Very Low Lipoprotein(a) Levels Are Associated With Higher Risks Of New-Onset Diabetes And Non-Alcoholic Liver Disease

Ming Wai Yeung[1], M. Abdullah Said[1], Yordi J. van de Vegte[1], Niek Verweij[1], Robin P.F. Dullaart[3*], Pim van der Harst[1,2*]

Author affiliations:
[1]University of Groningen, University Medical Center Groningen, Department of Cardiology, 9700 RB Groningen, the Netherlands
[2]Department of Cardiology, Division of Heart & Lungs, University Medical Center Utrecht, University of Utrecht, Utrecht, the Netherlands
[3]Department of Endocrinology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

* Contributed equally

Keywords: cardiovascular risk factor, Lipoprotein(a), Type 2 diabetes, NAFLD, Mendelian randomization, Prospective study, UK Biobank

Address for correspondence:
Ming Wai Yeung and Pim van der Harst
Department of Cardiology
University of Groningen, University Medical Center Groningen,
Hanzeplein 1, 9700 RB Groningen, The Netherlands
Telephone number: +31 (0)50 3612355
Email: m.w.yeung@umcg.nl and p.vanderharst@umcutrecht.nl

Abstract

Aim

We aimed to study the association of very low serum Lipoprotein(a) [Lp(a)] concentrations with new-onset type 2 diabetes (T2D) and non-alcoholic liver disease (NAFLD) in the context of lipid-lowering medication usage in the UK Biobank, a large prospective population cohort. We tested the association for potential causality between genetically determined Lp(a) concentrations and disease outcomes using a Mendelian randomization (MR) strategy.

Methods

Using an extended biomarker dataset, we identified 47,840 participants with very low Lp(a) concentrations (<3.8 nmol/L) from a total of 449,026 participants. With a median follow-up of 12.4 years (interquartile range 11.5-13.1), we assessed the risk of new-onset cardiometabolic diseases in these participants stratified by usage of lipid-lowering medication using Cox proportional hazards models. We performed two sample MR analyses to test causal relationship between genetically predicted Lp(a) and T2D and NAFLD.

Results

Compared to participants with Lp(a) within the standard reportable range of 3.8-189 nmol/L, participants with very low Lp(a) (<3.8 nmol/L) were at increased risk of developing both T2D (HR 1.07; 95% CI 1.01-1.12) and NAFLD (HR 1.36; 95% CI 1.26-1.46) as opposed to decreased risk for coronary artery disease. The risk for new-onset T2D was higher in participants using lipid-lowering medication (fully adjusted HR 1.16; 95% CI 1.06-1.26). On the other hand, there was higher risk of NAFLD in participants not on lipid-lowering medication (HR 1.39; 95%CI 1.27-1.51).

There was no evidence for causal links between genetically predicted Lp(a) and T2D or NAFLD in two sample MR analyses.

Conclusions

Very low Lp(a) was associated with higher risks of T2D and NAFLD in a prospective analysis of the UK Biobank. These associations were influenced by lipid lowering medication usage. MR analyses did not support causality for these inverse associations.

Introduction

Lipoprotein(a) [Lp(a)] is a specific lipoprotein produced in the liver that consists of a polymorphic glycoprotein, called apolipoprotein(a) [apo(a)] linked to apolipoprotein-B(100) on an LDL particle in a 1:1 ratio. Ample evidence supports a causal association between elevated Lp(a) and various cardiovascular diseases (CVD) [1–3].Interestingly, recent epidemiological studies suggested that very low Lp(a) is associated with increased risk of new-onset type 2 diabetes mellitus (T2D). [1]. However, it is unclear whether Lp(a) is a mere marker or whether there is a causal mechanism behind this inverse relationship. An inverse Lp(a) association has also been reported with fatty liver disease [4]. Closely linked to metabolic dysfunction, the causal relationship between non-alcoholic liver disease (NAFLD) and diabetes has been supported by evidence from both epidemiological studies and animal studies [5–7].The opposed direction of the Lp(a) association with new-onset T2D, NAFLD and CVDs also raises the question whether the risk of diabetes might be modified by lipid-lowering medication treatment [8].

