# Lipoprotein(a) and cardiovascular disease: prediction, attributable risk fraction and estimating benefits from novel interventions 

## Brief Title: Lipoprotein(a) and cardiovascular disease in UK biobank

Paul Welsh*a, Claire Welsh*b ${ }^{\text {b }}$, Carlos A Celis-Morales ${ }^{\text {a }}$, Rosemary Brown ${ }^{\text {a }}$, Lyn D Ferguson ${ }^{\text {a }}$, Patrick B Mark ${ }^{\text {a }}$, James Lewsey ${ }^{\text {c }}$, Stuart R Gray ${ }^{\text {a }}$, Donald M Lyall ${ }^{\text {c }}$, Jason MR Gill $^{\mathrm{a}}$, Jill P Pellc ${ }^{\mathrm{c}}$, James A de Lemos ${ }^{\mathrm{d}}$, Peter Willeit ${ }^{\mathrm{e}}$, Naveed Sattar ${ }^{\text {a }}$<br>*Joint first authors

## Affiliations

${ }^{\text {a }}$ Institute of Cardiovascular and Medical Sciences, University of Glasgow
${ }^{\mathrm{b}}$ Institute for Ageing, Newcastle University
${ }^{\text {c }}$ Institute of Health and Wellbeing, University of Glasgow
${ }^{\text {d }}$ Division of Cardiology, Department of Medicine, University of Texas Southwestern Medical Center
${ }^{e}$ Department of Neurology, Medical University of Innsbruck, Innsbruck, Austria; Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom.

## Funding

The work in this study was supported by a grant from Chest, Heart, and Stroke Association Scotland [Res16/A165], and by the British Heart Foundation Research Excellence Award (RE/18/6/34217).

## Disclosures

Peter W reports personal fees from Novartis Pharmaceuticals outside the submitted work. NS has consulted for Amgen and Sanofi. All other authors report no financial disclosures.

## Address for correspondence

Naveed Sattar
Institute of Cardiovascular and Medical Sciences, BHF Glasgow Cardiovascular Research Centre, University of Glasgow,
126 University Place, Glasgow G12 8TA, UK.
Tel: 0044 (0) 1413303419
E-mail address: Naveed.sattar@glasgow.ac.uk
Twitter: @MetaMedTeam

## Acknowledgments

This research was conducted using the UK Biobank resource. We thank the participants of the UK Biobank. The work was performed under UK biobank project number 9310.


#### Abstract

\section*{Background}

Lipoprotein (a) (Lp(a)) is a CVD risk factor amenable to intervention and might help guide risk prediction.


## Objectives

To investigate the population attributable fraction due to elevated $\operatorname{Lp}(a)$ and its utility in risk prediction.

## Methods

Using a prospective cohort study, 413,724 participants from UK Biobank, associations of serum Lp(a) with composite fatal/nonfatal CVD ( $\mathrm{n}=10,065$ events), fatal CVD ( $\mathrm{n}=3247$ ), coronary heart disease ( $\mathrm{n}=16,649$ ), ischaemic stroke ( $\mathrm{n}=3191$ ), and peripheral vascular disease ( $\mathrm{n}=2716$ ) were compared using Cox models. Predictive utility was determined by Cindex changes. The population attributable fraction was estimated.

## Results

Median $\operatorname{Lp}(\mathrm{a})$ was $19.7 \mathrm{nmol} / \mathrm{L}$ (interquartile interval 7.6-75.3nmol/L). $20.8 \%$ had $\mathrm{Lp}(\mathrm{a})$ values $>100 \mathrm{nmol} / \mathrm{L} ; 9.2 \%$ had values $>175 \mathrm{nmol} / \mathrm{L}$. After adjustment for classical risk factors, in participants with no baseline CVD and not taking a statin, 1 standard deviation increment in $\log \operatorname{Lp}$ (a) was associated with a HR for fatal/nonfatal CVD of 1.09 ( $95 \%$ CI 1.07-1.11). Associations were similar for fatal CVD, coronary heart disease, and peripheral vascular disease. Adding Lp(a) to a prediction model containing traditional CVD risk factors improved the C-index by +0.0017 ( $95 \%$ CI $0.0009,0.0026$ ). We estimated that having $\mathrm{Lp}(\mathrm{a})$ values $>100 \mathrm{nmol} / \mathrm{L}$ accounts for $5.7 \%$ of CVD events in the whole cohort. We modelled that an ongoing trial to lower $\operatorname{Lp}(a)$ in patients with CVD and $\operatorname{Lp}(a)$ above $\sim 175 \mathrm{nmol} / \mathrm{L}$ may reduce CVD risk by $20.3 \%$, assuming causality, and an achieved $\mathrm{Lp}(\mathrm{a})$ reduction of $80 \%$.

## Conclusions

Population screening for elevated $\operatorname{Lp}(a)$ may help to predict CVD and target $\mathrm{Lp}(\mathrm{a})$ lowering drugs, if such drugs prove efficacious, to those with markedly elevated levels.

| Abbreviations |  |
| :---: | :---: |
| ACC | American College of Cardiology |
| AF | atrial fibrillation |
| AHA | American Heart Association |
| ASCVD | atherosclerotic cardiovascular disease |
| CHD | coronary heart disease |
| CI | confidence interval |
| CVD | cardiovascular disease |
| EAS | European Atherosclerosis Society |
| ERFC | Emerging Risk Factors Collaboration |
| ESC | European Society of Cardiology |
| HDL | high-density lipoproteins |
| HES | Hospital Episode Statistics |
| HF | heart failure |
| HR | hazard ratio |
| ICD | International Classification of Diseases |
| LDL | low-density lipoproteins |
| Lp(a) | Lipoprotein (a) |
| $\mathrm{mmol} / \mathrm{L}$ | millimoles per litre |
| NHS | National Health Service |
| nmol/L | nanomoles per litre |
| NRI | net reclassification index |
| PAF | paroxysmal atrial fibrillation |
| PCSK9 | Proprotein convertase subtilisin/kexin type 9 |
| PCE | Pooled Cohort Equations |
| PVD | peripheral vascular disease |
| REC | Research Ethics Committee |
| SCORE | Systematic Coronary Risk Estimation |

## Introduction

Lipoprotein (a) (Lp(a)) is a low-density lipoprotein particle made by the liver, comprised of both an apolipoprotein(a) and an apolipoprotein B protein. Its structure is highly heterogeneous, but levels of $\operatorname{Lp}(a)$ are $80 \%-90 \%$ genetically determined and relatively stable across the life-course.

Epidemiological evidence shows strong associations of circulating Lp(a) with atherogenesis and consequent risk of cardiovascular disease (CVD). For instance, in a recent meta-analysis of statin trial data, those with Lp (a) concentrations above $50 \mathrm{mg} / \mathrm{dl}$ were at $35 \%$ higher risk of incident CVD events $(95 \%$ CI $11-66 \%)$ compared to those with Lp (a) $<15 \mathrm{mg} / \mathrm{dl}$ (approx. $30 \mathrm{nmol} / \mathrm{L}$ ) after adjusting for confounders (1). Similar data have been reported in population studies $(2,3)$. Furthermore, genetic data as well as basic science support the notion that the association is causal (4). This has led to interest both in the potential for $\operatorname{Lp}$ (a) to serve as a biomarker to enhance CVD risk prediction (5), and as a therapeutic target. Indeed, Lp (a) lowering drugs may be a viable therapeutic option, with at least one drug moving to phase 3 in the Lp (a) HORIZON trial (6). Further, proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors, which lower $\mathrm{Lp}(\mathrm{a})$ concentrations by around $27 \%$, may be particularly beneficial in reducing risk in patients with previous CVD and substantially raised $\operatorname{Lp}(a)$ (7).

