

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/168896/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Holmes, Michael V., Kartsonaki, Christiana, Boxall, Ruth, Lin, Kuang, Reeve, Nicola, Yu, Canqing, Lv, Jun, Bennett, Derrick A., Hill, Michael R., Yang, Ling, Chen, Yiping, Du, Huaidong, Turnbull, Iain, Collins, Rory, Clarke, Robert J., Tobin, Martin D., Li, Liming, Millwood, Iona Y., Chen, Zhengming, Walters, Robin G., Chen, Junshi, Chen, Zhengming, Clarke, Robert, Collins, Rory, Guo, Yu, Li, Liming, Wang, Chen, Lv, Jun, Peto, Richard, Walters, Robin, Avery, Daniel, Bennett, Derrick, Boxall, Ruth, Burgess, Sushila, Chan, Ka Hung, Chen, Yiping, Chen, Zhengming, Clarke, Johnathan, Clarke, Robert, Du, Huaidong, Edris, Ahmed, Fry, Hannah, Gilbert, Simon, Hill, Mike, Im, Pek Kei, Iona, Andri, Kakkoura, Maria, Kartsonaki, Christiana, Lam, Hubert, Lin, Kuang, Mazidi, Mohsen, Millwood, Iona, Morris, Sam, Nie, Qunhua, Pozarickij, Alfred, Ryder, Paul, Said, Saredo, Schmidt, Dan, Sherliker, Paul, Stevens, Becky, Turnbull, Iain, Walters, Robin, Wang, Baihan, Wang, Lin, Wright, Neil, Yang, Ling, Yang, Xiaoming, Yao, Pang, Han, Xiao, Hou, Can, Xia, Qingmei, Liu, Chao, Lv, Jun, Pei, Pei, Yu, Canqing, Dong, Caixia, Ge, Pengfei, Ren, Xiaolan, Li, Zhongxiao, Mao, Enke, Wang, Tao, Zhang, Hui, Zhang, Xi, Chen, Jinyan, Hu, Ximin, Wang, Xiaohuan, Guo, Zhendong, Li, Huimei, Li, Yilei, Weng, Min, Wu, Shukuan, Yan, Shichun, Zou, Mingyuan, Zhou, Xue, Guo, Ziyang, Kang, Quan, Li, Yanjie, Yu, Bo, Xu, Qinai, Chang, Liang, Fan, Lei, Feng, Shixian, Zhang, Ding, Zhou, Gang, Gao, Yulian, He, Tianyou, He, Pan, Hu, Chen, Sun, Huarong, Zhang, Xukui, Chen, Biyun, Fu, Zhongxi, Huang, Yuelong, Liu, Huilin, Xu, Qiaohua, Yin, Li, Long, Huajun, Xu, Xin, Zhang, Hao, Zhang, Libo, Chen, Naying, Liu, Duo, Tang, Zhenzhu, Chen, Ningyu, Jiang, Qilian, Lan, Jian, Li, Mingqiang, Liu, Yun, Meng, Fanwen, Meng, Jinhuai, Pan, Rong, Qin, Yulu, Wang, Ping, Wang, Sisi, Wei, Liuping, Zhou, Liyuan, Cheng, Liang, Du, Ranran, Gao, Ruqin, Li, Feifei, Li, Shanpeng, Liu, Yongmei, Ning, Feng, Pang, Zengchang, Sun, Xiaohui, Tian, Xiaocao, Wang, Shaojie, Zhai, Yaoming, Zhang, Hua, Hou, Wei, Lv, Silu, Wang, Junzheng, Chen, Xiaofang, Wu, Xianping, Zhang, Ningmei, Zhou, Weiwei, Chen, Xiaofang, Li, Jianguo, Liu, Jiaqiu, Luo, Guojin, Sun, Qiang, Zhong, Xunfu, Su, Jian, Tao, Ran, Wu, Ming, Yang, Jie, Zhou, Jinyi, Zhou, Yonglin, Hu, Yihe, Hua, Yujie, Jin, Jianrong, Liu, Fang, Liu, Jingchao, Lu, Yan, Ma, Liangcai, Tang, Aiyu, Zhang, Jun, Gong, Weiwei, Hu, Ruying, Wang, Hao, Wang, Meng, Yu, Min, Chen, Lingli, Gu, Qijun, Pan, Dongxia, Wang, Chunmei, Xie, Kaixu and Zhang, Xiaoyi 2024. PCSK9 genetic variants and risk of vascular and non-vascular diseases in Chinese and UK populations. *European Journal of Preventive Cardiology* 31 (8), pp. 1015-1025. 10.1093/eurjpc/zwae009

Publishers page: <http://dx.doi.org/10.1093/eurjpc/zwae009>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of



this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

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

Michael V. Holmes^{1,2†}, Christiana Kartsonaki^{1,2†}, Ruth Boxall ^{1,2†}, Kuang Lin², Nicola Reeve³, Canqing Yu ^{4,5,6}, Jun Lv ^{4,5,6}, Derrick A. Bennett^{1,2}, Michael R. Hill^{1,2}, Ling Yang^{1,2}, Yiping Chen^{1,2}, Huaidong Du ^{1,2}, Iain Turnbull², Rory Collins ², Robert J. Clarke², Martin D. Tobin^{3,7}, Liming Li^{4,5,6}, Iona Y. Millwood^{1,2‡}, Zhengming Chen ^{1,2*‡}, and Robin G. Walters ^{1,2*‡},
on behalf of the China Kadoorie Biobank Collaborative Group

¹Medical Research Council Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, Old Road Campus, Roosevelt Drive, Oxford OX3 7LF, UK; ²Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, Old Road Campus, Roosevelt Drive, Oxford OX3 7LF, UK; ³Department of Population Health Sciences, University of Leicester, Leicester, UK; ⁴Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China; ⁵Peking University Center for Public Health and Epidemic Preparedness and Response, Peking University, Beijing, China; ⁶Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing, China; and ⁷National Institute for Health and Care Research, Leicester Biomedical Research Centre, University of Leicester, Leicester, UK

Received 25 August 2023; revised 15 December 2023; accepted 20 December 2023; online publish-ahead-of-print 10 January 2024

Aims

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 biobanks can help to identify unexpected effects of PCSK9 inhibition.

Methods and results

In the prospective China Kadoorie Biobank, we constructed a genetic score using three variants at the PCSK9 locus associated with directly measured LDL-C [PCSK9 genetic score (PCSK9-GS)]. Logistic regression gave estimated odds ratios (ORs) for PCSK9-GS associations with CVD and non-CVD outcomes, scaled to 1 SD lower LDL-C. PCSK9-GS was associated with lower risks of carotid plaque [$n = 8340$ cases; OR = 0.61 (95% confidence interval: 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 hospitalization 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 amongst individuals with pre-existing COPD [$n = 730$; 3.61 (1.71–7.60); $P = 7.3 \times 10^{-4}$]. We also replicated associations for a PCSK9 variant, reported in UK Biobank, with increased risks of acute upper respiratory tract infection (URTI) [pooled OR after meta-analysis of 1.87 (1.38–2.54); $P = 5.4 \times 10^{-5}$] and self-reported asthma [pooled OR of 1.17 (1.04–1.30); $P = 0.0071$]. There was no association of a polygenic LDL-C score with COPD hospitalization, COPD exacerbation, or URTI.