In the current study, we aimed to investigate Lp(a) associations with new-onset T2D and NAFLD in the UK Biobank, a large population-based cohort, and examined effect modification by treatment with lipid-lowering medications. We further examined the potential causal relationship of Lp(a) with T2D and NAFLD using a two-sample MR approach.

Methods

Study Population

The UK Biobank study is a population-based prospective cohort in the United Kingdom in which over 500,000 participants aged between 40 and 69 years were included between 2006-2010. All participants have given informed consent for this study. The UK Biobank has ethical approval from North West - Haydock Research Ethics Committee (REC reference: 16/NW/0274). Details of the UK Biobank study has been described in detail previously [9]. This research has been conducted using the UK Biobank Resource under Application Number 74395. Demographics including age, sex, smoking status and ethnic background were collected at the baseline visit.

Ascertainment of health outcomes

Ascertainment of diseases and medication usage in the UK Biobank was based on a combination of hospital inpatient records, primary care records and self-reported records during an interview with a trained staff (**Table S1**). Briefly, T2D diagnosis was captured using ICD10 code E10-E14 or equivalent diagnosis codes; NAFLD diagnosis was captured using K721, K740, K741, K742, K746, K758 and K760 or equivalent diagnosis codes[10]. Participants with prevalent disease at baseline per outcome were excluded from regression analysis. For T2D, we additionally excluded participants with baseline serum glycated haemoglobin (HbA1c) ≥6.5% (48 mmol/mol). Health outcome data was processed and extracted using the ukbpheno v1.0 package in R [11]. Participant follow-up started at inclusion and ended at date of event, death or last recorded follow-up (**Table S2**), whichever occurred first.

Serum biomarker measurement and extreme Lp(a) concentration

Lp(a) (in nmol/L) was measured using an immuno-turbidimetric assay (Randox Bioscience, UK). The reportable range of the assay is 3.8-189 nmol/L. To identify participants with extreme Lp(a) concentration, we obtained Lp(a) concentrations outside of the reportable range from the UK Biobank (Return 2321). We therefore considered participants to have very low Lp(a) if they had Lp(a) concentrations <3.8 nmol/L; and those with Lp(a) concentrations >189 nmol/L very high Lp(a).

Non-fasting venous blood samples were drawn during participants' baseline visit to the Assessment Centre [12]. Low density lipoprotein cholesterol (LDL-C) was measured using an enzymatic selective protection method while high density lipoprotein cholesterol (HDL-C) was measured using an enzyme immuno-inhibition method. Triglycerides were measured using an enzymatic method namely the glycerol-3-phosphate (GPO)-peroxidase (POD) chromogenic method. All biomarkers were measured on a Beckman Coulter AU5800 (Beckman Coulter [UK], Ltd). Glycated haemoglobin (HbA1c) was measured in mmol/mol with high performance liquid chromatography method (Bio-Rad Laboratories, Inc.) on Bio-Rad Variant II Turbo analysers[13].

Statistical analyses

Data are presented as median (IQR) for continuous variables and as counts with percentages for discrete and categorical variables, respectively.

Cox regression analyses

We performed two Cox proportional hazards regression analyses to test the association of Lp(a) with coronary artery disease (CAD), as compared to T2D and

NAFLD separately, stratified by lipid-lowering medication use. We first examined the disease risk by Lp(a) categories. We estimated HRs with 95% confidence intervals in participants with either very low or very high Lp(a) concentrations against participants with Lp(a) concentrations between 3.8 and 189 nmol/L. In the second analysis, we directly modelled the inverse rank normalized serum Lp(a) concentrations linearly only in participants with Lp(a) in the reported range. In both analyses, we considered four models. Model 1 was adjusted for age at baseline and sex; model 2 was further adjusted for BMI and hypertension; model 3 was further adjusted for LDL-C and HDL-C; and finally model 4 was further adjusted for triglycerides. Cox regressions were performed using Stata 17.1 (StataCorp LLC). We considered results with $p<0.001$ to be statistically significant to correct for multiple testing.

