of diabetes associated with PCSK9 inhibition<sup>7</sup>. As with statins<sup>8</sup>, confirmation (or refutation) of any excess risk of diabetes will likely require meta-analysis of multiple large trials with patients treated and followed up for longer durations. Little is known about the long-term effects of PCSK9 inhibition on a wide range of other major disease outcomes.

Human genetics can be used to predict the likely effects of therapeutic modification of a drug target on disease outcomes<sup>9-11</sup>. Several well-established examples have either predicted or recapitulated findings from randomised clinical trials of lipid lowering (such as HMGCR<sup>12</sup> and CETP<sup>13</sup> inhibitors) and other therapies (including inhibitors of interleukin 6 receptor<sup>14, 15</sup>, secretory phospholipase A2-IIA<sup>16</sup>, and lipoprotein-associated phospholipase A2<sup>17, 18</sup>). This approach has provided evidence corroborating findings from randomised clinical trials that PCSK9 inhibition leads to reduced LDL-C<sup>19</sup>, with a corresponding reduction in risk of CVD<sup>20, 21</sup>. Similarly, genetic studies provide strong evidence that, as with statins, LDL-C lowering by PCSK9 inhibition may increase the risk of diabetes<sup>20, 22-24</sup>. In addition to diabetes, analyses in UK Biobank (UKB) of a functional variant of *PCSK9* (rs11591147) suggested possible excess risks of several non-vascular disease outcomes,

including asthma and respiratory tract infections<sup>22,25</sup>, although these associations were only nominally significant and require replication.

In this study we sought to further investigate the potential impact of therapeutic PCSK9 inhibition. In approximately 100,000 genotyped individuals from the prospective China Kadoorie Biobank (CKB), we assessed the associations of a *PCSK9* genetic score (*PCSK9*-GS) with lipids, lipoproteins and a range of vascular and non-vascular disease outcomes, including (where available) those previously reported in UKB. Where applicable, we further undertook meta-analyses of CKB and UKB data for respiratory disease and other relevant disease outcomes.

## Methods

### *Study design and population*

CKB is a prospective cohort study of 512,713 adults aged 30-79 years, recruited between 2004 and 2008 from 5 urban and 5 rural areas across China<sup>26</sup>. In brief, baseline information was collected via a laptop-based questionnaire (including demographic and lifestyle factors and medical history) and physical measurements (including anthropometry, blood pressure, and lung function). Duplicate lung function measurements were conducted using a portable handheld 'Micro spirometer' (Micro Medical Limited, Rochester, Kent, England), as previously described<sup>27</sup>, from which forced expiratory volume in 1s (FEV1) and forced vital capacity (FVC) were derived; the distributions of values for FEV1/FVC, with large numbers of individuals with values of 1, indicated systematic errors in data collection in two recruitment regions (Qingdao and Haikou, see **Supplementary Figure 1**), so participants in these regions were excluded from analyses requiring spirometry data. A non-fasting blood sample was collected (with time since last meal recorded) and separated into plasma and buffy-coat fractions for long-term storage. Resurveys of random 5% subsets of the cohort were conducted at periodic intervals, and the second resurvey in 2013-2014 included measurements of carotid intima media thickness (CIMT) and plaque<sup>28</sup>.

Incident disease outcomes were identified from long-term follow up through electronic linkage of each participant's unique national identification number to the Chinese national health insurance system, and to established regional registries for death and major diseases (cancer, ischaemic

heart disease, stroke, and diabetes). Health insurance records included detailed information about each hospital admission (e.g. disease description, International Statistical Classification of Diseases and Related Health Problems, 10th Revision [ICD-10] code, and procedure or examination codes). All reported cases of disease outcomes from different sources were centrally checked, reviewed and standardised by clinicians.

#### *Biomarker assays*

Plasma concentrations of total cholesterol, LDL-C, high-density lipoprotein cholesterol (HDL-C), triglycerides, apolipoprotein B (ApoB), and apolipoprotein A-I from baseline samples were quantified in 18,181 CKB participants using clinical chemistry assays at the Wolfson Laboratory (AU 680 clinical chemistry analysers, Beckman Coulter [UK] Ltd., Wycombe, United Kingdom) using manufacturers' reagents, calibrators, and settings<sup>29</sup>. In a subset of 4,442 individuals, a high-throughput targeted <sup>1</sup>H-NMR metabolomics platform<sup>30, 31</sup> was used to generate spectra from which 225 lipid and other metabolic measures were simultaneously quantified by Nightingale Health Ltd. (Helsinki, Finland; previously known as Brainshake Ltd).

#### *GWAS genotyping*

Genome-wide genotyping data were available for a subset of 100,706 CKB participants, comprising ~30,000 participants selected for nested case-control studies of CVD and respiratory disease, and ~70,000 being randomly-selected from the remaining participants<sup>32</sup>. Region-specific principal component analysis identified 6,107 individuals with ancestry not local to the region in which they were recruited (of whom 774 had clinical biochemistry measurements), who were excluded from region-stratified analyses. To avoid biases due to over-representation of disease cases in the genotyped dataset, we constructed a subset of 70,914 individuals representative of the full CKB cohort in which such over-representation was eliminated; this was a preliminary version of the population-representative sample described elsewhere<sup>32</sup>.

### *Instrument selection and derivation of weighted PCSK9 locus score*

In 17,687 samples with both genotyping and LDL-C data, LDL-C was regressed on age, age<sup>2</sup>, sex, study area, fasting time, and fasting time<sup>2</sup>. The residuals were rank inverse-normal transformed (RINT) and genome wide association analysis of the transformed variable was conducted using BOLT-LMM v2.3.2<sup>33</sup>. Summary statistics for variants within a 1Mbp window around the *PCSK9* structural gene were evaluated using FINEMAP v1.1<sup>34</sup>, identifying the model with the highest posterior probability of explaining association at the locus with LDL-C, comprising a set of three single nucleotide polymorphisms (SNPs; rs151193009, rs2495477, rs11206517: log<sub>10</sub> Bayes Factor=47.6). These were used to construct a weighted genetic score, with SNP weights corresponding to the beta coefficients for association of the SNPs with RINT LDL-C. To account for linkage disequilibrium between the SNPs, the per-allele beta coefficients were derived from a region-stratified, multivariable (mutually adjusted) model including dosages for the 3 SNPs, age, age<sup>2</sup>, sex, study area, fasting time, fasting time<sup>2</sup>, and region-specific principal components, with exclusion of 774 individuals with non-local ancestry (**Supplementary Table 1**), which were applied to individuals without LDL-C data. To avoid potential bias from using internally-derived weights, beta coefficients for participants with LDL-C data were instead derived using 100-fold block-jack-knifing, from regressions that excluded random 1% subsets of participants to whom those beta coefficients were assigned<sup>35</sup>. For each individual with genotyping data, the sum of SNP dosages, weighted by their corresponding beta coefficients, gave a *PCSK9* gene score calibrated to the predicted effect on LDL-C.

### *Derivation of weighted LDL-C genetic score*

An LDL-C genetic risk score was derived using SNPs that were previously identified as independently-associated with LDL-C (at  $P < 5 \times 10^{-8}$ ) in GWAS of Europeans in the Global Lipids Genetics Consortium (GLGC)<sup>36</sup>. Of 76 independent association signals, 2 SNPs were monomorphic in CKB; 64 of the remaining 74 SNPs showed an association with LDL-C in CKB that was consistent with that in GLGC (i.e. directionally consistent or no significant effect size heterogeneity at  $P < 0.05/74$ ). These 64 SNPs were used to construct an LDL-C GS in CKB (further

described in the **Supplementary Methods** and **Supplementary Table 2**), allowing a comparison of disease associations arising from LDL-C lowering by *PCSK9* variants with those from LDL-C lowering overall. This LDL-C score did not include variants at the *PCSK9* locus.