Currently, most guidelines and consensus statements do not advocate widespread screening for elevated $\operatorname{Lp}(a)$ and suggest focus should be on measuring $\operatorname{Lp}(a)$ in people at high risk of CVD or in those in whom risk is intermediate when high $\operatorname{Lp}(a)$ would serve as a risk enhancer (8-10). However, the recently released ESC/EAS guidelines suggested a "one-off measurement of $\operatorname{Lp}(a)$ may help to identify people with very high inherited $\operatorname{Lp}(a)$ levels who may have a substantial lifetime risk of atherosclerotic cardiovascular disease (ASCVD)" (11). Although Lp(a) is sometimes measured in patients with suspected familial hypercholesterolaemia (12), it is currently not routinely measured in general practice (9). In addition, there is conflicting advice on what constitutes a "high" $\mathrm{Lp}(\mathrm{a})$ level. Several guidelines and consensus statements advocate the $50 \mathrm{mg} / \mathrm{dl}(\sim 125 \mathrm{nmol} / \mathrm{L})$ cutoff $(8,9,13)$ as this corresponds to the $80^{\text {th }}$ percentile in one cohort study (14), but the 2016 Canadian Cardiovascular Society Guidelines use $30 \mathrm{mg} / \mathrm{dl}$ on the basis of elevated CVD risk ( $\sim 75 \mathrm{nmol} / \mathrm{L}$ ) (15). The Lp(a) HORIZON trial uses 70mg/dl ( $\sim 175 \mathrm{nmol} / \mathrm{L}$ ) as an inclusion criterion (16). The lack of data from a single large cohort with consistent phenotyping is a
significant limitation in interpreting the existing literature, impacting our understanding of the prevalence of high $\operatorname{Lp}(a)$, and its consequences for CVD risk.

UK Biobank is a large prospective population-based cohort study carried out in the UK, with information on baseline biochemistry measurements including routine lipids and $\mathrm{Lp}(\mathrm{a})$ measured in core laboratories. We aimed to use this resource to explore the shape of the association of $\mathrm{Lp}(\mathrm{a})$ with a range of distinct CVD outcomes to investigate the population attributable risk fraction for CVD that might be explained by elevated $\operatorname{Lp}(a)$, and to predict what might be the effect of novel $\operatorname{Lp}(a)$-lowering therapies based on these data and recent relevant trials.

## Methods

UK Biobank was conducted across 22 assessment centres across the UK between April 2007 and December 2010 and recruited 502,624 participants aged 37 to 73 . Baseline biological measurements were recorded and touch-screen questionnaires were administered according to a standardised protocol ( 17,18 ). UK Biobank received ethical approval from the North West Multi-Centre Research Ethics Committee (REC reference: 11/NW/03820). All participants gave written informed consent before enrolment in the study, which was conducted in accord with the principles of the Declaration of Helsinki.

For the present analysis, ethnicity was coded as white, South Asian, black, or mixed/other. Smoking status was categorised into never or former/current smoking. Systolic and diastolic blood pressure were measured at the baseline visit, preferentially using an automated measurement, but using manual measurement where this was not available. Blood collection sampling procedures for the study have been previously described and validated (19). The definition of baseline diabetes included self-reported type 1 or type 2 diabetes, those with a primary or secondary hospital diagnoses relating to diabetes at baseline (ICD-10 codes E10E14.9), and those who reported using diabetes medications. Statin (categorised to include other cholesterol lowering medications) and blood pressure medication use were also recorded from self-report. Baseline cardiovascular disease was defined as self-reported myocardial infarction, stroke, or transient ischaemic attack.

Biochemistry measures were performed at a dedicated central laboratory on around 480,000 samples between 2014 and 2017. These included serum total cholesterol and HDL cholesterol
(Beckman Coulter, UK on an AU5800 platform) and Lp(a) (Randox Bioscience, UK on an AU5800 platform). Data were adjusted by UK Biobank centrally before release to adjust for pre-analytical variables. For $\operatorname{Lp}(a)$, low, medium, and high-quality control materials ran with coefficients of variation of $\leq 6.1 \%$. Further details of these measurements can be found in the UK Biobank online showcase and protocol (http://www.ukbiobank.ac.uk). Lp(a) was reported in $\mathrm{nmol} / \mathrm{L}$ rather than $\mathrm{mg} / \mathrm{dl}$, reflecting the concentration of particles rather than the mass of particles. The minimum reported concentration of $\mathrm{Lp}(\mathrm{a})$ was $3.8 \mathrm{nmol} / \mathrm{L}$ and the maximum was $189 \mathrm{nmol} / \mathrm{L}$; participants who had levels below the lower level ( $\mathrm{n}=48,360$ ) were coded as having an Lp (a) concentration of $2.88 \mathrm{nmol} / \mathrm{L}$, and above the upper level ( $\mathrm{n}=34,195$ ) coded as $250 \mathrm{nmol} / \mathrm{L}$ for continuous analyses.

Date and cause of death were obtained from death certificates held by the National Health Service (NHS) Information Centre for participants from England and Wales and the NHS Central Register Scotland for participants from Scotland. Only primary causes of death listed on the death certificate were included in this analysis. Nonfatal outcomes were ascertained by linkage of participant study data to Hospital Episode Statistics (HES) from the National Health Service. The primary outcome of interest was ASCVD, a composite CVD outcome that reflects the ACC/AHA guidelines prediction score including death from CHD or stroke (ICD-10 I20-25, I60-64) or hospitalisation for myocardial infarction or stroke (ICD-10 I21, I22, I60, I61, I62, I64) (20). A secondary outcome was fatal cardiovascular disease as defined by primary cause of death from events included in the European SCORE clinical guidelines (I10-15, I44-51, I20-25, I61-73) (21). We also investigated secondary outcomes of fatal or nonfatal coronary heart disease (ICD10 I20-25, or operation code K40-49, K50, K75, U19)), peripheral vascular disease (ICD10 I70-78, or operation code K551), heart failure (ICD10 I150, I42.0, I42.6, I42.7, I42.9, I11.0), and ischaemic stroke (ICD10 I63, I64).

End of follow-up for each participant was recorded as the date of death or the date of end of follow-up for the assessment centre attended, whichever came first. The period at risk of each participant began on the date of their assessment.

## Statistical analyses

The association of continuous log-transformed $\mathrm{Lp}(\mathrm{a})$ with other lipid variables was tested using Pearson correlation coefficients. $\mathrm{Lp}(\mathrm{a})$ was analysed using a number of different models, reflecting existing uncertainty regarding cut-offs for "abnormal" levels. Lp(a) was
categorised into groups with cutoffs at $20 \mathrm{nmol} / \mathrm{L}$ and $100 \mathrm{nmol} / \mathrm{L}$ or $125 \mathrm{nmol} /$ or $150 \mathrm{nmol} / \mathrm{L}$ or $175 \mathrm{nol} / \mathrm{L}$. Sex and ethnicity specific centiles ( $50^{\text {th }}, 75^{\text {th }}, 80^{\text {th }}, 90^{\text {th }}$ and $95^{\text {th }}$ centiles) for Lp (a) were also developed, using binomial exact confidence intervals to yields $95 \%$ confidence intervals (CI) for the centiles. The sex and ethnicity specific $80^{\text {th }}$ centile was chosen to create a binary "high" category for $\operatorname{Lp}(a)$. Log transformed $\operatorname{Lp}(a)$ was also analysed as a continuous variable. Classical risk factors were expressed as mean (standard deviation) if symmetrically distributed, median (interquartile range) if skewed, and number (\%) if categorical. The distribution of classical risk factors by categories of outcome or exposure of interest were assessed using unpaired 2 tail t-tests, a Wilcoxon rank-sum, or a chi-squared test, respectively.

The cohort was analysed as a whole cohort, and was also stratified as a primary prevention cohort (participants without baseline CVD and not taking a statin) and as a high risk cohort (participants with baseline CVD and/or taking a statin). Rates of the primary composite CVD outcome were investigated in unadjusted models, splitting the cohort by the specified $\operatorname{Lp}(a)$ categories. Associations of continuous and categorical $\mathrm{Lp}(\mathrm{a})$ with outcomes of interest were investigated using Cox-proportional hazard models in the whole cohort adjusting for classical risk factors (age, sex, total cholesterol, HDL cholesterol, ethnicity, smoking, systolic blood pressure, blood pressure medications, baseline diabetes), and also adjusting for baseline CVD and baseline statin use in analyses including high risk participants. The proportional hazard assumption was checked by inspection of Schöenfeld residuals. We checked for interactions of $\operatorname{Lp}$ (a) effect with median age, sex, ethnicity, baseline CVD, statin use, and high total cholesterol (at $8.0 \mathrm{mmol} / \mathrm{L}$ ). Associations of $\mathrm{Lp}(\mathrm{a})$ with outcomes of interest were then conducted separately in the primary prevention and high risk cohorts. We explored the shape of the association of $\mathrm{Lp}(\mathrm{a})$ with outcomes using restricted cubic splines, and using linear and categorical models. Follow-up time was calculated as days from enrolment to death or the end of follow-up, whichever occurred first. Analyses were repeated in the group of people with baseline CVD or statin use at baseline, additionally adjusting for baseline CVD and statin use.