MR analyses

We performed two-sample Mendelian randomization (MR) analyses to assess potential causal effects of Lp(a) on T2D and NAFLD. We took the lead SNPs identified in the UK Biobank GWAS as instrumental variables for Lp(a) [3]. We took estimates from the summary statistics data of the DIAMANTE Consortium (1,159,055 controls and 180,834 [13.5%] cases) for diabetes[14] and EPoS Consortium for NAFLD (17,781 controls and 1,483 [7.7%] histologically confirmed cases)[15]. SNPs (genetic instruments) that were not available in the outcome GWAS were replaced with proxies in linkage disequilibrium (LD) of r2>0.8 respectively or excluded from the MR if no eligible proxies were available. Harmonization of SNP effects and the MR analyses were performed using R (version 4.0.3), the TwoSampleMR package (version 0.5.6) [16], MendelianRandomization

(version 0.6.0)[17], MRMix(version 0.1.0)[18], MR-PRESSO(Pleiotropy Residual Sum and Outlier) (version 1.0)[19] and mr.raps(version 0.2)[20].

MR estimates were obtained using an inverse variance weighted (IVW)-fixed effects model followed by sensitivity analyses to assess the robustness of the findings of the IVW estimates. For this, we used IVW-random effects, MR-Egger[21], median- , mode-based estimator MR analyses, MR-PRESSO, MR-Lasso[22], MR-Mix and MR-RAPS.

Weak instrument bias was tested using F-statistics[23], with F-statistics >10 indicating low risk of weak instrument bias. SNPs with F-statistics ≤10 were excluded from the analyses. The I2 index (> 25%)[24] and Cochran's Q statistic (P < 0.05)[25] were determined to assess potential heterogeneity within IVW-fixed effects estimates, indicating at least balanced horizontal pleiotropy for some of the SNPs. The MR-Egger method allows for a non-zero intercept hence capturing the unbalanced horizontal pleiotropic effects, if these pleiotropic effects are independent of their association with the exposure of interest (Instrument Strength Independent of Direct Effect [InSIDE] assumption)[26]. Rücker's Q'3 was calculated to assess heterogeneity within MR-Egger analyses, and the difference between Rücker's Q' statistic and Cochran's Q statistic (Q-Q') was tested[25]. The MR-Egger intercept was tested for a deviation from zero. We consider presence of unbalanced horizontal pleiotropy in the in the case of a significant Q-Q' and a non-zero MR-Egger intercept. The MR-Egger estimate was considered the causal estimate rather than the MR-IVW method in presence of unbalanced horizontal pleiotropy if the general InSIDE assumption holds.

We additionally assessed the potential weak instrument bias in the MR-Egger analyses by calculating the I2GX, which is the true variance of the SNP-exposure association and indicates low risk of measurement error if >95%[26].

The median/mode-based methods take, respectively, the median or mode of the ratio estimates instead of the weighted mean in IVW method. They are therefore robust against the presence of few invalid instruments. MR-PRESSO method performs the IVW method after exclusion of SNPs with MR estimates that differ substantially from the rest, thereby adjusting for the potential pleiotropy[19]. Similarly, the MR-Lasso method is another augmented version of the IVW method with outliers removed. MR-RAPS and MR-Mix methods adjust the estimates to address the violation of valid instrument assumption.

Results

We studied 449,026 participants (54.5% women with median age 58.3 years) with Lp(a) measurements and information on lipid-lowering medication usage. Among these participants, 377,564 had Lp(a) measurements within the reportable range of 3.8 -189 nmol/L (median 21.1; IQR 9.6-61.9); 47,840 participants had very low Lp(a) concentrations (<3.8 nmol/L) while 33,102 participants had very high Lp(a) (>189 nmol/L). A total of 78,044 (17.4%) participants were on lipid-lowering medication at baseline. Participants with very low Lp(a) concentrations were more likely to be men and to be White, had lower LDL-C and HDL-C but higher triglycerides (**Table 1**). During a median follow-up of 12.4 years (IQR 11.5-13.1) 40,000 participants were diagnosed with T2D and 7,078 with NAFLD. There were 45,484 cases of new-onset CAD. There was more new-onset diabetes (11.0% vs. 8.6% vs. 9.4%) as well as NAFLD (2.3% vs. 1.5% vs. 1.3%) among participants with very low Lp(a) concentrations (**Table 1**).