Conclusion

The LDL-C-lowering PCSK9 genetic variants are associated with lower risk of subclinical and clinical atherosclerotic vascular disease but higher risks of respiratory diseases. Pharmacovigilance studies may be required to monitor patients treated with therapeutic PCSK9 inhibitors for exacerbations of respiratory diseases or respiratory tract infections.

Lay summary

Genetic analyses of over 100 000 participants of the China Kadoorie Biobank, mimicking the effect of new drugs intended to reduce cholesterol by targeting the PCSK9 protein, have identified potential severe effects of lower PCSK9 activity in patients with existing respiratory disease.

* Corresponding authors. Tel: +44 1865 743977, Email: robin.walters@ndph.ox.ac.uk (R.G.W.); Tel: +44 1865 743839, Email: zhengming.chen@ndph.ox.ac.uk (Z.C.)

† The first three authors contributed equally to the study.

‡ The last three authors are joint senior authors.

© The Author(s) 2024. Published by Oxford University Press on behalf of the European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

- PCSK9 genetic variants that are associated with lower cholesterol and reduced rates of cardiovascular disease are also associated with increased risk of a range of respiratory diseases, including asthma, upper respiratory tract infections, and hospitalization with chronic obstructive pulmonary disease (COPD).
- These genetic variants are not associated with whether or not individuals have COPD; instead, they are specifically associated with an increase in the chance of those who already have COPD being hospitalized and even dying, suggesting that careful monitoring of such patients should be considered during development of and treatment with anti-PCSK9 medication.

Keywords

PCSK9 inhibition • Upper respiratory tract infection • Chronic obstructive pulmonary disease • Exacerbation • Mendelian randomization • Target-mediated adverse effects

Introduction

Lowering of low-density lipoprotein cholesterol (LDL-C) is an established, efficacious approach for treatment and prevention of occlusive cardiovascular disease (CVD).¹ Whilst 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. ezetimibe²) and of proprotein convertase subtilisin kinase 9 (PCSK9).^{3,4}

Compared with statins, PCSK9 inhibitors (e.g. evolocumab and inclisiran) are relatively new, with much less comprehensive long-term data concerning their benefits and potential side-effects. Two large randomized trials have reported that inhibition of PCSK9 over a period of around 2–3 years significantly reduced the risk of major CVD events in high-risk patients with prior history of vascular disease.^{3,4} Moreover, these trials have demonstrated a reduction in atheroma volume⁵ and that the reduction in CVD events was proportional to the reduction in LDL-C achieved.⁶ However, there was suggestive evidence of excess risk of diabetes associated with PCSK9 inhibition.⁷ As with statins,⁸ 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 potential long-term effects in the general population of PCSK9 inhibition on a wide range of other major disease outcomes, although open-label extension studies of trials of both evolocumab and inclisiran have identified very low rates of adverse events.^{9,10}

Human genetics can be used to predict the likely effects of therapeutic modification of a drug target on disease outcomes.^{11–13} Several well-established examples have either predicted or recapitulated findings from randomized clinical trials of lipid lowering (such as 3-Hydroxy-3-Methylglutaryl-CoA Reductase¹⁴ and Cholesteryl ester transfer protein (CETP)¹⁵ inhibitors) and other therapies (including inhibitors of interleukin 6 receptor,^{16,17} secretory phospholipase A2-IIA,¹⁸ and lipoprotein-associated phospholipase A2^{19,20}). This approach has provided evidence corroborating findings from randomized clinical trials that PCSK9 inhibition leads to reduced LDL-C,²¹ with a corresponding reduction in risk of CVD.^{22,23} Similarly, genetic studies provide strong evidence that, as with statins, LDL-C lowering by PCSK9 inhibition may increase the risk of diabetes.^{22,24–26} 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,^{24,27} 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

The overall study design is illustrated in [Supplementary material online, Figure S1](#). The CKB is a prospective cohort study of 512 713 adults aged 30–79 years, recruited between 2004 and 2008 from five urban and five rural areas across China.²⁸ 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, UK), as previously described,²⁹ from which forced expiratory volume in 1 s (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 material online, Figure S2](#)), 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 2013–14 included measurements of carotid intima-media thickness (CIMT) and plaque.³⁰

Incident disease outcomes up to 1 January 2017 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 standardized 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 (17 687 with genotyping data) using clinical chemistry assays at the Wolfson Laboratory [AU 680 clinical chemistry analysers, Beckman Coulter (UK) Ltd, Wycombe, UK] using manufacturers’ reagents, calibrators, and settings.³¹ In a subset of 4442 individuals, a high-throughput targeted ¹H-NMR metabolomic platform^{32,33} 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).

Genome-wide association studies 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.³⁴ Region-specific principal component analysis identified 6107 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 data set, 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.³⁴

Instrument selection and derivation of weighted PCSK9 locus score

In the 17 687 samples with both genotyping and LDL-C data, LDL-C was regressed on age, age-squared, sex, study area, fasting time, and fasting time-squared. The residuals were rank inverse-normal transformed (RINT) and genome-wide association analysis of the transformed variable was conducted using BOLT-LMM v2.3.2.³⁵ Summary statistics for variants within a 1 Mbp window around the PCSK9 structural gene were evaluated using FINEMAP v1.1,³⁶ 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, and rs11206517: \log_{10} Bayes Factor = 47.6). These were used to construct a weighted genetic score, PCSK9-GS, with SNP weights corresponding to the beta coefficients for the 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 three SNPs, age, age-squared, sex, study area, fasting time, fasting time-squared, and region-specific principal components, with exclusion of 774 individuals with non-local ancestry (see [Supplementary material online, Table S1](#)), which were applied to individuals without LDL-C data. The association of these SNPs with LDL-C was validated by performing lookups in summary statistics for East Asian populations from the Global Lipids Genetics Consortium (GLGC)³⁷ (see [Supplementary material online, Table S2](#)). To avoid potential bias from using internally derived weights and to provide internal cross-validation of the association of PCSK9-GS with LDL-C, beta coefficients for participants with LDL-C data were derived using 100-fold block jackknifing, from regressions that excluded random 1% subsets of participants to whom those beta coefficients were assigned.³⁸ 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 low-density lipoprotein cholesterol 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 genome-wide association studies (GWAS) of Europeans in the GLGC.³⁹ Of 76 independent association signals, two 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 material online, Table S3](#)), 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 material online, Table S4](#). Vascular disease outcomes included major coronary events (MCEs: non-fatal myocardial infarction, fatal ischaemic heart disease, or coronary revascularization); fatal/nonfatal ischaemic stroke; fatal/nonfatal intracerebral haemorrhage; major occlusive vascular events (MOVEs: consisting of fatal/nonfatal MCEs or ischaemic stroke); fatal occlusive vascular events (FOVEs); fatal CVD; and major vascular events (MVEs: fatal/nonfatal myocardial infarction, coronary revascularization procedures, stroke, or fatal CVD). These analyses used a common set of controls that excluded all individuals who self-reported prior coronary heart disease (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 as previously used for a drug target Mendelian randomization study of CETP.¹⁵ These were organized by anatomical site and consisted of combined incident, self-reported, and screen-detected cases of diabetes; 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 material online, Table S4](#)). 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 material online, Table S5](#)), 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 equations⁴⁰ according to participants' ancestry, age, height, and sex (see [Supplementary material online, Figure S3](#)). Alternatively, we defined prevalent COPD cases as those with FEV1/FVC < 0.7. 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 the cause of death) or non-fatal. These analyses excluded all participants in the two 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 2007–08.