### *Disease endpoints*

The detailed vascular and non-vascular disease outcomes used in the present study and their corresponding ICD-10 codes are provided in **Supplementary Table 3**. Vascular disease outcomes included major coronary events (MCE: non-fatal myocardial infarction, fatal ischaemic heart disease, or coronary revascularisation); fatal/nonfatal ischaemic stroke; fatal/nonfatal intracerebral haemorrhage; major occlusive vascular events (MOVE: consisting of fatal/nonfatal major coronary events or ischaemic stroke); fatal CVD; and major vascular events (MVE: fatal/nonfatal myocardial infarction, coronary revascularisation procedures, stroke, or fatal CVD). These analyses used a common set of controls that excluded all individuals who self-reported prior CHD, stroke or transient ischaemic attack at baseline, or who experienced any form of MVE during the follow-up period.

The main non-vascular outcomes were organised by anatomical site and consisted of: combined incident, self-reported, and screen-detected cases of diabetes; and incident events of each of chronic obstructive pulmonary disease (COPD); chronic kidney disease; chronic liver disease; eye disease; malignancy; and non-vascular mortality (see **Supplementary Table 3**). Controls for these diseases excluded individuals with self-reported history of that disease at baseline (where available). Phenome-wide analysis was conducted in the population subset using incident disease outcomes in the ICD-10 code range A00 to N99 grouped together (defined in **Supplementary Table 4**), with no exclusions for prevalent disease for either cases or controls.

To dissect the principal incident COPD outcomes, which comprised a mixture of incident and recurrent disease, we defined prevalent COPD cases at baseline as being those with FEV1/FVC less than the lower limit of normal as predicted from the Global Lung Function 2012 spirometry reference equations<sup>37</sup> according to participants' ancestry, age, height, and sex (**Supplementary Figure 2**). Amongst those with spirometry-defined prevalent COPD at baseline (where available),



we further defined those experiencing one or more incident COPD outcomes as exacerbations of COPD. Moreover, we further analysed COPD outcomes according to disease severity, according to whether incident COPD cases or cases of COPD exacerbation were fatal (COPD identified as cause of death) or non-fatal. These analyses excluded all participants in the 2 recruitment regions with systematic errors in spirometry data collection. Analyses of upper respiratory tract infections (URTI) excluded cases occurring before 2009, to avoid biases due to a spike in reported cases (>10% of all CKB cases across all years) in Zhejiang in 2007-2008.

### *Statistical analysis*

For continuous variables measured in the CVD nested case-control subset, such as blood biochemistry and NMR metabolomics, data transformation and analyses were stratified by recruitment region: following linear regression of each variable on age, age<sup>2</sup>, sex, case ascertainment category, and up to nine region-specific genetic principal components, the residuals underwent stratum-specific RINT. For CIMT, which was measured at the second resurvey in a randomly-selected subset of surviving participants, adjustments and data transformation were as above except without adjustment for case ascertainment. Associations of the weighted *PCSK9*-GS with these continuous traits were assessed by linear regression within each region with inverse-variance-weighted fixed-effect (IVW-FE) meta-analysis of the resulting estimates. SNP associations with spirometry were extracted from previous GWAS in CKB of FEV1/FVC<sup>38</sup> with IVW-FE meta-analysis of the resulting estimates. LDL-C variance explained and F-statistics for the *PCSK9* and LDL-C genetic scores were determined by comparison of linear models including or excluding the scores, using the R `anova()` function.

Disease outcome associations with *PCSK9*-GS were assessed by logistic regression within each region with adjustment for sex, age, age<sup>2</sup> and up to nine region-specific genetic principal components, and estimates were combined using IVW-FE meta-analysis. For all disease outcomes, cases and controls were selected from the population-representative subset of 70,914 genotyped individuals, plus additional cases identified among the remaining genotyped individuals. However, to avoid potential ascertainment bias, additional cases were excluded from the analysis if

they were selected for genotyping on the basis of a different disease outcome that might separately be associated with altered *PCSK9* expression (see **Supplementary Table 5**).

For confirmation of the therapeutic effects of inhibition of *PCSK9* in reducing risk of CVD outcomes<sup>4, 39</sup>, we assessed a combination of simple and composite outcomes, with a Bonferroni-adjusted threshold of significance ( $P < 0.05/3 = 0.0167$ ) based on 3 independent tests (determined by spectral decomposition of the correlation matrix for these related endpoints<sup>40</sup>). Similarly, we used a Bonferroni-adjusted threshold of significance ( $P < 0.05/(7 + 3 \text{ prior tests}) = 0.005$ ) for non-vascular disease associations (type 2 diabetes, COPD, chronic kidney disease, chronic liver disease, malignant neoplasms, eye disease, and non-vascular mortality), and also for the phenome-wide scan ( $P < 0.05/(41 + 10 \text{ prior tests}) = 9.8 \times 10^{-4}$ ). For sensitivity analyses of subgroups, analyses were performed in the full dataset with additional adjustment for recruitment region.

For disease outcomes previously reported as nominally associated with a functional *PCSK9* variant in a recent study of UKB participants by Rao et al<sup>22</sup>, associations of *PCSK9*-GS with seven (out of 12) diseases that were available in CKB were investigated as follows. We used identical case/control definitions (or the closest possible match), based on ICD-10 codes or responses to the baseline questionnaire (**Supplementary Table 6**); for cerebrovascular disease, the analysis was limited strictly to the population-representative subset – this was to minimise risk of bias due to over- or under-representation of particular disease subtypes, since the case-control part of the genotyped dataset is not representative of the full spectrum of cerebrovascular disease.

Replication of these prior associations was assessed at 5% false discovery rate (FDR, Benjamini-Hochberg). Heterogeneity was assessed using Cochran's Q-statistic with a Bonferroni-adjusted significance threshold ( $P < 0.05/7$ ).

In order to compare effect estimates, these were scaled to the same difference in LDL-C (1-SD lower) as follows: the UKB study by Rao et al.<sup>22</sup> (and Nelson et al.<sup>25</sup>) used a missense variant in *PCSK9* (rs11591147) to quantify the associations with risk of diseases. We took the per-allele rs11591147 LDL-C estimate from the Global Lipids Genetics Consortium<sup>41</sup> (per-allele beta coefficient: 0.497 SD) and scaled the disease association estimates reported by Rao et al<sup>22</sup> so that

they were equivalent to 1-SD lower LDL-C, as used for the CKB effect estimates. The scaled estimates derived from UKB and CKB were combined in IVW-FE meta-analyses, and Cochran's Q statistic was used to assess heterogeneity.

All analyses used release version 15 of the CKB database and were performed using SAS software (version 9.3; SAS Institute, Inc) or R v4.2.1.

## Results

### *Participant characteristics*

In the population-representative subset of 70,914 genotyped participants, the mean age at study baseline was 52.1 (SD 10.7), 59.7% were female, and 45.8% were from urban regions. Stratifying individuals according to whether they were carriers of a *PCSK9* functional variant (see below), no appreciable differences between these two groups were observed at baseline for a wide range of CVD risk factors (**Table 1**), including height, adiposity, blood pressure, physical activity, alcohol drinking, and smoking. Similar proportions of the two groups had prior diagnoses of CVD, hypertension, or diabetes, or were taking statins.

### *Characterisation of the PCSK9 genetic score*

In 17,687 participants with both genotyping and LDL-C data, three SNPs in/around *PCSK9* (rs151193009, rs2495477 and rs11206517) were independently associated with LDL-C at genome-wide significance ( $P < 5 \times 10^{-8}$ ) in a multivariable model, with per-allele effects on LDL-C of 0.65 SD (SE=0.05), 0.10 SD (SE=0.01), and 0.16 SD (SE=0.03), respectively (**Supplementary Table 1**). The SNP with the strongest effect, rs151193009C>T, is a known loss-of-function variant that disrupts binding of the PCSK9 protein prodomain to heparan sulphate proteoglycans (HSPGs), an essential step in PCSK9-mediated degradation of LDL receptors (LDLR)<sup>42</sup>. When combined into the weighted *PCSK9*-GS, the three SNPs explained 1.2% of the variance of LDL-C (F-statistic=206), and showed a consistent association with LDL-C between women and men, and across geographical regions, age groups, and smoking or alcohol behaviour categories (**Supplementary Figure 3**). Scaled to a 1-SD lowering of LDL-C (equivalent to 27.4 mg/dl), *PCSK9*-GS was associated with lower concentrations of ApoB (-0.95 SDs, 95%CI: -1.08, -0.82;  $P=6.7 \times 10^{-46}$ ) and, more weakly, a lowering of Lp(a) (-0.14 [-0.27, -0.01];  $P=0.038$ ). No significant associations were detected with triglycerides (-0.01 [-0.14, 0.12];  $P=0.87$ ) or HDL-C (-0.10 [-0.23, 0.03];  $P=0.13$ ) (**Figure 1**). Using detailed measures of lipoproteins and lipids quantified by NMR-spectroscopy, we identified associations of *PCSK9*-GS with lower concentrations of cholesterol among the smaller subclasses of very low-density lipoproteins and across intermediate- and low-

density lipoprotein subclasses (**Supplementary Figure 4**). After accounting for multiple testing, there was no significant association with inflammation biomarkers including glycoprotein acetyls, C-reactive protein, or fibrinogen, nor were associations identified with other non-lipid blood biochemistry measures (**Supplementary Figures 4-5**).