In all participants combined, Cox regression analyses showed that very low Lp(a) concentrations at baseline were associated with increased risk of T2D marginally (fully adjusted HR:1.07; 95% CI 1.01-1.12; p=0.011) and NAFLD (fully adjusted HR:1.36; 95% CI 1.26-1.46; p<0.001), in contrast to the decreased risk of CAD (fully adjusted HR 0.86; 95% CI 0.83-0.90; p<0.001) (**Figure 1**) . This inverse association with new-onset T2D was more prominent in participants using lipid-lowering medication at baseline (age and sex adjusted interaction p=0.04). Additional adjustments attenuated the association and it lost significance after further adjustment for triglycerides. Notably in the final model, the association remained significant among participants on lipid-lowering medication (HR1.16; 95% CI 1.17-1.38; p=0.001) but was lost in those not on lipid-lowering medication (HR 1.05; 95% CI 0.99-1.12; p=0.110) (**Table 2**). In all

participants combined, the association with NAFLD remained statistically significant in the fully adjusted model (**Figure 1**). The magnitude of HR was slightly higher in participants not using lipid-lowering medication than in participants on lipid-lowering medication (**Table 3**).

Reverse trends were observed in participants with very high Lp(a) concentrations (**Figure S1**). We furthermore modelled inverse rank normalized Lp(a) concentrations among participants whose Lp(a) concentrations were within the reportable range (N=369,425). As expected, the magnitude of risk was smaller, and neither association with diabetes nor NAFLD was significant in the fully adjusted Model 4 (**Figure S2**).

Given the ethnic difference in Lp(a), we assessed the associations with new-onset T2D (**Table S3**) and NAFLD (**Table S4**) by ethnic groups in participants not on lipid-lowering medication. We found no evidence of interaction effects between either disease and White ethnicity (age and sex adjusted interaction p=0.149 for T2D and p=0.456 for NAFLD, respectively).

Finally, a total of 37 and 35 genetic instruments for Lp(a) were available in the DIAMANTE and EPoS GWAS respectively for the MR analyses. We did not find evidence for an association of genetically predicted Lp(a) with T2D (IVW OR:1.02; 95% CI: 0.99-1.04, p=0.149) (**Figure S3**) nor NAFLD (IVW OR:1.00; 95% CI: 0.96-1.03, p=0.825) (**Figure S4**).

Discussion

Previous work equivocally suggested that higher Lp(a) concentration may be associated with a lower risk of diabetes [1, 27]. By using an extended biomarker dataset of the UK Biobank, we found individuals with very low Lp(a) concentrations (<3.8 nmol/L) were at increased risks of developing T2D as well as NAFLD, in contrast to its protective effect on CVDs. Notably, we found that the risk of diabetes was enhanced in individuals on lipid-lowering medication while the risk of NAFLD was attenuated in the context of lipid-lowering medication.

Our results from the two sample MR analyses using the 37 lead SNPs, which included rs10455872 (a common intronic SNP strongly associated with Lp(a) concentration in White individuals), as genetic instruments, did not support a causal relationship between Lp(a) and T2D nor NAFLD. Our results agree with previous work which solely used rs10455872 [28]. Sensitivity analyses suggested presence of horizontal pleiotropy in the association between Lp(a) with both traits. However, we were unable to investigate the role of Lp(a) isoform size as these data are not available in the UK Biobank, which is among the limitations of the current study.

Several studies have implicated a role of triglyceride-rich lipoproteins in the metabolism of Lp(a) [8, 29, 30], whereas conversely genetically predicted Lp(a) may influence very low density lipoprotein (VLDL) metabolism [31]. These findings corroborate to study the association of Lp(a) with new-onset T2D and NAFLD in the context of lipid-lowering medications use, and to adjust for triglycerides when examining the association of Lp(a) with T2D and NAFLD. Experimental studies are needed to study the intricate interactions between Lp(a) and VLDL as well as its involvement in the associations with diabetes and non-alcoholic liver disease reported in observational studies.

Conclusions

In prospective analyses of the UK Biobank, very low Lp(a) concentrations were associated with higher risks of T2D and non-alcoholic liver disease. This inverse association was modified by lipid-lowering medications usage. We found no evidence for an association between genetically predicted Lp(a) and T2D or NAFLD.