Statistical analysis

For assessment of the association of the PCSK9-GS with potential confounders, the population-representative subset was divided into four score categories defined by empirically determined thresholds of 1.33, 1.43, and 1.53 (see [Supplementary material online, Figure S4](#)). Differences between categories were tested by analysis of variance using the R `anova()` function.

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-squared, 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. Associations of PCSK9 SNPs with spirometry were extracted from previous GWAS in CKB of FEV1/FVC⁴¹ 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 analysis of variance to compare the fit 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-squared, 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 amongst 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 material online, Table S6](#)).

For confirmation of the therapeutic effects of inhibition of PCSK9 in reducing the risk of CVD outcomes,^{3,4} we assessed a combination of simple and composite outcomes, with a Bonferroni-adjusted threshold of significance ($P < 0.05/3 = 0.0167$) based on three independent tests (determined by spectral decomposition of the correlation matrix for these related endpoints⁴²). 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 data set 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.*,²⁴ 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 (see [Supplementary material online, Table S7](#)); for cerebrovascular disease, the analysis was limited strictly to the population-representative subset—this was to minimize the risk of bias due to over- or under-representation of particular disease subtypes, since the case-control part of the genotyped data set is not representative of the full spectrum of cerebrovascular disease. Replication of these prior associations and tests for heterogeneity using Cochran's *Q*-statistic were each assessed 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.*²⁴ (and Nelson *et al.*²⁷) 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 GLGC⁴³ (per-allele beta coefficient: 0.497 SD) and scaled the disease association estimates reported by Rao *et al.*²⁴ 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 the study baseline was 52.1 (SD 10.7), 59.7% were female, and 45.8% were from urban regions. Stratifying individuals into four groups according to their PCSK9-GS values, no appreciable differences between these 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. Of the risk factors assessed, only LDL-C exhibited a significant difference across the PCSK9-GS categories after adjustment for multiple testing.

Characterization 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 (see [Supplementary material online, Table S1](#)). The SNP with the strongest effect, rs151193009C > T, is a known loss-of-function variant that disrupts the binding of the

Table 1 Baseline characteristics of the population-representative subset of genotyped China Kadoorie Biobank participants, divided into four categories according to their PCSK9 genetic score

PCSK9-GS range	<1.33 (n = 4557)	1.33 to <1.43 (n = 20 274)	1.43 to <1.53 (n = 39 786)	≥1.53 (n = 6297)	P
PCSK9-GS, mean (SD)	1.11 (0.24)	1.40 (0.01)	1.49 (0.02)	1.57 (0.03)	—
LDL-C (mmol/L), mean (SD) ^a	2.13 (0.65)	2.32 (0.65)	2.39 (0.71)	2.39 (0.72)	1.6 × 10 ⁻⁸
Age (years), mean (SD)	52.2 (10.8)	52.0 (10.7)	52.2 (10.7)	52.3 (10.8)	0.23
Female, %	59.9	59.9	59.9	58.0	0.03
Urban region, %	45.3	45.6	46.1	45.1	0.26
Height (cm), mean (SD)	159.1 (8.2)	158.8 (8.3)	158.7 (8.3)	158.9 (8.4)	0.03
BMI (kg/m ²), mean (SD)	23.9 (3.5)	23.8 (3.4)	23.8 (3.4)	23.9 (3.5)	0.03
Waist/hip ratio, mean (SD)	0.88 (0.07)	0.88 (0.07)	0.88 (0.07)	0.88 (0.07)	0.02
SBP (mmHg), mean (SD)	131.7 (21.4)	131.3 (21.3)	131.5 (21.5)	131.9 (21.1)	0.14
DBP (mmHg), mean (SD)	78.3 (11.2)	77.9 (11.2)	78.0 (11.3)	78.2 (11.1)	0.09
Physical activity (MET h/d), mean (SD)	20.5 (14.1)	20.8 (14.0)	20.5 (14.0)	20.7 (14.1)	0.10
High school or higher education, %	21.3	20.6	21.4	20.0	0.02
Household income ≥20 000 yuan, %	39.6	40.5	40.0	39.8	0.60
Current regular alcohol drinking, %	15.8	15.3	15.4	16.2	0.31
Current regular smoker, %	25.0	26.0	25.7	27.1	0.06
Stroke/TIA, ^a %	1.6	1.7	1.7	1.2	0.04
Coronary heart disease, ^b %	3.7	3.4	3.7	3.5	0.42
Hypertension, %	12.6	11.4	11.6	11.3	0.14
Diabetes, ^c %	6.1	5.9	6.2	6.2	0.65
Taking statins, %	1.8	1.6	1.9	2.4	0.50

BMI, body mass index; DBP, diastolic blood pressure; LDL-C, low-density lipoprotein cholesterol; MET, metabolic equivalent of task; PCSK9-GS, PCSK9 genetic score; SBP, systolic blood pressure; TIA, transient ischaemic attack.

^aResults for 4181 members of the population-representative subset with LDL-C measurements.

^bSelf-reported physician diagnosed.

^cSelf-reported physician diagnosed or screen-detected at baseline.

PCSK9 protein prodomain to heparan sulfate proteoglycans (HSPGs), an essential step in PCSK9-mediated degradation of LDL receptors (LDLRs).⁴⁴ When combined into the weighted PCSK9-GS, in 100-fold cross-validation analyses, 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 (see [Supplementary material online, Figure S5](#)). 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% confidence interval: -1.08 to -0.82); $P = 6.7 \times 10^{-46}$] and, more weakly, a lowering of Lp(a) [-0.14 (-0.27 to -0.01); $P = 0.038$]. No significant associations were detected with triglycerides [-0.01 (-0.14 to 0.12); $P = 0.87$] or HDL-C [-0.10 (-0.23 to 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 amongst the smaller subclasses of very low-density lipoproteins and across intermediate- and low-density lipoprotein subclasses (see [Supplementary material online, Figure S6](#)). 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 (see [Supplementary material online, Figures S6 and S7](#)).

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 that was significant after adjustment for multiple testing [$n = 15\,752$; odds ratio (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, FOVE, 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 to -0.12); $P = 0.00012$] and 39% lower odds of carotid plaque [$n = 8340$ cases; 0.61 (0.45–0.83); $P = 0.0015$] (see [Supplementary material online, Table S8](#)). Conversely, no association of PCSK9-GS was observed with the risk of intracerebral haemorrhage [$n = 5906$; 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 material online, Table S3](#)). For example, using the LDL-C genetic score, 1 SD lower LDL-C was associated with a lower risk of MCE [0.74 (0.66–0.84); $P = 1.5 \times 10^{-6}$] 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 lower 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); $P = 8.9 \times 10^{-3}$], 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](#)). The association of PCSK9-GS was also investigated for a wide range of other disease outcomes available in CKB using a phenome-wide approach (see [Supplementary material online, Figure S8](#)), with no association meeting the pre-specified significance threshold.