The ability of $\mathrm{Lp}(\mathrm{a})$ to improve prediction of CVD was tested by assessing improvement over a base model containing all elements from the Pooled Cohort Equation. The model for the composite CVD outcome tested whether $\operatorname{Lp}($ a) improved prediction after inclusion of classical risk factors of age, sex, ethnicity, total cholesterol, HDL cholesterol, systolic blood
pressure, diabetes, smoking, statin and blood pressure medication use. Improvement in prediction was tested using Harrells C-Index for survival data, testing for increased concordance after the addition of $\operatorname{Lp}($ a) to the model as either a continuous or categorical variable. We used a categorical net reclassification index (NRI), to investigate the changes in predicted risk classification after addition of $\mathrm{Lp}(\mathrm{a})$ to models, using a pre-specified binary $10-$ year CVD risk boundary of $7.5 \%$, reflecting high or low risk categorisation according to guidelines.

Population attributable fractions in the exposed, with $95 \%$ confidence intervals, were estimated using two adjusted Cox models and the punafcc postestimate command in STATA. The first model added $\mathrm{Lp}(\mathrm{a})$ as a binary variable (detectable/undetectable), and was designed to estimate the causal effect of $\operatorname{Lp}(\mathrm{a})$ in the whole cohort. The second model added $\mathrm{Lp}(\mathrm{a})$ as a 4-category variable with concentration $>175 \mathrm{nmol} / \mathrm{L}$ (representing the Lp (a) HORIZON trial recruitment criterion (6)), $40-175 \mathrm{nmol} / \mathrm{L}, 30-40 \mathrm{nmol} / \mathrm{L}$ (the estimated attained $\mathrm{Lp}(\mathrm{a}$ ) assuming an $80 \%$ reduction on-treatment (22)), and $<30 \mathrm{nmol} / \mathrm{L}$. This model was then used to test the estimated proportional reduction in CVD events among those with $\mathrm{Lp}(\mathrm{a})>175 \mathrm{nmol} / \mathrm{L}$ when concentration was lowered to the $30-40 \mathrm{nmol} / \mathrm{L}$ range. The model was run in the whole cohort and in those with baseline CVD.

All analyses were performed using STATA 14 (StataCorp LP) and R (version 3.5.1).

## Results

## Cross sectional associations

Of 502,624 people included in the study, complete data on covariates, including Lp(a) were available in 413,724 participants. Median $\mathrm{Lp}(\mathrm{a})$ in the cohort was $19.7 \mathrm{nmol} / \mathrm{L}$ (interquartile interval 7.6-75.3nmol/L) and in participants without baseline CVD and not taking a statin this was $19.1 \mathrm{nmol} / \mathrm{L}$ (interquartile interval $7.6-70.5 \mathrm{nmol} / \mathrm{L}$ ). The $80^{\text {th }}$ centile in the whole cohort was $104.5 \mathrm{nmol} / \mathrm{L}$ ( $95 \%$ CI 103.8-105.3). In the whole cohort, 85,930 (20.8\%) had $\mathrm{Lp}(\mathrm{a})$ above $100 \mathrm{nmol} / \mathrm{L}, 68,601(16.6 \%)$ above $125 \mathrm{nmol} / \mathrm{L}, 52,157$ ( $12.6 \%$ ) above $150 \mathrm{nmol} / \mathrm{L}$, and $38,109(9.2 \%)$ above $175 \mathrm{nmol} / \mathrm{L}$. Sex and ethnicity specific cut-offs show that women and participants with black ethnicity more frequently had elevated Lp(a) (Figure 1).
$\mathrm{Lp}(\mathrm{a})$ had weak positive associations with total cholesterol ( $\mathrm{r}=0.11$ ), HDL-cholesterol ( $\mathrm{r}=0.04$ ), and LDL cholesterol ( $\mathrm{r}=0.12$ ) ( p for all<0.0001). These associations were only
nominally stronger in the population who did not report taking statins $(\mathrm{r}=0.14,0.05,0.16$, respectively). Participants with elevated $\operatorname{Lp}(a)$ were generally slightly older, had slightly higher systolic blood pressure and total cholesterol, and were more likely to have baseline CVD (Table 1).

## Univariable association of $\operatorname{Lp}(a)$ with outcomes

Median follow-up time for the composite CVD outcome was 8.9 years (IQR 8.2-9.5) in both the whole cohort and among participants in the cohort without baseline CVD and not taking a statin. The composite CVD outcome occurred in 10,065 participants (2.4\%), and fatal CVD occurred in 3247 participants ( $0.8 \%$ ) in the whole cohort. The composite CVD outcome occurred in 6125 participants (1.8\%), and fatal CVD occurred in 1627 participants ( $0.5 \%$ ) in the subgroup without baseline CVD and not taking a statin.

Baseline Lp (a) was higher among the participants who went on to experience the composite CVD outcome and the fatal CVD outcome (Supplementary Table 1). Lp(a) was also higher among those who went on to experience CHD or peripheral vascular disease (PVD), but was not higher among those who went on to experience ischaemic stroke, heart failure (Supplementary Table 1). Rates of the primary composite CVD outcome were higher in participants with higher categories of baseline Lp(a) (Figure 1).

## Multivariable association of $\operatorname{Lp}(a)$ with outcomes

Among those without baseline CVD and not taking a statin, after adjusting for classical risk factors, the shape of the association of $\operatorname{Lp}(a)$ with composite CVD, fatal CVD, and CHD was positive and broadly linear (Figure 2). There was also a weaker positive association with PVD, but no evidence of an association with stroke, heart failure or atrial fibrillation (Figure 3)

In the whole cohort, there was an independent association of one standard deviation increase in $\log \operatorname{Lp}(a)$ with the primary composite CVD outcome (HR 1.09 (95\% CI 1.07-1.11) after adjusting for classical risk factors statin use and baseline CVD. There was no interaction with age (above or below the median of 57 years) ( $\mathrm{p}=0.11$ ), sex ( $\mathrm{p}=0.40$ ), ethnicity ( $\mathrm{p}>0.18$ for each ethnic group compared to white), baseline CVD ( $\mathrm{p}=0.12$ ), statin use ( $\mathrm{p}=0.23$ ), or total cholesterol (above or below $8.0 \mathrm{mmol} / \mathrm{L}$ cut-off) $(\mathrm{p}=0.34)$.

Although there was no formal interaction, for clarity data were split into the primary prevention cohort and the high risk cohort to further explore associations with outcomes. Using continuous and categorical models of $\mathrm{Lp}(\mathrm{a})$ there was a positive association with composite primary CVD outcome (Table 2). Lp(a) was also associated with fatal CVD, CHD and PVD as well as demonstrating a borderline association with ischaemic stroke after adjusting for classical risk factors, in both the primary prevention cohort and the high risk cohort (Supplementary Table 2). There was no association with heart failure (Supplementary Table 2).

## $L p(a)$ and prediction of $C V D$

Prediction of incident CVD was specifically explored in the primary prevention cohort. In a model of CVD prediction based on pooled cohort equation risk factors plus current treatments, classical risk factors yielded a C-index of 0.7439 ( $95 \%$ CI 0.7381-0.7497). Addition of $\operatorname{Lp}(a)$ as a continuous variable to these risk factors increased the C -index by $+0.0017(95 \% \mathrm{CI} 0.0009,0.0026)$ (Table 3). On addition of $\operatorname{Lp}(a)$ to the model the improvement in the categorical NRI was $1.39 \%$ ( $95 \%$ CI 0.69-2.12\%) and most of the improvement was due to up classification of risk among cases (Table 3). Similar small improvements in prediction were obtained when $\mathrm{Lp}(\mathrm{a})$ was added as a categorical variable, with no clear advantage of one model over another (Table 3).

## Population attributable fraction of $L p(a)$

After adjusting for confounders, the proportion of CVD attributable due to any detectable $\mathrm{Lp}(\mathrm{a})$ in the whole cohort was $8.6 \%$ ( $95 \%$ CI 7.4-9.8\%) ; this is the reduction in risk expected if the whole cohort had Lp (a) below the limit of detection.