Reference

1. Kronenberg F, Mora S, Stroes ESG, et al (2022) Lipoprotein(a) in atherosclerotic cardiovascular disease and aortic stenosis: a European Atherosclerosis Society consensus statement. Eur Heart J 0:1–22. https://doi.org/10.1093/EURHEARTJ/EHAC361

2. Dullaart RPF (2020) Lipoprotein(a): The Renaissance of an Enigmatic Lipoprotein. J Clin Endocrinol Metab 105(3):e896–e898. https://doi.org/10.1210/clinem/dgz111

3. Said MA, Yeung MW, Van De Vegte YJ, et al (2021) Genome-Wide Association Study and Identification of a Protective Missense Variant on Lipoprotein(a) Concentration: Protective Missense Variant on Lipoprotein(a) Concentration - Brief Report. Arterioscler Thromb Vasc Biol 41:1792–1800. https://doi.org/10.1161/ATVBAHA.120.315300

4. Jung I, Kwon H, Park SE, et al (2020) Serum lipoprotein(a) levels and insulin resistance have opposite effects on fatty liver disease. Atherosclerosis 308:1–5. https://doi.org/10.1016/j.atherosclerosis.2020.07.020

5. Kwok R, Choi KC, Wong GL-H, et al (2016) Screening diabetic patients for non-alcoholic fatty liver disease with controlled attenuation parameter and liver stiffness measurements: a prospective cohort study. Gut 65(8):1359–1368. https://doi.org/10.1136/gutjnl-2015-309265

6. De Silva NMG, Borges MC, Hingorani AD, et al (2019) Liver Function and Risk of Type 2 Diabetes: Bidirectional Mendelian Randomization Study. Diabetes 68(8):1681–1691. https://doi.org/10.2337/db18-1048

7. Liu Z, Zhang Y, Graham S, et al (2020) Causal relationships between NAFLD, T2D and obesity have implications for disease subphenotyping. J Hepatol 73(2):263–276. https://doi.org/10.1016/j.jhep.2020.03.006

8. Nestel P, Loh WJ, Ward NC, Watts GF (2022) New Horizons: Revival of Lipoprotein (a) as a Risk Factor for Cardiovascular Disease. J Clin Endocrinol Metab 107(11):e4281–e4294. https://doi.org/10.1210/clinem/dgac541

9. Sudlow C, Gallacher J, Allen N, et al (2015) UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLOS Med 12(3):e1001779. https://doi.org/10.1371/journal.pmed.1001779

10. Verweij N, Haas ME, Nielsen JB, et al (2022) Germline Mutations in CIDEB and Protection against Liver Disease. N Engl J Med 387(4):332–344. https://doi.org/10.1056/NEJMoa2117872

11. Yeung MW, van der Harst P, Verweij N (2022) ukbpheno v1.0: An R package for phenotyping health-related outcomes in the UK Biobank. STAR Protoc 3(3):101471. https://doi.org/10.1016/J.XPRO.2022.101471

12. UK Biobank (2011) Blood Sample Collection, Processing and Transport. https://biobank.ndph.ox.ac.uk/ukb/ukb/docs/Bloodsample.pdf

13. Fry D, Almond R, Moffat S, Gordon M, Singh P (2019) UK Biobank Biomarker Project Companion Document to Accompany Serum Biomarker Data

14. Mahajan A, Spracklen CN, Zhang W, et al (2022) Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. Nat Genet 54(5):560–572. https://doi.org/10.1038/s41588-022-01058-3

15. Anstee QM, Darlay R, Cockell S, et al (2020) Genome-wide association study of non-alcoholic fatty liver and steatohepatitis in a histologically characterised cohort☆. J Hepatol 73(3):505–515. https://doi.org/10.1016/j.jhep.2020.04.003

16. Hemani G, Zheng J, Elsworth B, et al (2018) The MR-Base platform supports systematic causal inference across the human phenome. eLife 7. https://doi.org/10.7554/eLife.34408

17. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM (2020) A robust and efficient method for Mendelian randomization with hundreds of genetic variants. Nat Commun 11(1):376. https://doi.org/10.1038/s41467-019-14156-4

18. Qi G, Chatterjee N (2019) Mendelian randomization analysis using mixture models for robust and efficient estimation of causal effects. Nat Commun 10(1):1941. https://doi.org/10.1038/s41467-019-09432-2

19. Verbanck M, Chen C-Y, Neale B, Do R (2018) Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 50(5):693–698. https://doi.org/10.1038/s41588-018-0099-7