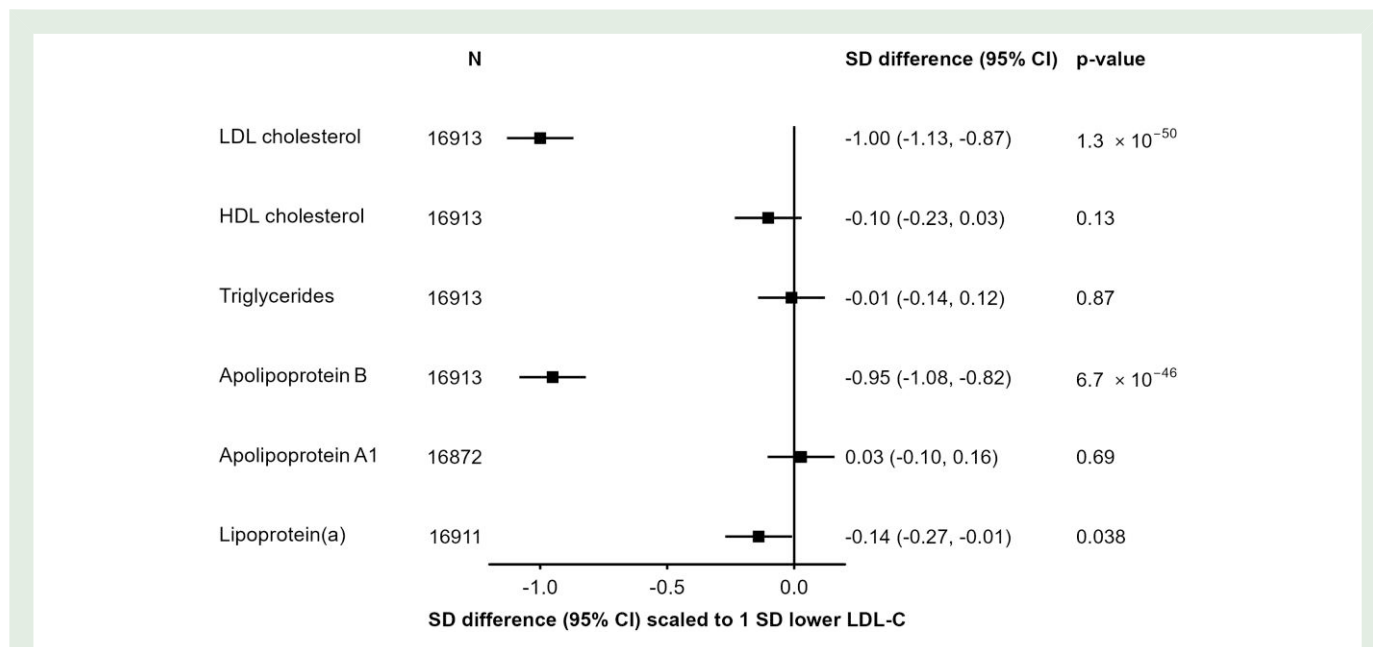


Figure 1 Associations of PCSK9 genetic score with major blood lipids and apolipoproteins. Estimates are standardized beta coefficients from linear regressions of the weighted PCSK9 genetic score, scaled to a 1 SD lowering of low-density lipoprotein cholesterol, calculated as an inverse-variance-weighted average of region-specific estimates with adjustment for age, age-squared, sex, ascertainment, and region-specific principal components. CI, confidence interval; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

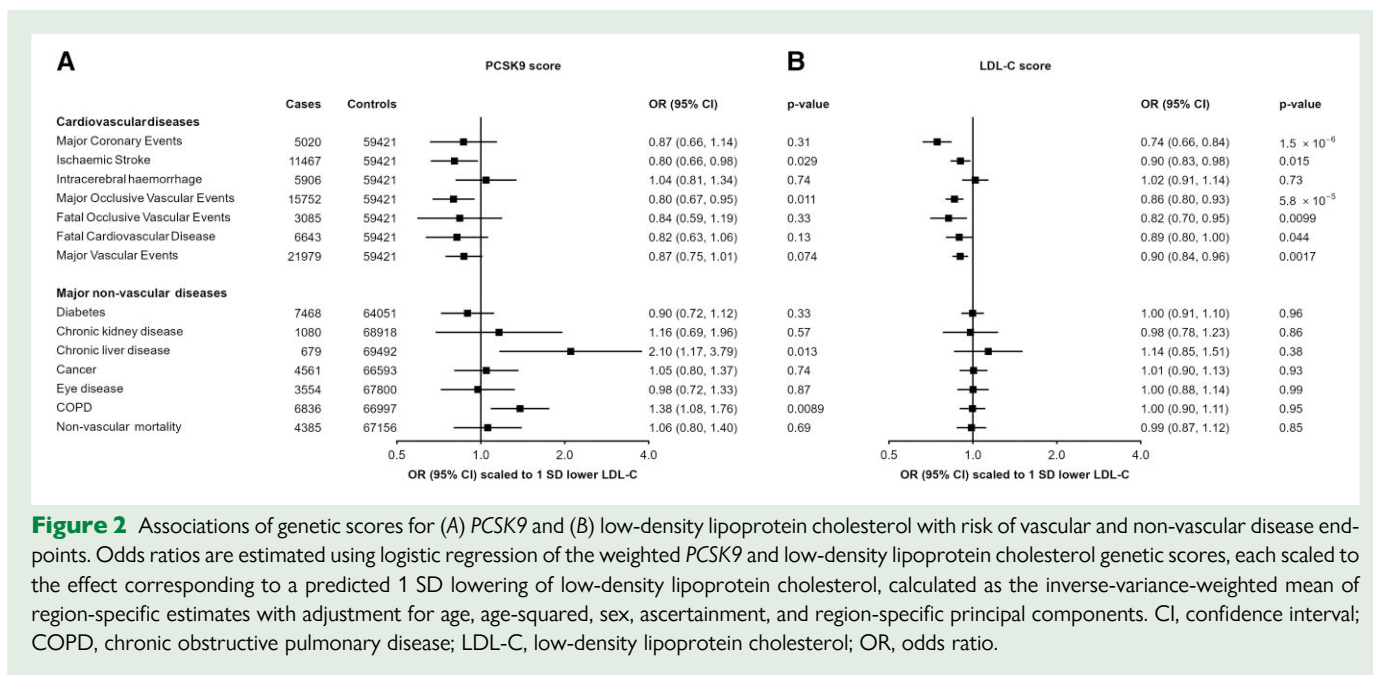


Figure 2 Associations of genetic scores for (A) *PCSK9* and (B) low-density lipoprotein cholesterol with risk of vascular and non-vascular disease endpoints. Odds ratios are estimated using logistic regression of the weighted *PCSK9* and low-density lipoprotein cholesterol genetic scores, each scaled to the effect corresponding to a predicted 1 SD lowering of low-density lipoprotein cholesterol, calculated as the inverse-variance-weighted mean of region-specific estimates with adjustment for age, age-squared, sex, ascertainment, and region-specific principal components. CI, confidence interval; COPD, chronic obstructive pulmonary disease; LDL-C, low-density lipoprotein cholesterol; OR, odds ratio.