In the whole cohort, an $\operatorname{Lp}(a)$ above $100 \mathrm{nmol} / \mathrm{L}$ accounted for $5.7 \%$ ( $95 \%$ CI 4.8, 6.6\%) of the composite CVD outcome (Figure 3). Lp(a) elevation above 100nmol/L accounted for $5.7 \% ~(4.0,7.3 \%)$ of fatal CVD, $6.7 \% ~(6.0,7.4 \%)$ of CHD, $2.5 \%$ ( $0.7,4.2 \%$ ) of stroke, $3.0 \%$ $(0.7,5.2 \%)$ of heart failure, and $8.3 \%(6.6,9.9 \%)$ of peripheral vascular disease. Of these outcomes, the only one where the $\operatorname{Lp}(a)$ attributable fraction varied substantially in the high risk cohort compared to the primary prevention cohort was peripheral vascular disease, with a population attributable fraction of $4.7 \%(2.4,7.0 \%)$ in the primary prevention cohort and $12.1 \%(9.6,14.6 \%)$ in the high risk cohort. Moving the threshold for "high" $\operatorname{Lp}(a)$ to higher cutoffs resulted in somewhat lower, but still substantial, attributable fractions due to reduced
prevalence of the higher cutoffs (Figure 3). The overall PAF lowered to 3.0\% (95\%CI 2.4$3.6 \%$ ) for a cutoff of $\mathrm{Lp}(\mathrm{a})$ above $175 \mathrm{nmol} / \mathrm{L}$ due to reduced prevalence.

## Expected benefit of Lp(a) reduction

We then specifically modelled the scenario in the ongoing phase 3 trial of an $\operatorname{Lp}(a)$ lowering agent. For all participants with an $\mathrm{Lp}(\mathrm{a})$ above $175 \mathrm{nmol} / \mathrm{L}$, reducing this by approximately $80 \%$ so that participants have Lp (a) $30-40 \mathrm{nmol} / \mathrm{L}$ range, yields an estimated risk reduction of $23.2 \%$ ( $95 \%$ CI 15.0-30.6\%). For participants with baseline CVD and an $\operatorname{Lp}(a)$ above $175 \mathrm{nmol} / \mathrm{L}$, similar Lp (a) reductions are estimated to reduce risk by $20.3 \%$ ( $95 \% \mathrm{CI} 2.5-$ $34.9 \%$ ).

## Discussion

In this large cohort of over 400,000 individuals, a high proportion of the UK biobank cohort had what might conventionally be called high Lp (a) levels; $20.8 \%$ above $100 \mathrm{nmol} / \mathrm{L}$ and $9.2 \%$ above $175 \mathrm{nmol} / \mathrm{L}$. We noted a broadly linear relationship of $\mathrm{Lp}(\mathrm{a})$ with composite fatal or nonfatal CVD, fatal CVD, and fatal or nonfatal CHD, with associations largely unaffected by other risk factors. Therefore, population attributable fractions for $\mathrm{Lp}(\mathrm{a})$ were notable. Further, we estimate that targeting Lp(a) lowering therapy in ongoing trials to those with Lp (a) concentrations above $175 \mathrm{nmol} / \mathrm{L}$ would reduce CVD incidence by around $20 \%$ (regardless of baseline CVD status). Collectively, our results seem to justify the recent ESC/EAS guidelines (11) suggesting consideration for at least a one-time Lp (a) measurement in all people being screened for cardiovascular disease.

Genetic data as well as basic science support the notion that Lp(a) causes CVD (4). Although genetic data suggested large reductions in $\operatorname{Lp}(a)$ would be required to show clinical benefit, recent phase 2 trial data show that the drug AKCEA-APO(a)-LRx (also called TQJ230) reduces $\operatorname{Lp}$ (a) substantially, with 80-90\% reductions in patients with established CVD and high $\operatorname{Lp}(a)$ levels, depending on dosing (22)(23). This antisense oligonucleotide inhibits the production of apolipoprotein(a), thereby reducing $\operatorname{Lp}(a)$. Phase 3 outcome trials are underway (6) and specifically target those with elevated $\operatorname{Lp}(a)$. If this or other drugs targeting $\mathrm{Lp}(\mathrm{a})$ prove successful in reducing CVD events, population screening of $\operatorname{Lp}($ a) may be one way to target those who are most likely to benefit. Our data suggest that a drug that prevents CVD through $\mathrm{Lp}(\mathrm{a})$ lowering may also have benefits for individual outcome components of the

CVD composite, and for PVD outcomes. Our models suggest potential benefit in both primary and secondary prevention. Although $\operatorname{Lp}(a)$ only adds moderate information to risk discrimination metrics, and is more expensive than traditional lipid measurements (24), the fact that the marker is i) causal, ii) largely orthogonal to other risk factors, iii) stable across lifecourse, iv) has a substantial population attributable fraction, and v) may help guide therapy allocation, enhances arguments that measurement of $\mathrm{Lp}(\mathrm{a})$ should become more common in the evaluation of CVD risk. Notably, the reported improvement in C-statistics with Lp (a) was around four times higher than previously reported for CRP (25).

The ethnicity-specific centiles we report confirm and extend observations in other cohorts, most noticeably higher Lp(a) in black people (26). However, we note no interaction of Lp(a) with demographics or other risk factors suggesting that $\mathrm{Lp}(\mathrm{a})$ is similarly associated with risk in different subgroups. We would expect a higher PAF for CVD outcomes in black ethnicities due to higher prevalence of the exposure, but low numbers of black participants restrict our ability to formally test this hypothesis in this cohort. The PAFs we report for $\mathrm{Lp}(\mathrm{a})$ may usefully be put in context of estimates of PAFs for other risk factors. The ARIC study reported PAFs at examination four (among white participants) of $21 \%$ for hypertension, 13\% for diabetes, $10 \%$ for hypercholesterolaemia and $12 \%$ for smoking (27). Similarly, in the Emerging Risk Factors Collaboration, the PAF for diabetes in vascular death has been estimated at $11 \%$ (assuming a $10 \%$ diabetes prevalence (28)), In terms of CVD risk prediction, our data are also broadly in line with the Emerging Risk Factors Collaboration (ERFC), where data from 165,544 participants in 37 prospective studies showed Lp(a) improved the C-index by +0.0016 ( $95 \% \mathrm{CI}, 0.0009-0.0023$ ) (29), lending strong external validity to these new reported findings. It is also in agreement with data from other large cohort studies (30-33). Our data extend these findings in a large cohort with substantial power, using a single method of $\operatorname{Lp}(a)$ measurement, where we also estimated the PAF for CVD outcomes independently due to $\operatorname{Lp}(a)$ elevation.

## Study strengths and limitations

The strengths of our study include the large cohort size at an age relevant to CVD risk scoring, and biochemistry assays performed in a single dedicated central laboratory. We were also able to extensively adjust our models for classical risk factors and separately analysed participants already on statins as well as those with previous CVD. We were also able to look at other cardiac outcomes such as heart failure. Potential limitations include the relatively low
average CVD risk of participants, although risk prediction models performed broadly in line with expectations. UK Biobank is not representative of the whole UK population (30), and while this is generally not a concern in investigating risk associations (34) it will have an impact on calculated population attributable fractions. The population attributable fractions we observe here cannot be taken as representative of the UK population as a whole. However, due to the under-representation of black people in UK Biobank, it may be that our estimates are conservative. In addition, the $80^{\text {th }}$ centile we report here of $105 \mathrm{nmol} / \mathrm{L}$ corresponds broadly to previously reported $80^{\text {th }}$ centiles in 3000 men and 3000 women from the Copenhagen General Population Study ( $50 \mathrm{mg} / \mathrm{dl}$ ) (9). Finally, our estimates of expected therapy effects among the exposed ( Lp (a) above $175 \mathrm{nmol} / \mathrm{L}$ ) should be robust to differences in representativeness, since we only consider those with high $\operatorname{Lp}(a)$ as exposed.