#### *Association of PCSK9 with subclinical and clinical vascular outcomes*

As expected, *PCSK9*-GS was associated with clinical CVD outcomes. Scaled to 1 SD lower LDL-C, there was reduced risk of MOVE which was significant after adjustment for multiple testing (n=15,752; OR=0.80 [0.67, 0.95]; P=0.011) and nominally significant reduced risk of ischaemic stroke (n=11,467; 0.80 [0.66, 0.98]; P=0.029); for MCE, MVE, and fatal CVD, the associations were directionally consistent but not statistically significant (**Figure 2A**). In addition, the score was associated with lower CIMT (n=20,896; -0.24 SDs [-0.36, -0.12]; P=0.00012) and 37% lower odds of carotid plaque (n=8340 cases; 0.61 [0.45, 0.83]; P=0.0015) (**Supplementary Table 7**). Conversely, no association of *PCSK9*-GS was observed with risk of intracerebral haemorrhage (n=5,906; 1.04 [0.81, 1.34]; P=0.74).

This pattern of associations with risk of vascular disease was very similar to that obtained when using a genetic score for LDL-C (**Figure 2B**) constructed with 64 SNPs from across the genome but which did not include SNPs at the *PCSK9* locus (variance explained=5.6%, F-statistic=1002; **Supplementary Table 2**). For example, using the LDL-C genetic score, 1 SD lower LDL-C was associated with lower risk of MCE (0.74 [0.66, 0.84]; P=1.5x10<sup>-6</sup>) and ischaemic stroke (0.90 [0.83, 0.98]; P=0.015), but not intracerebral haemorrhage (1.02 [0.91, 1.14]; P=0.73).

#### *Association of genetic scores with major non-vascular outcomes*

We further assessed the associations of the *PCSK9* and LDL-C genetic scores with major non-vascular outcomes. No associations were identified for *PCSK9*-GS with diabetes, chronic kidney diseases, eye diseases, malignant neoplasms, or non-vascular mortality (**Figure 2A**). There was a nominally significant positive association of higher LDL-C with risk of liver disease (2.10 [1.17, 3.79]; P=0.013), and a strongly suggestive positive association with risk of incident COPD (1.38 [1.08, 1.76]; P=8.9x10<sup>-3</sup>), although these were not significant after accounting for multiple testing.

The LDL-C genetic score showed no association with any of these outcomes (**Figure 2B**).

Association of *PCSK9*-GS was also investigated for a wide range of other disease outcomes available in CKB using a phenome-wide approach (**Supplementary Figure 6**), with no association meeting the pre-specified significance threshold.

#### *Replication of reported PCSK9 associations*

For seven traits previously reported as being nominally-associated with a missense *PCSK9* variant (rs11591147) in UKB<sup>22</sup>, we tested association with *PCSK9*-GS in CKB (**Figure 3**). Scaled to a 1-SD lowering of LDL-C, association in CKB of *PCSK9*-GS with higher risk of URTI (n=1,095 cases; 2.18 [1.34, 3.35; P=1.6x10<sup>-3</sup>) was significant at 5% FDR and was consistent with the association observed in UKB (n=2,364 cases; 1.70 [1.15, 2.51]), providing an estimate after meta-analysis of 1.87 [1.38, 2.54] (P=5.4x10<sup>-5</sup>). A nominally significant association was also identified for self-reported doctor-diagnosed asthma (n=427 cases; 2.28 [1.12, 4.64]; P=0.02), again consistent with the association reported in UK Biobank (n=39,269 cases; 1.15 [1.02, 1.28]), with an estimate after meta-analysis of OR=1.17 [1.04, 1.30] (P=0.007). Although low case numbers limited power for replication in CKB, effect size estimates were consistent with those from UKB for the other disease outcomes, with the exception of breast cancer (P-het = 0.005).

#### *Association of PCSK9 genetic score with further respiratory endpoints*

The association of *PCSK9*-GS with multiple respiratory-related diseases, in both UKB and CKB, was further investigated in related outcomes in CKB (**Figure 4**). The primary incident COPD outcome, reflecting either death or hospital admission due to COPD, was consistently observed across fatal (n=1632; 1.47 [0.93, 2.34]; P=0.098) or non-fatal (n=5204; 1.38 [1.05, 1.80], P=0.019) events. However, there was no association of the genetic score with prevalent COPD, as defined by spirometry measurements at baseline (n=5,105; 0.96 [0.73, 1.26]; P=0.75), nor with baseline FEV1/FVC as a continuous measure (**Supplementary Table 8**). By contrast, the association with higher risk of an incident COPD event was stronger in those with pre-existing COPD (2212 exacerbations among 6,409 individuals with spirometry-defined COPD; 1.91 [1.12, 3.24]; P=0.017). The increased risk was most clearly observed in those for whom the COPD exacerbation was fatal

(730 deaths due to acute exacerbations of COPD: 3.61 [1.71, 7.60];  $P=7.1 \times 10^{-4}$ ), although this was not significantly different from risk of a non-fatal exacerbation (1,482 non-fatal exacerbations: 1.58 [0.86, 2.91];  $P=0.14$ ;  $P\text{-het}=0.24$ ). The association of *PCSK9*-GS with URTI persisted after excluding individuals with spirometry-defined COPD and no differences were found across geographical regions, age groups, sex, smoking status, and alcohol consumption (**Supplementary Figure 7**).

The association with risk of COPD exacerbation in CKB was not replicated in UKB using a different functional *PCSK9* variant (**Supplementary Table 9**), although power was limited due to a much smaller number of recorded cases (468 in UKB vs 2212 cases in CKB), and it is unclear to what extent the aetiology or definition of acute exacerbation of COPD differed between the two biobanks which derive from very different populations with contrasting exposures to other risk factors for respiratory disease<sup>43</sup>. After meta-analysis, the pooled OR was 1.45 [0.91, 2.32] ( $P=0.12$ ), but with significant heterogeneity between the 2 estimates ( $P\text{-het} = 0.032$ ).

By contrast with *PCSK9*-GS, the LDL-C GS shows no association with any of these respiratory disease outcomes (**Figure 4**). When scaled to the same predicted difference in LDL-C, there was significant heterogeneity between associations with *PCSK9*-GS and LDL-C GS for both URTI (*PCSK9*: 2.18 [1.34, 3.53]; LDL-C GS: 1.11 [0.88, 1.38];  $P\text{-het}=0.013$ ) and fatal COPD exacerbations (*PCSK9*: 3.61 [1.71, 7.60]; LDL-C GS: 0.92 [0.75, 1.19];  $P\text{-het}=0.0012$ ).

## Discussion

We used genetic variants, including a loss-of-function variant, to mimic pharmacological inhibition of *PCSK9* to gain insight into expected therapeutic effects on a broad range of disease outcomes in a Chinese population. As expected, genetic inhibition of *PCSK9* was associated with lower levels of LDL-C and ApoB, less subclinical atherosclerosis, lower risk of MOVE, including ischaemic stroke, and a directionally concordant effect on risk of MCE. While these associations are confirmatory of previously reported results in European ancestry populations, as the relationship of *PCSK9* inhibition with risk of vascular disease is well documented from large-scale cardiovascular outcome trials<sup>4, 39</sup> and other genetic studies<sup>20, 22, 44, 45</sup>, they demonstrate that *PCSK9*

inhibition is likely also to be effective for the treatment of CVD in East Asian populations, in whom mean LDL-C levels are lower than in Western populations. Conversely, the present study also indicated that PCSK9 inhibition is potentially associated with higher risks of URTI, acute exacerbation of COPD, and asthma.