Replication of reported *PCSK9* associations

For seven traits previously reported as being nominally associated with a missense *PCSK9* variant (rs11591147) in UKB,²⁴ 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 a higher risk of URTI [$n = 1095$ cases; 2.18 (1.34–3.35); $P = 1.6 \times 10^{-3}$] was significant after adjustment for multiple testing and was consistent with the association observed in UKB [$n = 2364$ cases; 1.70 (1.15–2.51)], providing an estimate after meta-analysis of 1.87 [(1.38–2.54); $P = 5.4 \times 10^{-5}$]. 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 UKB [$n = 39\,269$ 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 compared with a population reference [$n = 5105$; 0.96 (0.73–1.26); $P = 0.75$], nor were these SNPs associated with baseline FEV1/FVC as a continuous measure (see Supplementary material online, Table S9). By contrast, the association with higher risk of an incident COPD event was stronger in those with pre-existing COPD [2212 exacerbations amongst 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 the risk of a non-fatal exacerbation [1482 non-fatal exacerbations: 1.58 (0.86–2.91); $P = 0.14$; P -het = 0.24]. Similar results were obtained when defining prevalent COPD as FEV1/FVC < 0.7 (see Supplementary material online, Figure S9), although with wider confidence intervals due to fewer prevalent COPD cases by this definition. 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 (see Supplementary material online, Figure S10).

The association with risk of COPD exacerbation in CKB was not replicated in UKB using a different functional *PCSK9* variant (see Supplementary material online, Table S10), although power was limited due to a much smaller number of recorded cases (468 in UKB vs. 2212 cases in CKB). After meta-analysis, the pooled OR was 1.45 [(0.91–2.32); $P = 0.12$] but with significant heterogeneity between the two estimates (P -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 -het = 0.013] and fatal COPD exacerbations [*PCSK9*: 3.61 (1.71–7.60); LDL-C GS: 0.92 (0.75–1.19); P -het = 0.0012].

Discussion

We used genetic variants, including a loss-of-function variant, to mimic the 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 directionally concordant effects on risk of FOVE and MCE. Whilst 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^{3,4} and other genetic studies,^{22,24,45,46} they

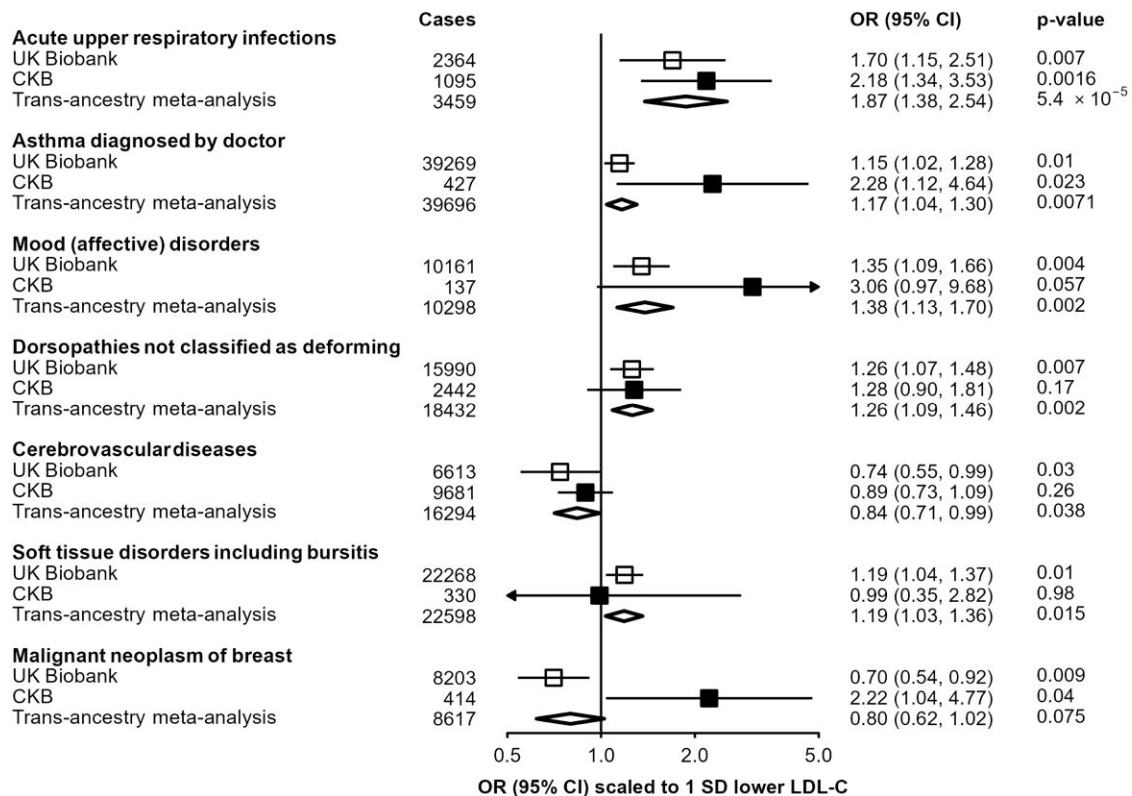


Figure 3 Replication in China Kadoorie Biobank of previously reported associations in UK Biobank of *PCSK9* genetic scores with disease outcomes. Estimates in UK Biobank (open boxes) originate from data published by Rao *et al.*²⁴ and are scaled to a corresponding 1 SD lowering of low-density lipoprotein cholesterol. Odds ratios in China Kadoorie Biobank (filled boxes) are estimated using logistic regression of the weighted *PCSK9* genetic score, scaled to a 1 SD lowering of low-density lipoprotein cholesterol, calculated as an inverse-variance-weighted average of region-specific estimates with adjustment for age, age-squared, sex, ascertainment, and region-specific principal components. Trans-ancestry estimates (diamonds) are derived from inverse-variance-weighted fixed-effect meta-analysis. CI, confidence interval; CKB, China Kadoorie Biobank; OR, odds ratio.

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.

Mendelian randomization 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%.⁴⁷ 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.¹ 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.⁶ Similarly, the GLAGOV trial illustrated a dose–response relationship between LDL-C lowering and volume of atherosclerosis.⁵ Thus, although the population studied here had much lower LDL-C levels at baseline as compared with European studies [e.g. 2.35 (0.69 SD) mmol/L CKB vs. 3.57 (0.87) UKB⁴⁷], triangulation of our genetic findings with studies in Europeans and randomized controlled trials 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 with already-low LDL-C. Thus, in high-risk individuals, it may be beneficial to adopt intensive therapy (e.g. adding a *PCSK9* inhibitor to statin) so as to reduce LDL-C below the normal target range.