## Conclusion

In conclusion, in this, the largest single prospective study of Lp (a) levels, our findings add strong support for recent guideline recommended one-time measurement of $\mathrm{Lp}(\mathrm{a})$ in CV risk assessment to identify a large proportion with markedly elevated levels sufficient to contribute to atherothrombotic risk. Our work also provides support to the ongoing programmes to develop efficacious $\operatorname{Lp}(a)$ lowering drugs

## References

1. Willeit P, Ridker PM, Nestel PJ, et al. Baseline and on-statin treatment lipoprotein(a) levels for prediction of cardiovascular events: individual patient-data meta-analysis of statin outcome trials. Lancet (London, England) 2018;392:1311-1320.
2. Waldeyer C, Makarova N, Zeller T, et al. Lipoprotein(a) and the risk of cardiovascular disease in the European population: results from the BiomarCaRE consortium. Eur. Heart J. 2017;38:2490-2498.
3. Cook NR, Mora S, Ridker PM. Lipoprotein(a) and Cardiovascular Risk Prediction Among Women. J. Am. Coll. Cardiol. 2018;72:287-296.
4. Burgess S, Ference BA, Staley JR, et al. Association of LPA Variants With Risk of Coronary Disease and the Implications for Lipoprotein(a)-Lowering Therapies. JAMA Cardiol. 2018;3:619.
5. Tsimikas S, Fazio S, Ferdinand KC, et al. NHLBI Working Group Recommendations to Reduce Lipoprotein(a)-Mediated Risk of Cardiovascular Disease and Aortic Stenosis. J. Am. Coll. Cardiol. 2018;71:177-192.
6. Anon. Assessing the Impact of Lipoprotein (a) Lowering With TQJ230 on Major Cardiovascular Events in Patients With CVD - Full Text View - ClinicalTrials.gov.
7. O’Donoghue ML, Fazio S, Giugliano RP, et al. Lipoprotein(a), PCSK9 Inhibition, and Cardiovascular Risk. Circulation 2019;139:1483-1492.
8. Wilson DP, Jacobson TA, Jones PH, et al. Use of Lipoprotein(a) in clinical practice: A biomarker whose time has come. A scientific statement from the National Lipid Association.
J. Clin. Lipidol. 2019;13:374-392.
9. Nordestgaard BG, Chapman MJ, Ray K, et al. Lipoprotein(a) as a cardiovascular risk factor: current status. Eur. Heart J. 2010;31:2844-2853.
10. Reiner Z, Capatano A, De Backer G, Graham I. ESC/EAS Guidelines for the management of dyslipidemias.
11. Mach F, Baigent C, Catapano AL, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. Eur. Heart J. 2019.
12. Ellis KL, Pérez de Isla L, Alonso R, Fuentes F, Watts GF, Mata P. Value of Measuring Lipoprotein(a) During Cascade Testing for Familial Hypercholesterolemia. J. Am. Coll. Cardiol. 2019;73:1029-1039.
13. Grundy SM, Stone NJ, Bailey AL, et al. 2018

AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of

Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation 2019;139:e1082-e1143.
14. Kamstrup PR, Tybjaerg-Hansen A, Steffensen R, Nordestgaard BG. Genetically Elevated Lipoprotein(a) and Increased Risk of Myocardial Infarction. JAMA 2009;301:2331. 15. Anderson TJ, Grégoire J, Pearson GJ, et al. 2016 Canadian Cardiovascular Society Guidelines for the Management of Dyslipidemia for the Prevention of Cardiovascular Disease in the Adult. Can. J. Cardiol. 2016;32:1263-1282.
16. Anon. Assessing the Impact of Lipoprotein (a) Lowering With TQJ230 on Major Cardiovascular Events in Patients With CVD (LP(a)HORIZON).
17. Collins R. What makes UK Biobank special? Lancet 2012;379:1173-4.
18. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLoS Med. 2015;12:e1001779.
19. Elliott P, Peakman TC. The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. Int. J. Epidemiol. 2008;37:234-244.
20. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk. Circulation 2014;129:S49-S73.
21. Rücker V, Keil U, Fitzgerald AP, et al. Predicting 10-Year Risk of Fatal Cardiovascular Disease in Germany: An Update Based on the SCORE-Deutschland Risk Charts Lazzeri C, editor. PLoS One 2016;11:e0162188.
22. Viney NJ, van Capelleveen JC, Geary RS, et al. Antisense oligonucleotides targeting apolipoprotein(a) in people with raised lipoprotein(a): two randomised, double-blind, placebo-controlled, dose-ranging trials. Lancet 2016;388:2239-2253.
23. Langsted A, Nordestgaard BG. Antisense Oligonucleotides Targeting Lipoprotein(a). Curr. Atheroscler. Rep. 2019;21:30.
24. Tsimikas S. A Test in Context: Lipoprotein(a): Diagnosis, Prognosis, Controversies, and Emerging Therapies. J. Am. Coll. Cardiol. 2017;69:692-711.
25. Emerging Risk Factors Collaboration, Kaptoge S, Di Angelantonio E, et al. C-reactive protein, fibrinogen, and cardiovascular disease prediction. N. Engl. J. Med. 2012;367:13101320.
26. Marcovina SM, Albers JJ. Lipoprotein (a) measurements for clinical application. J. Lipid Res. 2016;57:526-37.
27. Cheng S, Claggett B, Correia AW, et al. Temporal trends in the population attributable
risk for cardiovascular disease: the Atherosclerosis Risk in Communities Study. Circulation 2014;130:820-8.
28. Emerging Risk Factors Collaboration, Sarwar N, Gao P, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet (London, England) 2010;375:2215-2222.
29. Emerging Risk Factors Collaboration, Di Angelantonio E, Gao P, et al. Lipid-related markers and cardiovascular disease prediction. JAMA 2012;307:2499-2506.
30. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and HealthRelated Characteristics of UK Biobank Participants With Those of the General Population. Am. J. Epidemiol. 2017;186:1026-1034.
31. Kamstrup PR, Tybjærg-Hansen A, Nordestgaard BG. Extreme Lipoprotein(a) Levels and Improved Cardiovascular Risk Prediction. J. Am. Coll. Cardiol. 2013;61:1146-1156.
32. Willeit P, Kiechl S, Kronenberg F, et al. Discrimination and Net Reclassification of Cardiovascular Risk With Lipoprotein(a). J. Am. Coll. Cardiol. 2014;64:851-860.
33. Verbeek R, Sandhu MS, Hovingh GK, et al. Lipoprotein(a) Improves Cardiovascular Risk Prediction Based on Established Risk Algorithms. J. Am. Coll. Cardiol. 2017;69:15131515.
34. Keyes KM, Westreich D. UK Biobank, big data, and the consequences of nonrepresentativeness. Lancet 2019;393:1297.

## Figure Legends

Figure 1 Centiles of Lp(a), along with $95 \%$ confidence intervals, by sex and ethnicity in the whole cohort. Red denotes white ethnicity, green denotes black ethnicity, blue denotes South Asian, and purple denotes other or mixed ethnicity.

Figure 2 Association of $\operatorname{Lp}(a)$ with outcomes after adjusting for classical risk factors among participants without baseline CVD and not taking a statin.

Figure 3 Population attributable fractions (with $95 \%$ confidence intervals) of $\operatorname{Lp}(a)$ over $100 \mathrm{nmol} / \mathrm{L}$ for each outcome of interest within the whole cohort (orange), the primary prevention cohort (yellow) and the high risk cohort (history of CVD or taking a statin; green).

Table 1 Association of categories of $\operatorname{Lp}(a)$ with classical risk factors for CVD at baseline ( $\mathrm{n}=413,724$ )