MR studies in Europeans have shown that a 1-SD reduction in LDL-C (equivalent to ~1 mmol/L) leads to a reduction in risk of CHD of approximately 40%<sup>46</sup>. This compares to a more modest association in CKB for a comparable outcome of around 25% per 1-SD lower LDL-C, although this is at least partly attributable to a smaller SD for LDL-C in CKB (0.69 mmol/L) than in European populations. When scaled to the same mmol/L reduction in LDL-C as in Europeans, the reduction in risk in CKB is 34%, which suggests consistency of effect across ancestral groups once appropriate comparisons are made<sup>1</sup>. A pre-specified secondary analysis of the FOURIER trial identified that treatment with a PCSK9 inhibitor led to a monotonic relationship between LDL-C lowering and risk of vascular events<sup>6</sup>. Similarly, the GLAGOV trial illustrated a dose response relationship between LDL-C lowering and volume of atherosclerosis<sup>5</sup>. Thus, although the population studied here had much lower LDL-C levels at baseline as compared to European studies (e.g. 2.35 (0.69 SD) mmol/L CKB vs 3.57 (0.87) UKB<sup>46</sup>), triangulation of our genetic findings with studies in Europeans and RCTs of LDL-C-lowering therapies, including PCSK9 inhibitors, supports the clinical efficacy of PCSK9 inhibition on vascular disease outcomes in patients of East Asian ancestry. Moreover, these findings provide further support for intensive LDL-C lowering therapy in high risk individuals (e.g. adding PCSK-9 inhibitor to statin) in order to achieve maximal beneficial effects.

In addition to confirming the expected associations with various forms of occlusive CVD, we further identified that *PCSK9*-GS, orientated to a lowering of LDL-C, was associated with a higher risk of URTI, acute exacerbations of COPD, and self-reported asthma. For URTI and asthma, our study findings replicate nominally significant findings in a previous phenome scan of another functional variant in *PCSK9* conducted within the UKB<sup>22,25</sup>. In addition to these three respiratory disease outcomes, meta-analysis of our results with those from UKB strengthened evidence for an association of *PCSK9* genetic variants with mood disorders and with non-deforming dorsopathies



(diseases of the spine). Previous studies have identified association of *PCSK9* with severe forms of dorsopathy<sup>47</sup> which we were unable to investigate in CKB, a population-based study of middle-aged adults.

The findings of the present study linking PCSK9 inhibition with lung infections have biological plausibility. The mechanism by which PCSK9 influences LDL-C is by mediating lysosomal degradation of LDLR, so that inhibition of PCSK9 leads to higher levels of LDLR in the liver<sup>48</sup> and a concomitant reduction in blood LDL-C levels. However, PCSK9 binds to and mediates degradation of not only LDLR but also very low-density lipoprotein receptors (VLDLR)<sup>49, 50</sup>. PCSK9, LDLR and VLDLR are each also expressed in the lung<sup>51</sup>, where VLDLR and LDLR are involved in the initiation of infection by several classes of respiratory viruses (e.g. human rhinovirus, some types of coronavirus), which bind to VLDLR and LDLR on the cell surface, and are internalised via receptor-mediated endocytosis<sup>52, 53</sup>. Inhibition of the LDL receptor in cultured human tracheal epithelial cells lowers the infectivity of human rhinovirus<sup>54</sup>. Thus, it is possible that the relationship of genetic variants mimicking therapeutic inhibition of PCSK9 with a higher risk of URTI (which can underlie acute exacerbations of COPD<sup>55</sup>, and worsen asthma symptoms<sup>56</sup>) may arise from elevated levels of VLDLR or LDLR in lung tissue. Since the proposed mechanism is via higher susceptibility to respiratory viruses, leading to an increase in exacerbations of COPD (which are often infectious in nature), we would not expect an association with prevalent COPD, as observed – the pathophysiology of COPD development is itself separate, a potential example of disease incidence vs progression being related to discrete aetiological pathways.

A review article in 2013 suggested that dual PCSK9 and statin therapy might amplify potential safety issues, including those related to respiratory infections, given that both act to increase levels of LDLR and VLDLR<sup>57</sup>. What is unclear is whether PCSK9 inhibitors in clinical use (or under development) alter pulmonary VLDL receptor levels and, thus, whether genetic findings reflecting life-long systemic changes in PCSK9 have clinical relevance to patients receiving long-term treatment with tissue-localised PCSK9 inhibitors. While various parameters determine the tissue distribution of monoclonal antibodies<sup>58</sup>, current evidence suggests that evolocumab has limited tissue distribution<sup>59, 60</sup>. Furthermore, inclisiran, a small interfering RNA inhibitor of *PCSK9* that is

currently under development<sup>61</sup>, features N-acetylgalactosamine conjugation that renders it hepatotropic<sup>62</sup>. Consequently, limited tissue penetration of biological agents may mean that the genetic associations we identify do not have major clinical relevance.

The available evidence indicates that neither of two large phase III trials led to any excess risk of respiratory disease in individuals treated with PCSK9 inhibition for a mean ~2 to 3 years<sup>4, 39</sup>.

However, the association of *PCSK9*-GS with exacerbations of COPD may indicate that PCSK9 inhibition may confer higher risk only in individuals with existing respiratory disease, perhaps triggered by URTI. Clinical trials may not have included many such participants, thus limiting their power to detect any excess risk within the short period of trial treatment. While some phase II trials did show a non-significant excess risk of URTI in those treated with evolocumab<sup>63-65</sup>, such flu-like symptoms may be common following treatment with monoclonal antibodies in general<sup>66</sup> and, as such, it is not clear whether these findings reported in RCTs represent a reaction to the treatment modality as opposed to PCSK9 inhibition itself.

Despite the consistency of associations of genetic variants in *PCSK9* and risk of lung diseases in these two large biobanks, and the underlying biological plausibility, there are several other factors to consider in interpreting these findings. First, the genetic findings may be a false positive. For example, we failed, albeit based on a small number of cases and using a different functional *PCSK9* variant, to replicate the association observed in CKB with a similar COPD exacerbation endpoint in UKB. In the HUNT biobank in Norway, no association of *PCSK9* genetic variants with risks of asthma (6858 cases) or COPD (6685 cases) was identified, although there was a nominal association with rhinitis<sup>67</sup>; in that study, the association with URTI was not reported and it is also unclear whether COPD was prevalent or represented acute exacerbations of disease. Second, direct comparisons between different populations are complicated by functional differences between the genetic variants available in different ancestries. The variant with the strongest effect in the CKB *PCSK9* score modifies the PCSK9 HSPG-binding domain, directly impacting a key step in LDL receptor degradation<sup>42</sup>. By contrast, the UK Biobank functional variant (rs11591147) instead exerts its effect by modulating local protein structure<sup>68</sup>. While both lead to lower LDL-C, suggesting overall similarities in their biological consequences, their relative impact on differing PCSK9-

mediated pathways may vary. Third, the association of *PCSK9*-GS with URTI might arise due to the life-long effects of genetic perturbations that alter human biology<sup>69</sup>, meaning that treatment with a PCSK9 inhibitor at a specific time (e.g. in later life) need not lead to a similarly altered risk of lung disease. Fourth, as discussed above, limited tissue penetration of biological therapeutics may mean that VLDLR and LDLR levels in lung parenchyma are unaltered. Fifth, even if the genetic associations with respiratory disease foretell findings in humans from inhibition of PCSK9 that have yet to be identified, the relevance of the magnitude of these effects and how to frame these in terms of absolute risk are unclear. While there is a strong basis for calibration of genetic estimates for vascular disease risk of LDL-C lowering variants to the equivalent from a therapeutic trial<sup>70</sup>, no such calibration exists for non-vascular disease outcomes such as respiratory disease, which poses a translational challenge.