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 genome scan of another functional variant in *PCSK9* conducted within the UKB.^{24,27} In addition to these three respiratory disease outcomes, meta-analysis of our results with those from UKB strengthened the evidence for an association of *PCSK9* genetic variants with mood disorders and with non-deforming dorsopathies (diseases of the spine). Previous studies have identified an association of *PCSK9* with severe forms of dorsopathy,⁴⁸ 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 liver⁴⁹ 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 (VLDLRs).^{50,51} *PCSK9*, LDLR, and VLDLR are each also expressed in the lung,⁵² where VLDLR and

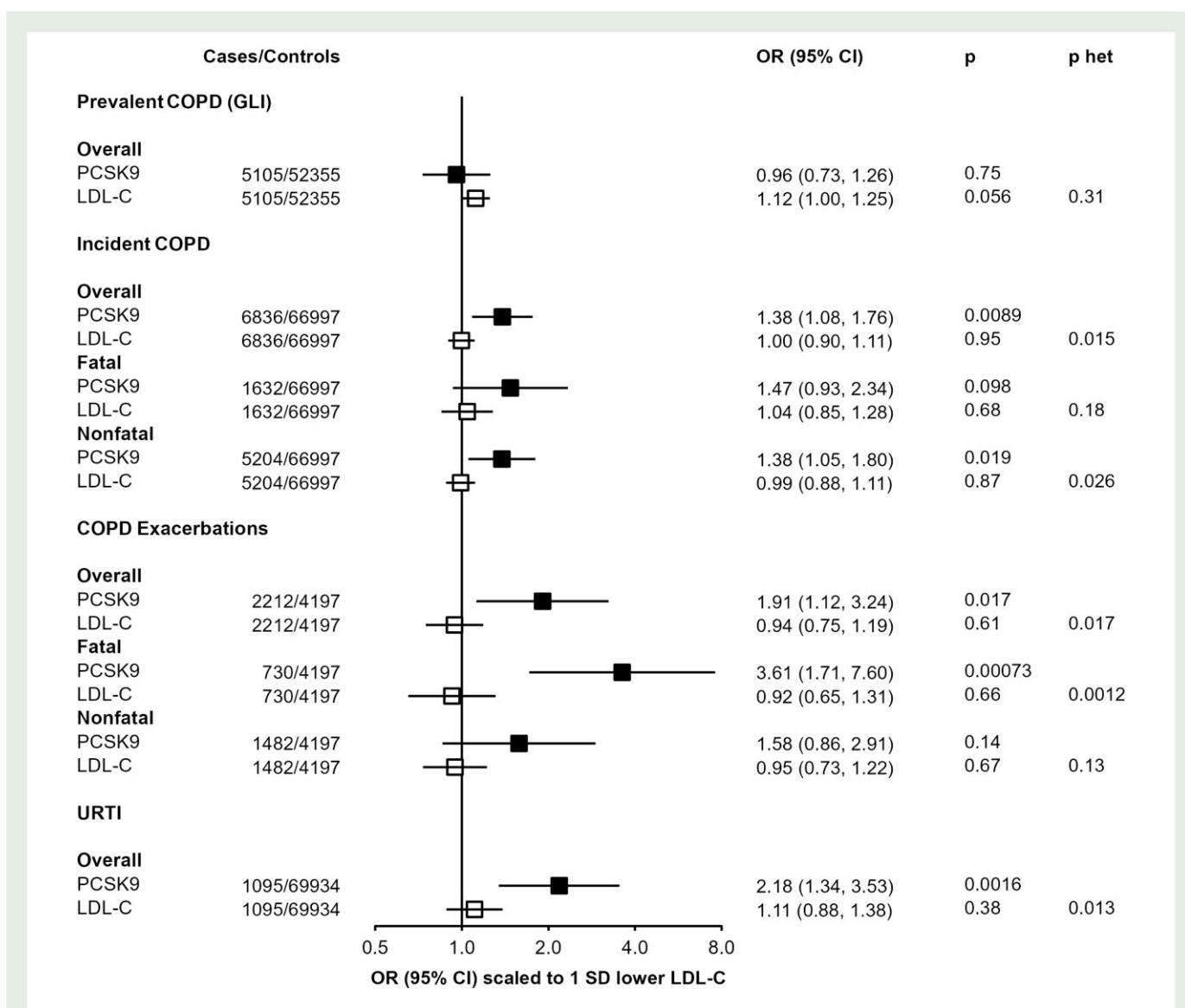


Figure 4 Associations of *PCSK9* and low-density lipoprotein cholesterol genetic scores with respiratory disease endpoints. Odds ratios are estimated using logistic regression of the weighted *PCSK9* genetic score, scaled to 1 SD lower low-density lipoprotein cholesterol, calculated as an inverse-variance-weighted average of region-specific estimates with adjustment for age, age-squared, sex, ascertainment, and region-specific principal components. Prevalent chronic obstructive pulmonary disease defined on basis of spirometry (see Methods). COPD exacerbations are incident events occurring in individuals with spirometry-defined prevalent chronic obstructive pulmonary disease. *P*-het is the *P*-value from a test for heterogeneity for estimates of *PCSK9* and low-density lipoprotein cholesterol genetic instruments. CI, confidence interval; COPD, chronic obstructive pulmonary disease; LDL-C, low-density lipoprotein cholesterol; OR, odds ratio; URTI, upper respiratory tract infection.

LDLR are involved in the initiation of infection by several classes of respiratory viruses (e.g. human rhinovirus and some types of coronavirus), which bind to VLDLR and LDLR on the cell surface, and are internalized via receptor-mediated endocytosis.^{53,54} Inhibition of LDLR in cultured human tracheal epithelial cells lowers the infectivity of human rhinovirus.⁵⁵ 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 COPD⁵⁶ and worsen asthma symptoms⁵⁷) 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.⁵⁸ 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-localized *PCSK9* inhibitors. Whilst various parameters

determine the tissue distribution of monoclonal antibodies,⁵⁹ current evidence suggests that evolocumab has limited tissue distribution.^{60,61} Furthermore, inclisiran, a small interfering RNA inhibitor of PCSK9,⁶² features N-acetylgalactosamine conjugation that renders it hepatotropic.⁶³ Consequently, limited tissue penetration of biological agents may mean that the genetic associations we identify do not have major clinical relevance.

Two large phase III trials and subsequent open-label extensions found no excess risk of respiratory disease in individuals treated with PCSK9 inhibition for up to 9 years.^{9,10} However, the association of PCSK9-GS with exacerbations of COPD indicates that PCSK9 inhibition potentially confers 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. Whilst some phase II trials did show a non-significant excess risk of URTI in those treated with evolocumab,^{64–66} such flu-like symptoms may be common following treatment with monoclonal antibodies in general,⁶⁷ 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 the 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 rhinitis⁶⁸; 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. It is unclear; however, to what extent the aetiology or definition of respiratory diseases differs between biobanks which derive from very different populations with contrasting exposures to other risk factors for respiratory disease.⁶⁹ 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 LDLR degradation.⁴⁴ By contrast, the UKB functional variant (rs11591147) instead exerts its effect by modulating local protein structure.⁷⁰ Whilst both lead to lower LDL-C, suggesting overall similarities in their biological consequences, PCSK9 mediates degradation of multiple proteins,⁵⁰ so that different SNPs may vary in their impact on differing PCSK9-mediated pathways, including those not directly related to LDL-C lowering. Indeed, it is possible that this accounts for the contrasting associations observed for breast cancer—one PCSK9 target, LRP8, has been implicated in triple-negative breast cancer.⁷¹ Third, the association of PCSK9-GS with URTI might arise due to the life-long effects of genetic perturbations that alter human biology,¹¹ 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 VLDL and LDL 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. Whilst 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,⁷² 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 randomized 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-characterized 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. Whilst 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.