|  | $\begin{gathered} \text { Lp(a): } \\ \text { <20nmol/L } \end{gathered}$ | $\begin{gathered} \text { Lp(a) } \\ \text { 20-99.9nmol/L } \end{gathered}$ | $\begin{gathered} \mathrm{Lp}(\mathrm{a}) \\ \geq 100 \mathrm{nmol} / \mathrm{L} \end{gathered}$ | P-value for trend |
| :---: | :---: | :---: | :---: | :---: |
|  | $\mathrm{n}=207,903$ | $\mathrm{n}=119,891$ | $\mathrm{n}=85,930$ |  |
| Age (years) | $56.3 \pm 8.2$ | $56.7 \pm 8.1$ | $56.9 \pm 8.0$ | <0.001 |
| Male sex (\%) | $\begin{aligned} & 101,212 \\ & (48.7 \%) \end{aligned}$ | 52,539 (43.8\%) | $\begin{gathered} 37,402 \\ (43.5 \%) \end{gathered}$ | <0.001 |
| Ethnicity (\%) |  |  |  |  |
| White | $\begin{aligned} & 201,066 \\ & (96.7 \%) \end{aligned}$ | 109949 (91.7\%) | $\begin{gathered} 80953 \\ (94.2 \%) \end{gathered}$ | <0.001 |
| Black | 717 (0.3\%) | 3327 (2.8\%) | 2404 (2.8\%) |  |
| South Asian | 2426 (1.2\%) | 2931 (2.4\%) | 1064 (1.2\%) |  |
| Other | 3694 (1.8\%) | 3684 (3.1\%) | 1509 (1.8\%) |  |
| Systolic blood pressure ( mmHg ) | $136.8 \pm 19.6$ | $139.7 \pm 19.7$ | $140.3 \pm 19.7$ | <0.001 |
| Diastolic blood pressure ( mmHg ) | $82.3 \pm 10.7$ | $82.2 \pm 10.7$ | $82.4 \pm 10.7$ | 0.026 |
| Ever smoker (\%) | $\begin{gathered} 22,088 \\ (10.6 \%) \end{gathered}$ | 12,761 (10.6\%) | 8,994 (10.5\%) | 0.37 |
| Total cholesterol ( $\mathrm{mmol} / \mathrm{L}$ ) | 5.59 (1.12) | 5.75 (1.15) | 5.86 (1.18) | <0.001 |
| HDL cholesterol ( $\mathrm{mmol} / \mathrm{L}$ ) | 1.44 (0.39) | 1.45 (0.38) | 1.48 (0.38) | $<0.001$ |
| Baseline diabetes (\%) | 11,462 (5.5\%) | 5764 (4.8\%) | 4453 (5.2\%) | <0.001 |
| Baseline CVD (\%) | 11,133 (5.4\%) | 6900 (5.8\%) | 6613 (7.7\%) | <0.001 |
| Statin use (\%) | $\begin{gathered} 31,228 \\ (15.0 \%) \end{gathered}$ | 18,620 (15.5\%) | $\begin{gathered} 17,374 \\ (20.2 \%) \end{gathered}$ | <0.001 |
| BP medication (\%) | $\begin{aligned} & 42,157 \\ & (20.3 \%) \end{aligned}$ | 24,578 (20.5\%) | $\begin{gathered} 19,569 \\ (22.8 \%) \end{gathered}$ | $<0.001$ |

BP: blood pressure; CVD: cardiovascular disease; HDL: high-density lipoproteins; mmHg : millimetres of mercury; mmol/L: millimoles per litre; nmol/L: nanomoles per litre

Table 2 Association of $\operatorname{Lp}($ a) (hazard ratio and $95 \% \mathrm{CI}$ ) as a continuous variable (per 1 SD increase in $\log \operatorname{Lp}(a))$ and as a categorical variable with the primary composite CVD outcome after adjusting for classical risk factors*.

| Lp(a) conc. | Prima | ention cohort | High | hort*** |
| :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \mathrm{N} \\ \text { (n events) } \end{gathered}$ | HR (95\% CI) | $\begin{gathered} \mathrm{N} \\ \text { (n events) } \end{gathered}$ | HR (95\% CI) |
| Per 1SD | $\begin{aligned} & 340333 \\ & (6125) \end{aligned}$ | $\begin{gathered} 1.09 \\ (1.07-1.11) \end{gathered}$ | 73391 (3940) | $\begin{gathered} 1.07 \\ (1.05-1.09) \end{gathered}$ |
| <20nmol/L | $\begin{aligned} & 173636 \\ & (2857) \end{aligned}$ | Ref | $\begin{aligned} & 34267 \\ & (1688) \end{aligned}$ | Ref |
| 20-99.9nmol/L | $\begin{aligned} & 99496 \\ & (1802) \end{aligned}$ | $\begin{gathered} 1.11 \\ (1.04-1.18) \end{gathered}$ | $\begin{aligned} & 20395 \\ & (1086) \end{aligned}$ | $\begin{gathered} 1.09 \\ (1.01-1.17) \end{gathered}$ |
| $\geq 100 \mathrm{nmol} / \mathrm{L}$ | $\begin{aligned} & 67201 \\ & (1466) \end{aligned}$ | $\begin{gathered} 1.36 \\ (1.28-1.45) \end{gathered}$ | $\begin{aligned} & 18729 \\ & (1166) \end{aligned}$ | $\begin{gathered} 1.28 \\ (1.18-1.37) \end{gathered}$ |
| <20nmol/L | $\begin{aligned} & 173636 \\ & (2857) \end{aligned}$ | Ref | $\begin{aligned} & 34267 \\ & (1688) \end{aligned}$ | Ref |
| 20-124.9nmol/L | $\begin{aligned} & 114090 \\ & (2086) \end{aligned}$ | $\begin{gathered} 1.12 \\ (1.06-1.19) \end{gathered}$ | $\begin{aligned} & 23130 \\ & (1241) \end{aligned}$ | $\begin{gathered} 1.09 \\ (1.01-1.18) \end{gathered}$ |
| $\geq 125 \mathrm{nmol} / \mathrm{L}$ | $\begin{aligned} & 52607 \\ & (1182) \end{aligned}$ | $\begin{gathered} 1.40 \\ (1.31-1.50) \end{gathered}$ | $\begin{aligned} & 15994 \\ & (1011) \end{aligned}$ | $\begin{gathered} 1.31 \\ (1.21-1.41) \end{gathered}$ |
| <20nmol/L | $\begin{aligned} & 173636 \\ & (2857) \end{aligned}$ | Ref | $\begin{aligned} & 34267 \\ & (1688) \end{aligned}$ | Ref |
| $\begin{aligned} & 20- \\ & 149.9 .9 \mathrm{nmol} / \mathrm{L} \end{aligned}$ | $\begin{aligned} & 127633 \\ & (2371) \end{aligned}$ | $\begin{gathered} 1.14 \\ (1.08-1.21) \end{gathered}$ | $\begin{aligned} & 26031 \\ & (1428) \end{aligned}$ | $\begin{gathered} 1.11 \\ (1.03-1.19) \end{gathered}$ |
| $\geq 150 \mathrm{nmol} / \mathrm{L}$ | $\begin{gathered} 39064 \\ (897) \end{gathered}$ | $\begin{gathered} 1.43 \\ (1.33-1.55) \end{gathered}$ | $\begin{gathered} 13093 \\ (824) \end{gathered}$ | $\begin{gathered} 1.32 \\ (1.21-1.43) \end{gathered}$ |
| <20nmol/L | $\begin{aligned} & 173636 \\ & (2857) \end{aligned}$ | Ref | $\begin{aligned} & 34267 \\ & (1688) \end{aligned}$ | Ref |
| 20-174.9nmol/L | $\begin{aligned} & 138833 \\ & (2626) \end{aligned}$ | $\begin{gathered} 1.16 \\ (1.10-1.23) \end{gathered}$ | $\begin{aligned} & 28879 \\ & (1608) \end{aligned}$ | $\begin{gathered} 1.13 \\ (1.05-1.21) \end{gathered}$ |
| $\geq 175 \mathrm{nmol} / \mathrm{L}$ | $\begin{gathered} 27864 \\ (642) \end{gathered}$ | $\begin{gathered} 1.44 \\ (1.32-1.58) \end{gathered}$ | $\begin{aligned} & 10245 \\ & (644) \end{aligned}$ | $\begin{gathered} 1.33 \\ (1.21-1.45) \end{gathered}$ |
| Below the sex and ethnicity specific $80^{\text {th }}$ centile | $\begin{aligned} & 275886 \\ & (4672) \end{aligned}$ | Ref | $\begin{aligned} & 54884 \\ & (2767) \end{aligned}$ | Ref |
| Above the sex and ethnicity specific $80^{\text {th }}$ centile | $\begin{aligned} & 64447 \\ & (1453) \end{aligned}$ | $\begin{gathered} 1.31 \\ (1.24-1.39) \end{gathered}$ | $\begin{aligned} & 18507 \\ & (1173) \end{aligned}$ | $\begin{gathered} 1.24 \\ (1.16-1.33) \end{gathered}$ |

*age, sex, total cholesterol, HDL cholesterol, ethnicity, smoking, systolic blood pressure, blood pressure medications, baseline diabetes
** Participants with baseline CVD or taking a statin; additionally adjusted for baseline CVD and statin use
CI: confidence interval; CVD: cardiovascular disease; HDL: high-density lipoproteins; HR: hazard ratio; $\mathrm{Lp}(\mathrm{a})$ : Lipoprotein (a); nmol/L: nanomoles per litre; SD: standard deviation

Table 3 Improvement in prediction of CVD among participants in the primary prevention cohort measured by the C-statistic and categorical net reclassification index (across the 7.5\% 10 -year risk boundary).