This study demonstrates the advantages of leveraging large-scale genetic data from prospective biobanks in diverse populations to inform drug target validation, repurposing, and design and conduct of randomised trials. Apart from providing improvements in statistical power, a key advantage is in accessing the considerable heterogeneity in the genetic and environmental characteristics of different populations, to inform the utility and application of therapies in varying contexts globally. Further, our study used functional variants that are present at appreciable frequency only in one ancestry, which, when findings are consistent across different functional variants in different ancestral groups with very different environmental exposures, reinforces the credibility of the data. However, despite this, our findings remain hypothesis generating and require further validation in future studies involving much larger numbers of well-characterised disease outcomes.

## **Conclusion**

This study provides evidence that genetic variants in *PCSK9* that lower LDL-C and risk of CVD are also associated with higher risks of respiratory disease, in particular acute URTI and acute exacerbations of COPD. While further genetic and clinical evidence is needed to confirm (or refute) these findings, they raise the possibility that, in individuals with pre-existing COPD, a careful risk-

benefit evaluation of LDL-C lowering through inhibition of PCSK9 may be required.

Pharmacovigilance may be warranted to ascertain whether these findings have relevance to patients being treated with PCSK9 inhibitors.

## Declarations

### *Ethics approval*

Ethical approval for CKB was obtained jointly from the University of Oxford, the Chinese Centre for Disease Control and Prevention (CCDC) and the regional CCDC from the 10 study areas. All participants provided written informed consent.

### *Data availability*

The China Kadoorie Biobank (CKB) is a global resource for the investigation of lifestyle, environmental, blood biochemical and genetic factors as determinants of common diseases. The CKB study group is committed to making the cohort data available to the scientific community in China, the UK and worldwide to advance knowledge about the causes, prevention and treatment of disease. For detailed information on what data is currently available to open access users and how to apply for it, visit: <https://www.ckbiobank.org/data-access>.

Researchers who are interested in obtaining raw data from the China Kadoorie Biobank study that underlies this paper should contact [ckbaccess@ndph.ox.ac.uk](mailto:ckbaccess@ndph.ox.ac.uk). A research proposal will be requested to ensure that any analysis is performed by bona fide researchers and - where data is not currently available to open access researchers - is restricted to the topic covered in this paper. Sharing of genotyping data is currently constrained by the Administrative Regulations on Human Genetic Resources of the People's Republic of China. Access to these and certain other data is available only through collaboration with CKB researchers.

### *Author contributions*

Data collection and analysis – MVH, CK, RB, KL, NR, CY, JL, DAB, MRH, LY, YC, HD, IT, MDT, IYM, RGW; funding – ZC, LL, RGW, IYM, RJC, MRH, RC; study oversight: MVH, RGW; manuscript draft and revision – MVH, RGW, IYM, ZC. All authors read and approved the final manuscript.

### *Funding*

The CKB baseline survey and the first re-survey were supported by the Kadoorie Charitable Foundation in Hong Kong. Long-term follow-up was supported by the Wellcome Trust (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z), the National Natural Science Foundation of China (82192900, 81941018, 91843302, 91846303), and the National Key Research and Development Program of China (2016YFC0900500, 2016YFC0900501, 2016YFC0900504, 2016YFC1303904). DNA extraction and genotyping was funded by GlaxoSmithKline, and the UK Medical Research Council (MC-PC-13049, MC-PC-14135). The project is supported by core funding from the UK Medical Research Council (MC\_UU\_00017/1, MC\_UU\_12026/2, MC\_U137686851), Cancer Research UK (C16077/A29186; C500/A16896), and the British Heart Foundation (CH/1996001/9454) to the Clinical Trial Service Unit and Epidemiological Studies Unit and to the MRC Population Health Research Unit at Oxford University. MVH was supported by a British Heart Foundation Intermediate Clinical Research Fellowship (FS/18/23/33512) and the National Institute for Health Research Oxford Biomedical Research Centre. MDT was partially supported through the National Institute for Health and Care Research (NIHR, Leicester Biomedical Research Centre and NIHR Senior Investigator Award). Views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. MDT was also supported via Wellcome Trust Awards WT202849/Z/16/Z and WT225221/Z/22/Z. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### *Acknowledgements*

The most important acknowledgement is to the participants in the study and the members of the survey teams in each of the 10 regional centres. We also thank the project development and management teams based at Beijing, Oxford, and the 10 regional centres. China's National Health Insurance provides electronic linkage to all hospital treatments.

### *Competing interests*

The Clinical Trial Service Unit and Epidemiological Studies Unit, University of Oxford (CTSU) is coordinating a PCSK9 phase III cardiovascular outcome trial (ORION 4). MVH has collaborated with Boehringer Ingelheim in research, and in accordance with the policy of CTSU did not accept any personal payment. MVH was at the University of Oxford when this work was conducted, and is presently employed by 23andMe and holds stock in 23andMe.

### *License*

This research was funded in whole, or in part, by the Wellcome Trust [212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z]. For the purpose of Open Access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

## References

1. Collins R, Reith C, Emberson J, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet* 2016; **388**: 2532-61.
2. Cannon CP, Blazing MA, Giugliano RP, et al. Ezetimibe Added to Statin Therapy after Acute Coronary Syndromes. *N Engl J Med* 2015; **372**: 2387-97.
3. Sabatine MS, Giugliano RP, Keech AC, et al. Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. *N Engl J Med* 2017; **376**: 1713-22.
4. Schwartz GG, Steg PG, Szarek M, et al. Alirocumab and Cardiovascular Outcomes after Acute Coronary Syndrome. *N Engl J Med* 2018.
5. Nicholls SJ, Puri R, Anderson T, et al. Effect of Evolocumab on Progression of Coronary Disease in Statin-Treated Patients: The GLAGOV Randomized Clinical Trial. *JAMA* 2016; **316**: 2373-84.
6. Giugliano RP, Pedersen TR, Park JG, et al. Clinical efficacy and safety of achieving very low LDL-cholesterol concentrations with the PCSK9 inhibitor evolocumab: a prespecified secondary analysis of the FOURIER trial. *Lancet* 2017; **390**: 1962-71.
7. de Carvalho LSF, Campos AM, Sposito AC. Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) Inhibitors and Incident Type 2 Diabetes: A Systematic Review and Meta-analysis With Over 96,000 Patient-Years. *Diabetes Care* 2018; **41**: 364-7.
8. Sattar N, Preiss D, Murray HM, et al. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. *Lancet* 2010; **375**: 735-42.
9. Holmes MV, Ala-Korpela M, Davey Smith G. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat Rev Cardiol* 2017; **14**: 577-90.
10. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018; **362**: k601.
11. Holmes MV. Human genetics and drug development. *N Engl J Med* 2019; **380**: 1076-9.
12. 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* 2015; **385**: 351-61.



13. Millwood IY, Bennett DA, Holmes MV, et al. Association of CETP gene variants with risk for vascular and nonvascular diseases among Chinese adults. *JAMA Cardiol* 2018; **1**: 34-43.
14. Bovijn J, Lindgren CM, Holmes MV. Genetic variants mimicking therapeutic inhibition of IL-6 receptor signaling and risk of COVID-19. *Lancet Rheumatol* 2020; **2**: e658-e9.
15. Group RC. Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet* 2021; **397**: 1637-45.
16. Holmes MV, Simon T, Exeter HJ, et al. Secretory Phospholipase A2-IIA and Cardiovascular Disease. *Journal of the American College of Cardiology* 2013; **62**: 1966-76.
17. Millwood IY, Bennett DA, Walters RG, et al. A phenome-wide association study of a lipoprotein-associated phospholipase A2 loss-of-function variant in 90 000 Chinese adults. *International Journal of Epidemiology* 2016; **45**: 1588-99.
18. Millwood IY, Bennett DA, Walters RG, et al. Lipoprotein-Associated Phospholipase A2 Loss-of-Function Variant and Risk of Vascular Diseases in 90,000 Chinese Adults. *J Am Coll Cardiol* 2016; **67**: 230-1.
19. Sliz E, Kettunen J, Holmes MV, et al. Metabolomic Consequences of Genetic Inhibition of PCSK9 Compared with Statin Treatment. *Circulation* 2018; **138**: 2499-512.
20. Ference BA, Robinson, J. G., Brook, R. D., Catapano AL, Chapman MJ, et al. Variation in PCSK9 and HMGCR and risk of cardiovascular disease and diabetes. *New England Journal of Medicine* 2016; **375**: 2144-53.
21. Schmidt AF, Holmes MV, Preiss D, et al. Phenome-wide association analysis of LDL-cholesterol lowering genetic variants in PCSK9. *BMC Cardiovasc Disord* 2019; **19**: 240.
22. Rao AS, Lindholm D, Rivas MA, Knowles JW, Montgomery SB, Ingelsson E. Large-Scale Phenome-Wide Association Study of PCSK9 Variants Demonstrates Protection Against Ischemic Stroke. *Circ Genom Precis Med* 2018; **11**: e002162.
23. Schmidt AF, Swerdlow DI, Holmes MV, et al. PCSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol* 2017; **5**: 97-105.
24. Lotta LA, Sharp SJ, Burgess S, et al. Association Between Low-Density Lipoprotein Cholesterol-Lowering Genetic Variants and Risk of Type 2 Diabetes: A Meta-analysis. *JAMA* 2016; **316**: 1383-91.