Supplementary material

Supplementary material is available at *European Journal of Preventive Cardiology*.

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.

Author contributions

M.V.H., C.K., R.B., K.L., N.R., C.Y., J.L., D.A.B., M.R.H., L.Y., Y.C., H.D., I.T., M.D.T., I.Y.M., and R.G.W.: data collection and analysis; Z.C., L.L., R.G.W., I.Y.M., R.J.C., M.R.H., and R.C.: funding; M.V.H. and R.G.W.: study oversight; and M.V.H., R.G.W., I.Y.M., Z.C.: manuscript draft and revision. 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, Hong Kong. Long-term follow-up was supported by the Wellcome Trust, UK (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, and 088158/Z/09/Z), the National Natural Science Foundation of China (82192901, 82192904, and 82192900), and the National Key Research and Development Program of China, China (2016YFC0900500). DNA extraction and genotyping were funded by GlaxoSmithKline and the Medical Research Council, UK (MC-PC-13049 and MC-PC-14135). The project is supported by core funding from the Medical Research Council, UK (MC_UU_00017/1, MC_UU_12026/2, and MC_U137686851), Cancer Research UK (C16077/A29186 and 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. M.V.H. was

supported by an Intermediate Clinical Research Fellowship from the British Heart Foundation (FS/18/23/33512) and by the National Institute for Health Research (Oxford Biomedical Research Centre). M.D.T. was partially supported through the National Institute for Health and Care Research, UK (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. M.D.T. was also supported via Wellcome Trust, UK 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.

Ethical 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.

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

Data availability

The 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 are 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 CKB study that underlies this paper should contact 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 are 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. 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

- Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet* 2016;**388**:2532–2561.
- Cannon CP, Blazing MA, Giugliano RP, McCagg A, White JA, Theroux P, et al. Ezetimibe added to statin therapy after acute coronary syndromes. *N Engl J Med* 2015;**372**:2387–2397.
- Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, et al. Evolocumab and clinical outcomes in patients with cardiovascular disease. *N Engl J Med* 2017;**376**:1713–1722.
- Schwartz GG, Steg PG, Szarek M, Bhatt DL, Bittner VA, Diaz R, et al. Alirocumab and cardiovascular outcomes after acute coronary syndrome. *N Engl J Med* 2018;**379**:2097–2107.
- Nicholls SJ, Puri R, Anderson T, Ballantyne CM, Cho L, Kastelein JJ, et al. Effect of evolocumab on progression of coronary disease in statin-treated patients: the GLAGOV randomized clinical trial. *JAMA* 2016;**316**:2373–2384.
- Giugliano RP, Pedersen TR, Park JG, De Ferrari GM, Gaciong ZA, Ceska R, 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**:1962–1971.
- 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**:364–367.
- Sattar N, Preiss D, Murray HM, Welsh P, Buckley BM, de Craen AJ, et al. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. *Lancet* 2010;**375**:735–742.
- O'Donoghue ML, Giugliano RP, Wiviott SD, Atar D, Keech A, Kuder JF, et al. Long-term evolocumab in patients with established atherosclerotic cardiovascular disease. *Circulation* 2022;**146**:1109–1119.
- Ray KK, Troquay RPT, Visseren FLJ, Leiter LA, Scott Wright R, Vikarunnessa S, et al. Long-term efficacy and safety of inclisiran in patients with high cardiovascular risk and elevated LDL cholesterol (ORION-3): results from the 4-year open-label extension of the ORION-1 trial. *Lancet Diabetes Endocrinol* 2023;**11**:109–119.
- Holmes MV, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat Rev Cardiol* 2017;**14**:577–590.
- Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018;**362**:k601.
- Holmes MV. Human genetics and drug development. *N Engl J Med* 2019;**380**:1076–1079.
- Swerdlow DI, Preiss D, Kuchenbaecker KB, Holmes MV, Engmann JE, Shah T, et al. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* 2015;**385**:351–361.
- Millwood IY, Bennett DA, Holmes MV, Boxall R, Guo Y, Bian Z, et al. Association of CETP gene variants with risk for vascular and nonvascular diseases among Chinese adults. *JAMA Cardiol* 2018;**1**:34–43.
- 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–e659.
- RECOVERY Collaborative Group. Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet* 2021;**397**:1637–1645.
- Holmes MV, Simon T, Exeter HJ, Folkersen L, Asselbergs FW, Guardiola M, et al. Secretory phospholipase A2-IIA and cardiovascular disease: a Mendelian randomization study. *J Am Coll Cardiol* 2013;**62**:1966–1976.
- Millwood IY, Bennett DA, Walters RG, Clarke R, Waterworth D, Johnson T, et al. A genome-wide association study of a lipoprotein-associated phospholipase A2 loss-of-function variant in 90 000 Chinese adults. *Int J Epidemiol* 2016;**45**:1588–1599.
- Millwood IY, Bennett DA, Walters RG, Clarke R, Waterworth D, Johnson T, 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**:230–231.
- Sliz E, Kettunen J, Holmes MV, Williams CO, Boachie C, Wang Q, et al. Metabolomic consequences of genetic inhibition of PCSK9 compared with statin treatment. *Circulation* 2018;**138**:2499–2512.
- Ference BA, Robinson JG, Brook RD, Catapano AL, Chapman MJ, Neff DR, et al. Variation in PCSK9 and HMGCR and risk of cardiovascular disease and diabetes. *N Engl J Med* 2016;**375**:2144–2153.
- Schmidt AF, Holmes MV, Patel R, Swerdlow DI, Denaxas S, Fatemifar G, et al. Phenome-wide association analysis of LDL-cholesterol lowering genetic variants in PCSK9. *BMC Cardiovasc Disord* 2019;**19**:240.
- 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.
- Schmidt AF, Swerdlow DI, Holmes MV, Patel RS, Fairhurst-Hunter Z, Lyall DM, et al. PCSK9 genetic variants and risk of type 2 diabetes: a Mendelian randomisation study. *Lancet Diabetes Endocrinol* 2017;**5**:97–105.
- Lotta LA, Sharp SJ, Burgess S, Perry JRB, Stewart ID, Willems SM, et al. Association between low-density lipoprotein cholesterol-lowering genetic variants and risk of type 2 diabetes: a meta-analysis. *JAMA* 2016;**316**:1383–1391.
- Nelson CP, Lai FY, Nath M, Ye S, Webb TR, Schunkert H, 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.
- Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* 2011;**40**:1652–1666.
- Kurmi OP, Li L, Smith M, Augustyn M, Chen J, Collins R, 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 Respir Res* 2014;**1**:e000025.