| Model | C-Index | Change in <br> C-Index | Overall <br> NRI | Case NRI | Control NRI |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Classical risk factors* | (0.7439 |  |  |  |  |
|  | $(0.7381$, |  |  |  |  |
| Classical risk factors + | $0.7497)$ |  |  |  |  |
| continuous log Lp(a) |  | +0.0017 | $+1.39 \%$ | $+1.46 \%$ | $-0.08 \%$ |
|  | $(0.0009$, | $(0.69$, | $(0.77$, | $(-0.13,-$ |  |
| Classical risk factors + | $0.0026)$ | $2.12 \%)$ | $2.19 \%)$ | $0.03 \%)$ |  |
| categorical Lp(a) at <20, | +0.0018 | $+0.98 \%$ | $+1.07 \%$ | $-0.09 \%$ |  |
| 20-99.9, and >100nmol/L |  | $(0.0009$, | $(0.26$, | $(0.37$, | $(-0.14,-$ |
| Classical risk factors + | $0.0026)$ | $1.71 \%)$ | $1.79 \%)$ | $0.05 \%)$ |  |
| categorical Lp(a) at <20, |  | +0.0016 | $+1.15 \%$ | $+1.23 \%$ | $-0.08 \%$ |
| 20-149.9, and >150nmol/L |  | $(0.0007$, | $(0.42$, | $(0.51$, | $(-0.13,-$ |
| Classical risk factors + | $0.0024)$ | $1.88 \%)$ | $1.96 \%)$ | $0.03 \%)$ |  |
| binary Lp(a) at sex and | +0.0015 | $+1.15 \%$ | $+1.21 \%$ | $-0.06 \%$ |  |
| ethnicity specific 80 th | $(0.0007$, | $(0.52$, | $(0.57$, | $(-0.11,-$ |  |
| percentile | $0.0023)$ | $1.80 \%)$ | $1.86 \%)$ | $0.01 \%)$ |  |

*age, sex, total cholesterol, HDL cholesterol, ethnicity, smoking, systolic blood pressure, blood pressure medications, baseline diabetes
CVD: cardiovascular disease; HDL: high-density lipoproteins; Lp(a): Lipoprotein (a); $\mathrm{nmol} / \mathrm{L}$ : nanomoles per litre; NRI: net reclassification index

Figure 1 Centiles of $\operatorname{Lp}(a)$, along with $95 \%$ confidence intervals, by sex and ethnicity in the whole cohort. Red denotes white ethnicity, green denotes black ethnicity, blue denotes South Asian, and purple denotes other or mixed ethnicity.



Figure 2 Association of $\operatorname{Lp}(a)$ with outcomes after adjusting for classical risk factors among participants without baseline CVD and not taking a statin.







Figure 3 Population attributable fractions (with $95 \%$ confidence intervals) of $\operatorname{Lp}(a)$ over a range of thresholds for each outcome of interest within the whole cohort.


Supplementary Table 1 Association of $\operatorname{Lp}(a)(\mathrm{nmol} / \mathrm{L})$ with outcomes of interest
$\left.\begin{array}{|lccc|}\hline \text { Cohort/Outcome } & \begin{array}{c}\text { No incident } \\ \text { event }\end{array} & \text { Incident event } & \begin{array}{c}\text { P-value } \\ \text { for trend }\end{array} \\ \hline \begin{array}{l}\text { Cohort with no baseline CVD or statins } \\ \text { (n=340,333) }\end{array} & & & \\ \hline \text { Composite CVD } & \mathrm{n}=334,208 & \mathrm{n}=6125 & <0.001 \\ & 19.0(7.6,70.1) & 23.0(8.6,94.0) & \\ \text { Fatal CVD } & \mathrm{n}=338,706 & \mathrm{n}=1627\end{array}\right)<0.001$

| Cohort with baseline CVD or prescribed statins ( $\mathrm{n}=73,391$ ) |  |  |  |
| :---: | :---: | :---: | :---: |
| Composite CVD | $\begin{gathered} \mathrm{n}=69,451 \\ 23.3(7.9,101.9) \end{gathered}$ | $\begin{gathered} \mathrm{n}=3940 \\ 30.1(8.8,127.9) \end{gathered}$ | <0.001 |
| Fatal CVD | $\begin{gathered} \mathrm{n}=71,771 \\ 23.5(7.9,102.8) \end{gathered}$ | $\begin{gathered} \mathrm{n}=1620 \\ 28.2(8.6,127.7) \end{gathered}$ | <0.001 |
| CHD | $\begin{gathered} \mathrm{n}=65,672 \\ 23.0(7.8,100.2) \end{gathered}$ | $\begin{gathered} \mathrm{n}=7719 \\ 30.3(8.8,128.2) \end{gathered}$ | <0.001 |
| Ischaemic stroke | $\begin{gathered} \mathrm{n}=72,085 \\ 23.6(7.9,103.2) \end{gathered}$ | $\begin{gathered} \mathrm{n}=1306 \\ 26.8(8.4,110.8) \end{gathered}$ | 0.091 |
| Heart failure | $\begin{gathered} \mathrm{n}=72,344 \\ 23.6(8.0,103.3) \end{gathered}$ | $\begin{gathered} \mathrm{n}=1047 \\ 24.7(7.4,112.0) \end{gathered}$ | 0.90 |
| Peripheral vascular disease | $\begin{gathered} \mathrm{n}=72,050 \\ 23.5(7.9,102.4) \end{gathered}$ | $\begin{gathered} \mathrm{n}=1341 \\ 37.6(9.1,153.5) \end{gathered}$ | $<0.001$ |

Values are n , median $\mathrm{Lp}(\mathrm{a})$ and (Q1-Q3)
CHD: coronary heart disease; CVD: cardiovascular disease; Lp(a): Lipoprotein (a)

Supplementary Table 2 Association of $\operatorname{Lp}$ (a) (hazard ratio and $95 \% \mathrm{CI}$ ) as a continuous variable (per 1 SD increase in $\log \mathrm{Lp}(\mathrm{a})$ ) and as a categorical variable with secondary outcomes of interest after adjusting for classical risk factors in stepped models.