25. Nelson CP, Lai FY, Nath M, et al. Genetic Assessment of Potential Long-Term On-Target Side Effects of PCSK9 (Proprotein Convertase Subtilisin/Kexin Type 9) Inhibitors. *Circ Genom Precis Med* 2019; **12**: e002196.
26. Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* 2011; **40**: 1652-66.
27. Kurmi OP, Li L, Smith M, et al. Regional variations in the prevalence and misdiagnosis of air flow obstruction in China: baseline results from a prospective cohort of the China Kadoorie Biobank (CKB). *BMJ Open Respiratory Research* 2014; **1**: e000025.
28. Clarke R, Du H, Kurmi O, et al. Burden of carotid artery atherosclerosis in Chinese adults: Implications for future risk of cardiovascular diseases. *European journal of preventive cardiology* 2017; **24**: 647-56.
29. Sun L, Clarke R, Bennett D, et al. Causal associations of blood lipids with risk of ischemic stroke and intracerebral hemorrhage in Chinese adults. *Nat Med* 2019; **25**: 569-74.
30. Soininen P, Kangas AJ, Wurtz P, et al. High-throughput serum NMR metabolomics for cost-effective holistic studies on systemic metabolism. *Analyst* 2009; **134**: 1781-5.
31. Soininen P, Kangas AJ, Wurtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ Cardiovasc Genet* 2015; **8**: 192-206.
32. Walters RG, Millwood IY, Lin K, et al. Genotyping and population characteristics of the China Kadoorie Biobank. *Cell Genomics* 2023; **3**: 100361.
33. Loh PR, Kichaev G, Gazal S, Schoech AP, Price AL. Mixed-model association for biobank-scale datasets. *Nat Genet* 2018; **50**: 906-8.
34. Benner C, Spencer CC, Havulinna AS, Salomaa V, Ripatti S, Pirinen M. FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* 2016; **32**: 1493-501.
35. Hartwig FP, Davies NM. Why internal weights should be avoided (not only) in MR-Egger regression. *Int J Epidemiol* 2016; **45**: 1676-8.
36. Do R, Willer CJ, Schmidt EM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet* 2013; **45**: 1345-52.

37. Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; **40**: 1324-43.
38. Shrine N, Izquierdo AG, Chen J, et al. Multi-ancestry genome-wide association analyses improve resolution of genes and pathways influencing lung function and chronic obstructive pulmonary disease risk. *Nature Genetics* 2023; **55**: 410-22.
39. Sabatine MS, Giugliano, R. P., Keech, A. C., Honarpour N, Wiviott SD, et al. Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. *New England Journal of Medicine* 2017; **376**: 1713-22.
40. Machado AMC. Multiple Testing Correction in Medical Image Analysis. *Journal of Mathematical Imaging and Vision* 2007; **29**: 107-17.
41. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013; **45**: 1274-83.
42. Gustafsen C, Olsen D, Vilstrup J, et al. Heparan sulfate proteoglycans present PCSK9 to the LDL receptor. *Nature Communications* 2017; **8**: 503.
43. Chan KH, Kurmi OP, Bennett DA, et al. Solid Fuel Use and Risks of Respiratory Diseases. A Cohort Study of 280,000 Chinese Never-Smokers. *American journal of respiratory and critical care medicine* 2019; **199**: 352-61.
44. Cohen JC, Boerwinkle E, Mosley TH, Jr., Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* 2006; **354**: 1264-72.
45. Ference BA, Yoo, W., Alesh, I., Mahajan N, Mirowska KK, et al. Effect of Long-Term Exposure to Lower Low-Density Lipoprotein Cholesterol Beginning Early in Life on the Risk of Coronary Heart Disease A Mendelian Randomization Analysis. *Journal of the American College of Cardiology* 2012; **60**: 2631-9.
46. Richardson TG, Sanderson E, Palmer TM, et al. Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS Med* 2020; **17**: e1003062.
47. Jerome RN, Pulley JM, Roden DM, et al. Using Human 'Experiments of Nature' to Predict Drug Safety Issues: An Example with PCSK9 Inhibitors. *Drug Saf* 2018; **41**: 303-11.

48. Grefhorst A, McNutt MC, Lagace TA, Horton JD. Plasma PCSK9 preferentially reduces liver LDL receptors in mice. *J Lipid Res* 2008; **49**: 1303-11.
49. Poirier S, Mayer G, Benjannet S, et al. The proprotein convertase PCSK9 induces the degradation of low density lipoprotein receptor (LDLR) and its closest family members VLDLR and ApoER2. *J Biol Chem* 2008; **283**: 2363-72.
50. Shan L, Pang L, Zhang R, Murgolo NJ, Lan H, Hedrick JA. PCSK9 binds to multiple receptors and can be functionally inhibited by an EGF-A peptide. *Biochem Biophys Res Commun* 2008; **375**: 69-73.
51. Uhlen M, Fagerberg L, Hallstrom BM, et al. Proteomics. Tissue-based map of the human proteome. *Science* 2015; **347**: 1260419.
52. Rankl C, Kienberger F, Wildling L, et al. Multiple receptors involved in human rhinovirus attachment to live cells. *Proc Natl Acad Sci U S A* 2008; **105**: 17778-83.
53. Hofer F, Gruenberger M, Kowalski H, et al. Members of the low density lipoprotein receptor family mediate cell entry of a minor-group common cold virus. *Proc Natl Acad Sci U S A* 1994; **91**: 1839-42.
54. Suzuki T, Yamaya M, Kamanaka M, et al. Type 2 rhinovirus infection of cultured human tracheal epithelial cells: role of LDL receptor. *Am J Physiol Lung Cell Mol Physiol* 2001; **280**: L409-20.
55. Sapey E, Stockley RA. COPD exacerbations . 2: aetiology. *Thorax* 2006; **61**: 250-8.
56. Busse WW, Lemanske RF, Jr., Gern JE. Role of viral respiratory infections in asthma and asthma exacerbations. *Lancet* 2010; **376**: 826-34.
57. Poirier S, Mayer G. The biology of PCSK9 from the endoplasmic reticulum to lysosomes: new and emerging therapeutics to control low-density lipoprotein cholesterol. *Drug Des Devel Ther* 2013; **7**: 1135-48.
58. Tabrizi M, Bornstein GG, Suria H. Biodistribution mechanisms of therapeutic monoclonal antibodies in health and disease. *AAPS J* 2010; **12**: 33-43.
59. Henry CA, Lyon RA, Ling H. Clinical efficacy and safety of evolocumab for low-density lipoprotein cholesterol reduction. *Vasc Health Risk Manag* 2016; **12**: 163-9.