20. Zhao Q, Wang J, Hemani G, Bowden J, Small DS (2019) Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score

21. Bowden J, Davey Smith G, Burgess S (2015) Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 44(2):512–525. https://doi.org/10.1093/ije/dyv080

22. Rees JMB, Wood AM, Dudbridge F, Burgess S (2019) Robust methods in Mendelian randomization via penalization of heterogeneous causal estimates. PLOS ONE 14(9):e0222362. https://doi.org/10.1371/journal.pone.0222362

23. Palmer TM, Lawlor DA, Harbord RM, et al (2012) Using multiple genetic variants as instrumental variables for modifiable risk factors. In: Statistical Methods in Medical Research. pp 223–242

24. Greco M FD, Minelli C, Sheehan NA, Thompson JR (2015) Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. Stat Med 34(21):2926–2940. https://doi.org/10.1002/sim.6522

25. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan N, Thompson J (2017) A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med 36(11):1783–1802. https://doi.org/10.1002/sim.7221

26. Bowden J, Del Greco M. F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR (2016) Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. Int J Epidemiol 45(6):dyw220. https://doi.org/10.1093/ije/dyw220

27. Månsson M, Kalies I, Bergström G, et al (2014) Lp(a) is not associated with diabetes but affects fibrinolysis and clot structure ex vivo. Sci Rep 4(1):5318. https://doi.org/10.1038/srep05318

28. Ye Z, Haycock PC, Gurdasani D, et al (2013) The Association Between Circulating Lipoprotein(a) and Type 2 Diabetes: Is It Causal? Diabetes 63(1):332–342. https://doi.org/10.2337/db13-1144

29. Moriarty PM, Varvel SA, Gordts PLSM, McConnell JP, Tsimikas S (2017) Lp(a) mass levels increase significantly according to APOE genotype: An analysis of 431,239 patients. Arterioscler Thromb Vasc Biol 37(3):580. https://doi.org/10.1161/ATVBAHA.116.308704

30. Croyal M, Blanchard V, Ouguerram K, et al (2020) VLDL (very-low-density lipoprotein)-Apo e (Apolipoprotein E) may influence Lp(a) (Lipoprotein [a]) synthesis or assembly. Arterioscler Thromb Vasc Biol 819–829. https://doi.org/10.1161/ATVBAHA.119.313877

31. Kettunen J, Demirkan A, Würtz P, et al (2016) Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun 2016 71 7(1):1–9. https://doi.org/10.1038/ncomms11122

Declaration

Ethics approval and consent to participate

The UK Biobank has ethical approval from North West - Haydock Research Ethics Committee (REC reference: 16/NW/0274). All participants have given informed consent.

Consent for publication

All authors consent this manuscript for publication.

Availability of data and materials

The data that support the findings of this study are available from the corresponding authors upon reasonable request. Source data may be requested from the UK biobank.

Competing interests

M.W.Y. is a full-time employee of GSK plc; this work was conducted before employment by GSK. N.V. is a full-time employee of Regeneron Pharmaceutical Inc. and receives stock options and restricted stock units as compensation. Other authors declare no competing interests.

Funding

None

Authors' contributions

R.P.F.D., M.W.Y. and M.A.S. contributed to the conception and design of the study. M.W.Y. contributed to acquistion of data and performed the analysis of data. M.W.Y. wrote this paper under the supervision of R.P.F.D. and P.v.d.H.. All authors contributed to interpretation of data, reviewed the draft paper and provided critical intellectual content and approved the final version of the manuscript.

Acknowledgments

This research has been conducted using the UK Biobank resource under application numbers 74395. We thank all UK Biobank participants for their contributions.