| $\mathrm{Lp}(\mathrm{a})$ conc. | Outcome | Primary prevention cohort |  | High risk cohort |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | N (n events) | HR (95\% CI) | N (n events) | HR (95\% CI) |
| Per 1SD | Fatal CVD | $\begin{gathered} 340333 \\ (1627) \\ \hline \end{gathered}$ | $\begin{gathered} 1.10 \\ (1.06-1.14) \end{gathered}$ | $\begin{aligned} & \hline 73391 \\ & (1620) \end{aligned}$ | $\begin{gathered} 1.07 \\ (1.04-1.11) \end{gathered}$ |
| <20nmol/L | Fatal CVD | $\begin{gathered} 173636 \\ (762) \\ \hline \end{gathered}$ | Ref | $\begin{aligned} & 34267 \\ & (698) \\ & \hline \end{aligned}$ | Ref |
| 20-99.9nmol/L |  | $\begin{gathered} 99496 \\ (480) \\ \hline \end{gathered}$ | $\begin{gathered} 1.14 \\ (1.01-1.28) \\ \hline \end{gathered}$ | $\begin{gathered} 20395 \\ (449) \\ \hline \end{gathered}$ | $\begin{gathered} 1.12 \\ (0.99-1.26) \\ \hline \end{gathered}$ |
| $\geq 100 \mathrm{nmol} / \mathrm{L}$ |  | $\begin{gathered} 67201 \\ (385) \\ \hline \end{gathered}$ | $\begin{gathered} 1.41 \\ (1.25-1.60) \end{gathered}$ | $\begin{aligned} & 18729 \\ & (473) \end{aligned}$ | $\begin{gathered} 1.27 \\ (1.13-1.43) \\ \hline \end{gathered}$ |
| <20nmol/L | Fatal CVD | $\begin{gathered} 173636 \\ (762) \end{gathered}$ | Ref | $\begin{aligned} & 34267 \\ & (698) \end{aligned}$ | Ref |
| 20-124.9nmol/L |  | $\begin{gathered} 114090 \\ (558) \\ \hline \end{gathered}$ | $\begin{gathered} 1.16 \\ (1.04-1.29) \\ \hline \end{gathered}$ | $\begin{gathered} 23130 \\ (507) \\ \hline \end{gathered}$ | $\begin{gathered} 1.10 \\ (0.98-1.24) \\ \hline \end{gathered}$ |
| $\geq 125 \mathrm{nmol} / \mathrm{L}$ |  | $\begin{gathered} 52607 \\ (307) \end{gathered}$ | $\begin{gathered} 1.44 \\ (1.26-1.65) \end{gathered}$ | $\begin{aligned} & 15994 \\ & (415) \end{aligned}$ | $\begin{gathered} 1.32 \\ (1.17-1.49) \end{gathered}$ |
| <20nmol/L | Fatal CVD | $\begin{gathered} 173636 \\ (762) \end{gathered}$ | Ref | $\begin{aligned} & 34267 \\ & (698) \end{aligned}$ | Ref |
| 20-149.9.9nmol/L |  | $\begin{gathered} 127633 \\ (624) \\ \hline \end{gathered}$ | $\begin{gathered} 1.16 \\ (1.04-1.29) \\ \hline \end{gathered}$ | $\begin{gathered} 26031 \\ (584) \\ \hline \end{gathered}$ | $\begin{gathered} 1.12 \\ (1.00-1.25) \\ \hline \end{gathered}$ |
| $\geq 150 \mathrm{nmol} / \mathrm{L}$ |  | $\begin{gathered} 39064 \\ (241) \\ \hline \end{gathered}$ | $\begin{gathered} 1.54 \\ (1.33-1.78) \\ \hline \end{gathered}$ | $\begin{aligned} & 12755 \\ & (338) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.33 \\ (1.17-1.52) \\ \hline \end{gathered}$ |
| <20nmol/L | Fatal CVD | $\begin{gathered} 173636 \\ (762) \\ \hline \end{gathered}$ | Ref | $\begin{array}{r} 34267 \\ (698) \\ \hline \end{array}$ | Ref |
| 20-174.9.9nmol/L |  | $\begin{gathered} 138833 \\ (692) \\ \hline \end{gathered}$ | $\begin{gathered} 1.18 \\ (1.07-1.31) \\ \hline \end{gathered}$ | $\begin{array}{r} 28879 \\ (658) \\ \hline \end{array}$ | $\begin{gathered} 1.14 \\ (1.02-1.26) \\ \hline \end{gathered}$ |
| $\geq 175 \mathrm{nmol} / \mathrm{L}$ |  | $\begin{gathered} 27864 \\ (173) \\ \hline \end{gathered}$ | $\begin{gathered} 1.57 \\ (1.32-1.85) \\ \hline \end{gathered}$ | $\begin{aligned} & 10245 \\ & (264) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.35 \\ (1.17-1.56) \\ \hline \end{gathered}$ |
| Above the sex and ethnicity specific $80^{\text {th }}$ centile | Fatal CVD | $\begin{gathered} 64447 \\ (386) \end{gathered}$ | $\begin{gathered} 1.36 \\ (1.21-1.52) \end{gathered}$ | $\begin{aligned} & 18507 \\ & (479) \end{aligned}$ | $\begin{gathered} 1.24 \\ (1.11-1.38) \end{gathered}$ |
| Per 1SD | CHD | $\begin{gathered} 340333 \\ (8930) \\ \hline \end{gathered}$ | $\begin{gathered} 1.13 \\ (1.11-1.15) \end{gathered}$ | $\begin{aligned} & \hline 73391 \\ & (7719) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.08 \\ (1.06-1.09) \end{gathered}$ |
| <20nmol/L | CHD | $\begin{gathered} 173636 \\ (3993) \\ \hline \end{gathered}$ | Ref | $\begin{array}{r} 34267 \\ (3283) \\ \hline \end{array}$ | Ref |
| 20-99.9nmol/L |  | $\begin{aligned} & 99496 \\ & (2682) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.17 \\ (1.11-1.23) \end{gathered}$ | $\begin{aligned} & 20395 \\ & (2151) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.09 \\ (1.04-1.16) \end{gathered}$ |
| $\geq 100 \mathrm{nmol} / \mathrm{L}$ |  | $\begin{aligned} & 67201 \\ & (2255) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.50 \\ (1.42-1.58) \\ \hline \end{gathered}$ | $\begin{array}{r} 18729 \\ (2285) \\ \hline \end{array}$ | $\begin{gathered} 1.29 \\ (1.22-1.36) \\ \hline \end{gathered}$ |
| <20nmol/L | CHD | $\begin{aligned} & 173636 \\ & (3993) \\ & \hline \end{aligned}$ | Ref | $\begin{array}{r} 34267 \\ (3283) \\ \hline \end{array}$ | Ref |

$\left.\begin{array}{|l|c|c|c|c|c|}\hline & & 110979 & 1.19 & 23130 & 1.10 \\ 20-124.9 \mathrm{nmol} / \mathrm{L} & & (3111) & (1.14-1.25) & (2465) & (1.04-1.16) \\ \hline & & 50781 & 1.55 & 15994 & 1.32 \\ \geq 125 \mathrm{nmol} / \mathrm{L} & & & (1826) & (1.46-1.64) & (1971)\end{array}\right]\left(\begin{array}{l}1.24-1.39) \\ \hline\end{array}\right.$

|  |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Above the sex and <br> ethnicity specific <br> $80^{\text {th }}$ centile | Ischaemic <br> stroke | 64447 <br> $(396)$ | 1.15 <br> $(1.03-1.29)$ | 18507 <br> $(356)$ | $(0.98-1.25)$ |
|  |  |  |  |  |  |
| Per 1SD | Heart <br> failure | 340333 <br> $(838)$ | 1.06 | 73391 | $(1.01-1.12)$ |


|  |  | (473) | (1.03-1.31) | (393) | (0.96-1.25) |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\geq 125 \mathrm{nmol} / \mathrm{L}$ |  | $\begin{aligned} & 52607 \\ & (263) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.46 \\ (1.26-1.69) \\ \hline \end{gathered}$ | $\begin{aligned} & 15601 \\ & (393) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.62 \\ (1.42-1.84) \\ \hline \end{gathered}$ |
| <20nmol/L | PVD | $\begin{gathered} 173636 \\ (639) \\ \hline \end{gathered}$ | Ref | $\begin{aligned} & 34267 \\ & (555) \\ & \hline \end{aligned}$ | Ref |
| 20-149.9.9nmol/L |  | $\begin{gathered} 127633 \\ (539) \\ \hline \end{gathered}$ | $\begin{gathered} \hline 1.19 \\ (1.06-1.33) \\ \hline \end{gathered}$ | $\begin{aligned} & 26031 \\ & (442) \\ & \hline \end{aligned}$ | $\begin{gathered} \hline 1.09 \\ (0.96-1.24) \\ \hline \end{gathered}$ |
| $\geq 150 \mathrm{nmol} / \mathrm{L}$ |  | $\begin{aligned} & 39064 \\ & (197) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.48 \\ (1.26-1.74) \end{gathered}$ | $\begin{aligned} & 12755 \\ & (344) \end{aligned}$ | $\begin{gathered} 1.76 \\ (1.54-2.02) \end{gathered}$ |
| <20nmol/L | PVD | $\begin{gathered} \hline 173636 \\ (639) \\ \hline \end{gathered}$ | Ref | $\begin{aligned} & 34267 \\ & (555) \\ & \hline \end{aligned}$ | Ref |
| 20-174.9nmol/L |  | $\begin{gathered} 127633 \\ (539) \end{gathered}$ | $\begin{gathered} \hline 1.19 \\ (1.06-1.34) \end{gathered}$ | $\begin{gathered} 28879 \\ (505) \\ \hline \end{gathered}$ | $\begin{gathered} \hline 1.12 \\ (0.99-1.27) \\ \hline \end{gathered}$ |
| $\geq 175 \mathrm{nmol} / \mathrm{L}$ |  | $\begin{aligned} & 39064 \\ & (197) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.57 \\ (1.31-1.88) \\ \hline \end{gathered}$ | $\begin{aligned} & 10245 \\ & (281) \\ & \hline \end{aligned}$ | $\begin{gathered} 1.87 \\ (1.62-2.16) \\ \hline \end{gathered}$ |
| Above the sex and ethnicity specific $80^{\text {th }}$ centile | PVD | $\begin{gathered} 64447 \\ (310) \end{gathered}$ | $\begin{gathered} 1.28 \\ (1.13-1.45) \end{gathered}$ | $\begin{aligned} & 18507 \\ & (451) \end{aligned}$ | $\begin{gathered} 1.53 \\ (1.37-1.72) \end{gathered}$ |

CI: confidence interval; CHD: coronary heart disease; CVD: cardiovascular disease; HR: hazard ratio; $\mathrm{Lp}(\mathrm{a})$ : Lipoprotein (a); nmol/L: nanomoles per litre; PVD: peripheral vascular disease.