60. Kasichayanula S, Grover A, Emery MG, et al. Clinical Pharmacokinetics and Pharmacodynamics of Evolocumab, a PCSK9 Inhibitor. *Clinical Pharmacokinetics* 2018; **57**: 769-79.
61. Mullard A. Nine paths to PCSK9 inhibition. *Nature Reviews Drug Discovery* 2017; **16**: 299-301.
62. Nair JK, Willoughby JL, Chan A, et al. Multivalent N-acetylgalactosamine-conjugated siRNA localizes in hepatocytes and elicits robust RNAi-mediated gene silencing. *Journal of the American Chemical Society* 2014; **136**: 16958-61.
63. Blom DJ, Hala T, Bolognese M, et al. A 52-week placebo-controlled trial of evolocumab in hyperlipidemia. *N Engl J Med* 2014; **370**: 1809-19.
64. Sabatine MS, Giugliano RP, Wiviott SD, et al. Efficacy and safety of evolocumab in reducing lipids and cardiovascular events. *N Engl J Med* 2015; **372**: 1500-9.
65. Raal FJ, Honarpour N, Blom DJ, et al. Inhibition of PCSK9 with evolocumab in homozygous familial hypercholesterolaemia (TESLA Part B): a randomised, double-blind, placebo-controlled trial. *Lancet* 2015; **385**: 341-50.
66. Baldo BA. Adverse events to monoclonal antibodies used for cancer therapy: Focus on hypersensitivity responses. *Oncoimmunology* 2013; **2**: e26333.
67. Brumpton BM, Fritsche LG, Zheng J, et al. Variation in Serum PCSK9 (Proprotein Convertase Subtilisin/Kexin Type 9), Cardiovascular Disease Risk, and an Investigation of Potential Unanticipated Effects of PCSK9 Inhibition. *Circ Genom Precis Med* 2019; **12**: e002335.
68. Sarkar SK, Foo ACY, Matyas A, et al. A transient amphipathic helix in the prodomain of PCSK9 facilitates binding to low-density lipoprotein particles. *Journal of Biological Chemistry* 2020; **295**: 2285-98.
69. Holmes MV A-KM, Davey Smith G. Mendelian randomization in cardiometabolic disease: challenges in interpretations of causality. *Nature Reviews Cardiology* 2017; **In Press**.
70. Ference BA. How to use Mendelian randomization to anticipate the results of randomized trials. *Eur Heart J* 2018; **39**: 360-2.

## Figure legends

### **Figure 1: Associations of *PCSK9* genetic score with major blood lipids, and apolipoproteins.**

Estimates are standardized beta coefficients from linear regressions of the weighted *PCSK9* genetic score, scaled to a 1-SD lowering of LDL-cholesterol, calculated as an inverse-variance-weighted average of region-specific estimates with adjustment for age, age<sup>2</sup>, sex, ascertainment, and region-specific principal components.

### **Figure 2: Associations of genetic scores for (A) *PCSK9* and (B) LDL cholesterol with risk of vascular and non-vascular disease endpoints.**

Odds ratios are estimated using logistic regression of the weighted *PCSK9* genetic score, scaled to a 1-SD lowering of LDL cholesterol, calculated as an inverse-variance-weighted average of region-specific estimates with adjustment for age, age<sup>2</sup>, sex, ascertainment, and region-specific principal components.

### **Figure 3. Replication in CKB of previously reported associations in UKB of *PCSK9* genetic scores with disease outcomes.**

Estimates in UKB (open boxes) originate from data published by Rao et al<sup>22</sup>, and scaled to a corresponding 1-SD lowering of LDL cholesterol. Odds ratios in CKB (filled boxes) are estimated using logistic regression of the weighted *PCSK9* genetic score, scaled to a 1-SD lowering of LDL cholesterol, calculated as an inverse-variance-weighted average of region-specific estimates with adjustment for age, age<sup>2</sup>, sex, ascertainment, and region-specific principal components. Trans-ancestry estimates (diamonds) are derived from inverse-variance-weighted fixed-effect meta-analysis.

**Figure 4. Associations of *PCSK9* and LDL-C genetic scores with respiratory disease endpoints.**

Odds ratios are estimated using logistic regression of the weighted *PCSK9* genetic score, scaled to a 1-SD lowering of LDL cholesterol, calculated as an inverse-variance-weighted average of region-specific estimates with adjustment for age, age<sup>2</sup>, sex, ascertainment, and region-specific principal components. Prevalent COPD defined on basis of spirometry (see Methods). COPD exacerbations are incident events occurring in individuals with spirometry-defined prevalent COPD. P<sub>het</sub> is the p-value from a test for heterogeneity for estimates of *PCSK9* and LDL-C genetic instruments.

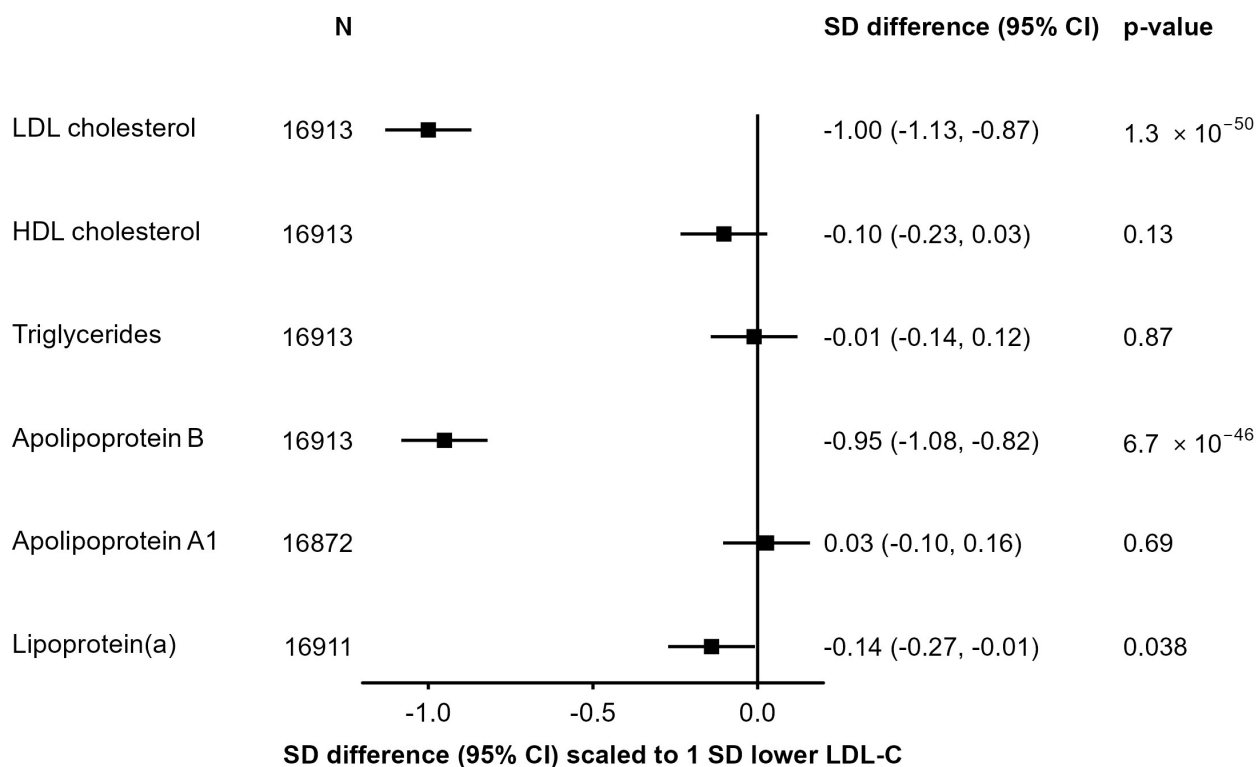
**Table 1:** Baseline characteristics of the study population, dichotomised as non-carriers/carriers of the loss-of-function *PCSK9* variant rs151193009-T