Table 1. Baseline characteristics of UKB participants by Lp(a) concentration

Variable	Lp(a) <3.8 nmol/L	3.8≤ Lp(a) ≤189 nmol/L	Lp(a) >189 nmol/L	*p*
N	47,131	369,425	32,470	
Age, years (IQR)	57.54 (49.66, 63.43)	58.22 (50.45, 63.70)	59.70 (52.75, 64.31)	<0.001
Sex				<0.001
Men	24,555 (52.1%)	167,457 (45.3%)	12,504 (38.5%)	
Women	22,576 (47.9%)	201,968 (54.7%)	19,966 (61.5%)	
Ethnicity				<0.001
White	45,882 (97.3%)	346,593 (93.8%)	30,597 (94.2%)	
Asian	543 (1.2%)	9192 (2.5%)	445 (1.4%)	
Black	86 (0.2%)	6083 (1.6%)	878 (2.7%)	
Mixed	166 (0.4%)	2307 (0.6%)	197 (0.6%)	
Other/Unknown	454 (1.0%)	5250 (1.4%)	353 (1.1%)	
Current smoker	5038 (10.7%)	39,198 (10.6%)	3379 (10.4%)	0.42
BMI, kg/m^2	27.08 (24.30, 30.46)	26.69 (24.10, 29.81)	26.89 (24.38, 30.00)	<0.001
Systolic blood pressure, mmHg	133.00 (121.50, 145.47)	131.71 (120.21, 144.14)	133.05 (121.13, 145.47)	<0.001
Diastolic blood pressure, mmHg	82.12 (76.47, 88.19)	81.72 (76.06, 87.39)	81.73 (76.06, 87.79)	<0.001
LDL-C, mg/dl	128.72 (107.41, 151.26)	135.64 (113.79, 158.53)	145.11 (120.25, 170.51)	<0.001
HDL-C, mg/dl	52.74 (43.85, 63.68)	54.05 (45.28, 64.69)	56.30 (47.37, 66.93)	<0.001
Triglycerides, mg/dl	136.49 (94.24, 201.06)	130.65 (92.21, 189.11)	127.02 (91.50, 180.69)	<0.001
Hypertension	15,023 (31.9%)	112,593 (30.5%)	12,094 (37.3%)	<0.001
On lipid-lowering medication	8029 (17.0%)	60,714 (16.4%)	9301 (28.6%)	<0.001
Statin	7746 (16.4%)	58,798 (15.9%)	9122 (28.1%)	<0.001
Non-statin	626 (1.3%)	4354 (1.2%)	805 (2.5%)	<0.001
New-onset T2D	5140 (11.0%)	31,804 (8.6%)	3056 (9.4%)	<0.001
New-onset non alcoholic liver disease	1096 (2.3%)	5553 (1.5%)	429 (1.3%)	<0.001
New-onset coronary artery disease	4403 (9.3%)	36,110 (9.8%)	4971 (15.3%)	<0.001
New-onset aortic stenosis	423 (0.9%)	3646 (1.0%)	639 (2.0%)	<0.001
New-onset heart failure	1593 (3.4%)	11,926 (3.2%)	1403 (4.3%)	<0.001
New-onset stroke	1972 (4.2%)	15,460 (4.2%)	1703 (5.2%)	<0.001
New-onset peripheral artery disease	592 (1.3%)	4883 (1.3%)	791 (2.4%)	<0.001

BMI: body mass index, HDL-C: high density lipoprotein cholesterol, IQR: interquartile range, LDL-C: low-density lipoprotein cholesterol.

Table 2. Multivariable logistic regression of Lp(a) concentration on Type 2 diabetes