30. Clarke R, Du H, Kurmi O, Parish S, Yang M, Arnold M, et al. Burden of carotid artery atherosclerosis in Chinese adults: implications for future risk of cardiovascular diseases. *Eur J Prev Cardiol* 2017;**24**:647–656.
31. Sun L, Clarke R, Bennett D, Guo Y, Walters RG, Hill M, et al. Causal associations of blood lipids with risk of ischemic stroke and intracerebral hemorrhage in Chinese adults. *Nat Med* 2019;**25**:569–574.
32. Soininen P, Kangas AJ, Wurtz P, Tukiainen T, Tynkynen T, Laatikainen R, et al. High-throughput serum NMR metabolomics for cost-effective holistic studies on systemic metabolism. *Analyst* 2009;**134**:1781–1785.
33. 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.
34. Walters RG, Millwood IY, Lin K, Schmidt Valle D, McDonnell P, Hacker A, et al. Genotyping and population characteristics of the China Kadoorie Biobank. *Cell Genom* 2023;**3**:100361.
35. Loh PR, Kichaev G, Gazal S, Schoech AP, Price AL. Mixed-model association for biobank-scale datasets. *Nat Genet* 2018;**50**:906–908.
36. 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**:1493–1501.
37. Graham SE, Clarke SL, Wu K-HH, Kanoni S, Zajac GJM, Ramdas S, et al. The power of genetic diversity in genome-wide association studies of lipids. *Nature* 2021;**600**:675–679.
38. Hartwig FP, Davies NM. Why internal weights should be avoided (not only) in MR-egger regression. *Int J Epidemiol* 2016;**45**:1676–1678.
39. Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet* 2013;**45**:1345–1352.
40. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, 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**:1324–1343.
41. Shrine N, Izquierdo AG, Chen J, Packer R, Hall RJ, Guyatt AL, et al. Multi-ancestry genome-wide association analyses improve resolution of genes and pathways influencing lung function and chronic obstructive pulmonary disease risk. *Nat Genet* 2023;**55**:410–422.
42. Machado AMC. Multiple testing correction in medical image analysis. *J Math Imaging Vis* 2007;**29**:107–117.
43. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013;**45**:1274–1283.
44. Gustafsen C, Olsen D, Vilstrup J, Lund S, Reinhardt A, Wellner N, et al. Heparan sulfate proteoglycans present PCSK9 to the LDL receptor. *Nat Commun* 2017;**8**:503.
45. 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**:1264–1272.
46. Ference BA, Yoo W, Alesh I, Mahajan N, Mirowska KK, Mewada A, 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. *J Am Coll Cardiol* 2012;**60**:2631–2639.
47. Richardson TG, Sanderson E, Palmer TM, Ala-Korpela M, Ference BA, Davey Smith G, 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.
48. Jerome RN, Pulley JM, Roden DM, Shirey-Rice JK, Bastarache LA, R. Bernard G, et al. Using human 'experiments of nature' to predict drug safety issues: an example with PCSK9 inhibitors. *Drug Saf* 2018;**41**:303–311.
49. Grefhorst A, McNutt MC, Lagace TA, Horton JD. Plasma PCSK9 preferentially reduces liver LDL receptors in mice. *J Lipid Res* 2008;**49**:1303–1311.
50. Poirier S, Mayer G, Benjannet S, Bergeron E, Marcinkiewicz J, Nassoury N, et al. The pro-protein convertase PCSK9 induces the degradation of low density lipoprotein receptor (LDLR) and its closest family members VLDLR and ApoER2. *J Biol Chem* 2008;**283**:2363–2372.
51. 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.
52. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science* 2015;**347**:1260419.
53. Rankl C, Kienberger F, Wildling L, Wruess J, Gruber HJ, Blaas D, et al. Multiple receptors involved in human rhinovirus attachment to live cells. *Proc Natl Acad Sci U S A* 2008;**105**:17778–17783.
54. Hofer F, Gruenberger M, Kowalski H, Machat H, Huettinger M, Kuechler E, 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**:1839–1842.
55. Suzuki T, Yamaya M, Kamanaka M, Jia YX, Nakayama K, Hosoda 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–L420.
56. Sapely E, Stockley RA. COPD exacerbations. 2: aetiology. *Thorax* 2006;**61**:250–258.
57. Busse WW, Lemanske RF Jr, Gern JE. Role of viral respiratory infections in asthma and asthma exacerbations. *Lancet* 2010;**376**:826–834.
58. 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**:1135–1148.
59. Tabrizi M, Bornstein GG, Suria H. Biodistribution mechanisms of therapeutic monoclonal antibodies in health and disease. *AAPS J* 2010;**12**:33–43.
60. Henry CA, Lyon RA, Ling H. Clinical efficacy and safety of evolocumab for low-density lipoprotein cholesterol reduction. *Vasc Health Risk Manag* 2016;**12**:163–169.
61. Kasichayanula S, Grover A, Emery MG, Gibbs MA, Somaratne R, Wasserman SM, et al. Clinical pharmacokinetics and pharmacodynamics of evolocumab, a PCSK9 inhibitor. *Clin Pharmacokinet* 2018;**57**:769–779.
62. Mullard A. Nine paths to PCSK9 inhibition. *Nat Rev Drug Discov* 2017;**16**:299–301.
63. Nair JK, Willoughby JL, Chan A, Charisse K, Alam MR, Wang Q, et al. Multivalent N-acetylgalactosamine-conjugated siRNA localizes in hepatocytes and elicits robust RNAi-mediated gene silencing. *J Am Chem Soc* 2014;**136**:16958–16961.
64. Blom DJ, Hala T, Bolognese M, Lillestol MJ, Toth PD, Burgess L, et al. A 52-week placebo-controlled trial of evolocumab in hyperlipidemia. *N Engl J Med* 2014;**370**:1809–1819.
65. Sabatine MS, Giugliano RP, Wiviott SD, Raal FJ, Blom DJ, Robinson J, et al. Efficacy and safety of evolocumab in reducing lipids and cardiovascular events. *N Engl J Med* 2015;**372**:1500–1509.
66. Raal FJ, Honarpour N, Blom DJ, Hovingh GK, Xu F, Scott R, et al. Inhibition of PCSK9 with evolocumab in homozygous familial hypercholesterolaemia (TESLA Part B): a randomised, double-blind, placebo-controlled trial. *Lancet* 2015;**385**:341–350.
67. Baldo BA. Adverse events to monoclonal antibodies used for cancer therapy: focus on hypersensitivity responses. *Oncoimmunology* 2013;**2**:e26333.
68. Brumpton BM, Fritsche LG, Zheng J, Nielsen JB, Mannila M, Surakka I, 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.
69. Chan KH, Kurmi OP, Bennett DA, Yang L, Chen Y, Tan Y, et al. Solid fuel use and risks of respiratory diseases: a cohort study of 280,000 Chinese never-smokers. *Am J Respir Crit Care Med* 2019;**199**:352–361.
70. Sarkar SK, Foo ACY, Matyas A, Asikhia I, Kosenko T, Goto NK, et al. A transient amphipathic helix in the prodomain of PCSK9 facilitates binding to low-density lipoprotein particles. *J Biol Chem* 2020;**295**:2285–2298.
71. Lin CC, Lo MC, Moody R, Jiang H, Harouaka R, Stevers N, et al. Targeting LRP8 inhibits breast cancer stem cells in triple-negative breast cancer. *Cancer Lett* 2018;**438**:165–173.
72. Ference BA. How to use Mendelian randomization to anticipate the results of randomized trials. *Eur Heart J* 2018;**39**:360–362.