Characteristics	rs151193009 genotype	
	CC (N=69,088)	CT/TT (N=1,826)
Age (years), mean (SD)	52.1 (10.7)	52.5 (10.9)
Female, %	59.7	61.2
Urban region, %	45.8	47.2
LDL-C (mmol/L), mean (SD)	2.36 (0.69)	1.99 (0.69)
Height (cm), mean (SD)	158.8 (8.3)	159.1 (8.2)
BMI (kg/m <sup>2</sup> ), mean (SD)	23.8 (3.4)	24.0 (3.5)
Waist/hip ratio, mean (SD)	0.88 (0.07)	0.88 (0.07)
SBP (mmHg), mean (SD)	131.5 (21.4)	132.5 (21.8)
DBP (mmHg), mean (SD)	78.0 (11.2)	78.8 (11.5)
Physical activity (MET h/d), mean (SD)	20.6 (14.0)	20.6 (14.3)
High school or higher education, %	21.0	21.8
Household income ≥20,000 yuan, %	40.1	40.4
Current regular alcohol drinking, %	15.5	15.5
Current regular smoker, %	25.9	24.2
Stroke/TIA <sup>a</sup> , %	1.7	1.7
Coronary heart disease <sup>a</sup> , %	3.6	4.2
Hypertension, %	11.6	12.9
Diabetes <sup>b</sup> , %	6.1	6.2
Taking statins, %	1.9	1.6

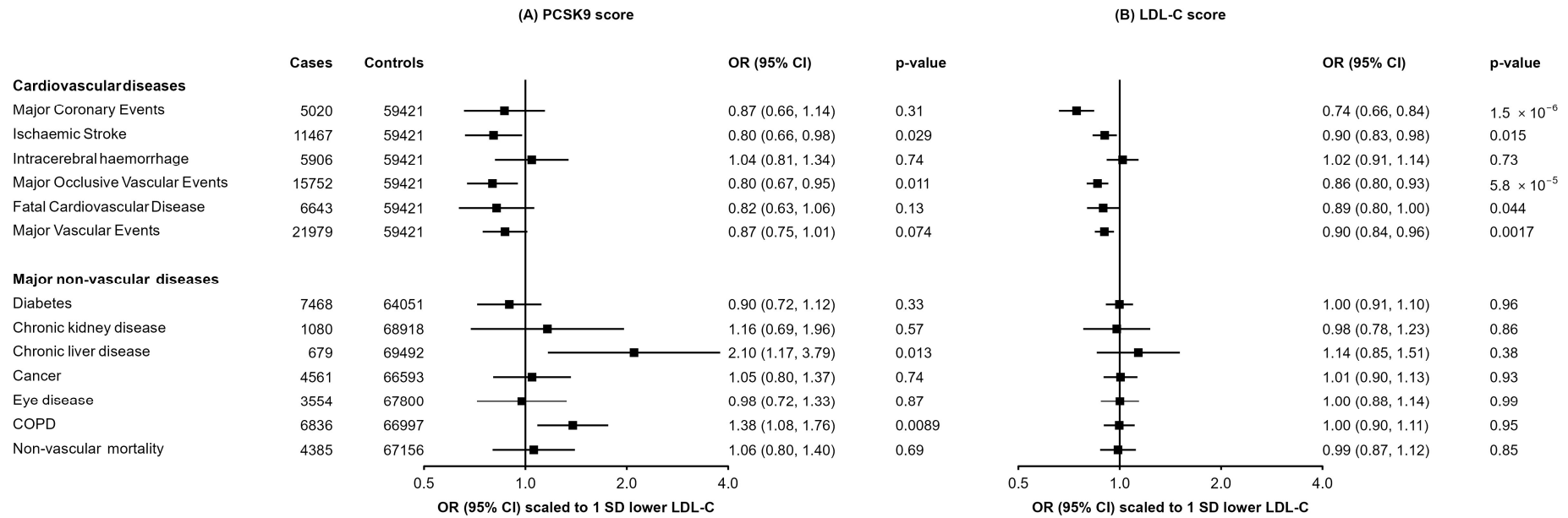
BMI: body mass index; SBP: systolic blood pressure; DBP: diastolic blood pressure; MET: metabolic equivalent of task  
<sup>a</sup>self-reported physician diagnosed. <sup>b</sup>self-reported physician diagnosed or screen-detected at baseline



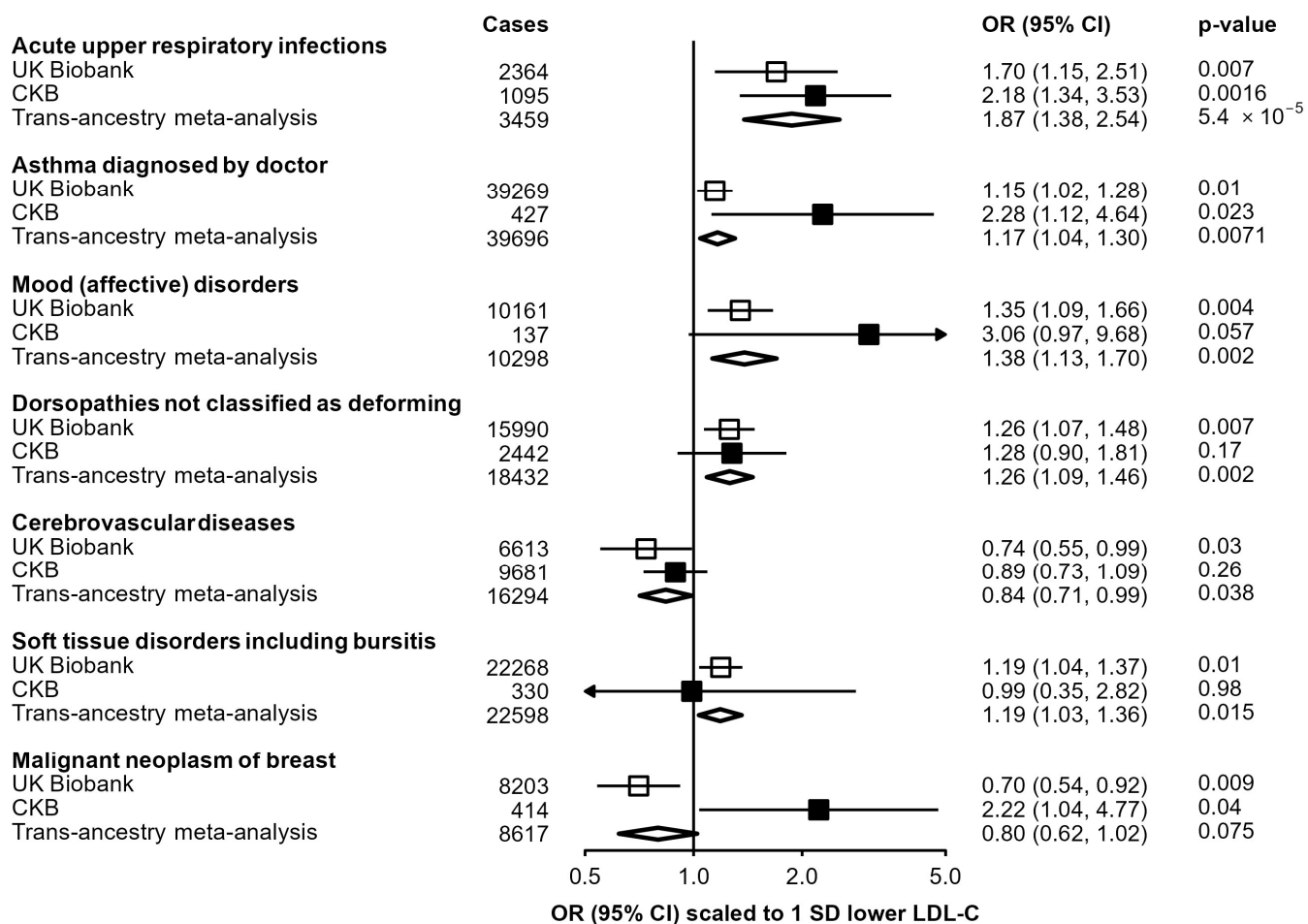
**Figure 1. Associations of the PCSK9 genetic score with major blood lipids and apolipoproteins in CKB.**



**Figure 2:** Associations of genetic scores for (A) *PCSK9* and (B) LDL cholesterol with risk of vascular and non-vascular disease endpoints.



**Figure 4. Replication by CKB of previously reported associations in UK Biobank of polygenic scores with disease outcomes.**



**Figure 4 Associations of PCSK9 and LDL-C genetic scores with respiratory disease endpoints**