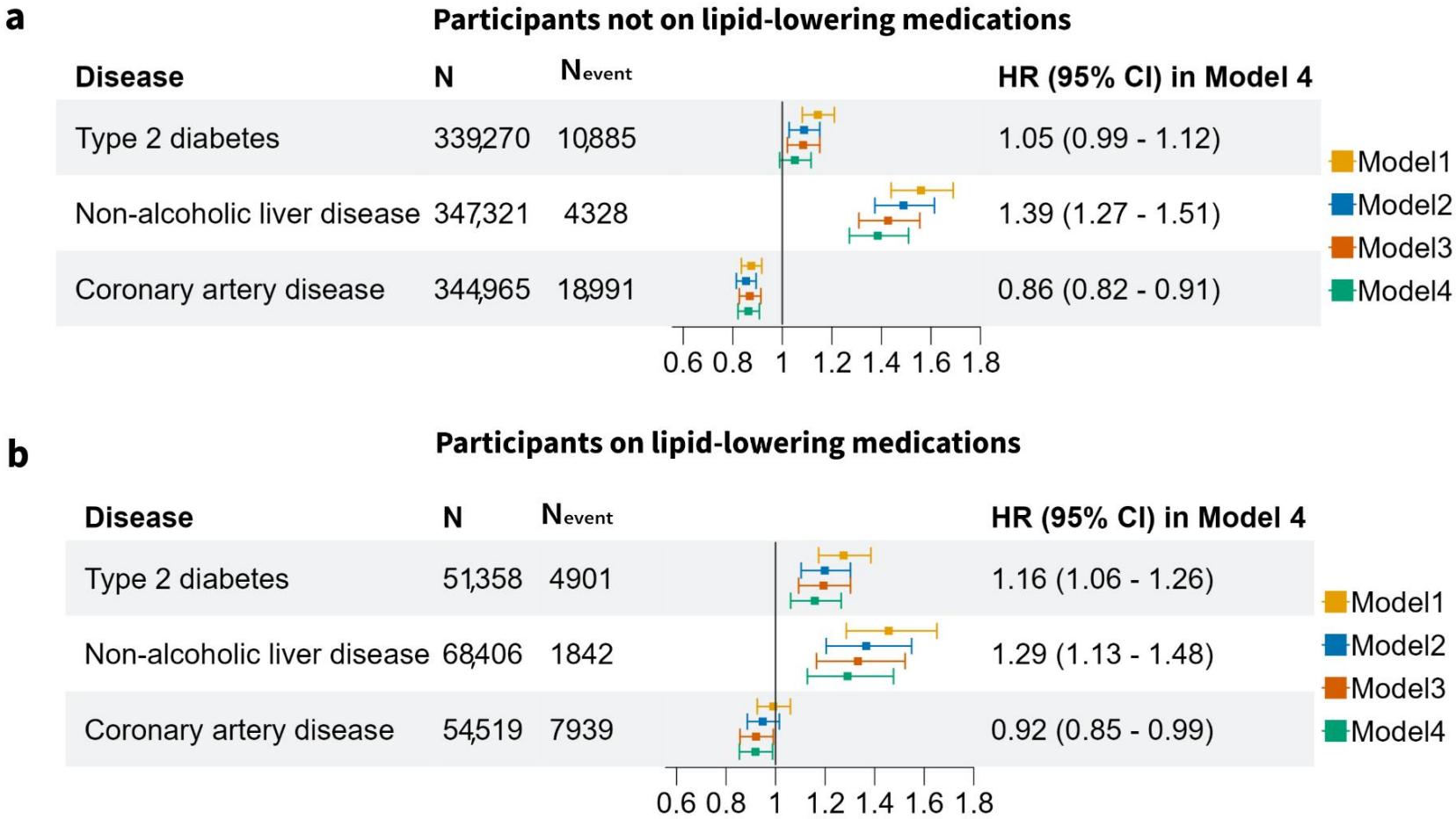
Lp(a) concentration	Model	Covariates	HR	95% CI	*p*	Lp(a) concentration	Model	HR	95% CI	*p*
participants not on lipid-lowering medication										
very low Lp(a)	Model 1	Age, Sex	1.14	1.08 - 1.21	<0.001	very high Lp(a)	Model 1	0.91	0.84 - 0.99	0.028
	Model 2	Model 1 + BMI and hypertension	1.09	1.03 - 1.15	0.004		Model 2	0.91	0.84 - 0.99	0.022
	Model 3	Model 2 + LDL-C + HDL-C	1.08	1.02 - 1.15	0.009		Model 3	0.95	0.87 - 1.03	0.206
	Model 4	Model 3 + triglycerides	1.05	0.99 - 1.12	0.110		Model 4	0.98	0.90 - 1.07	0.627
participants on lipid-lowering medication										
very low Lp(a)	Model 1	Age, Sex	1.27	1.17 - 1.38	<0.001	very high Lp(a)	Model 1	0.93	0.85 - 1.01	0.071
	Model 2	Model 1 + BMI and hypertension	1.20	1.10 - 1.30	<0.001		Model 2	0.95	0.87 - 1.03	0.231
	Model 3	Model 2 + LDL-C + HDL-C	1.19	1.09 - 1.30	<0.001		Model 3	0.98	0.89 - 1.06	0.575
	Model 4	Model 3 + triglycerides	1.16	1.06 - 1.26	0.001		Model 4	1.01	0.92 - 1.10	0.907

Very low Lp(a): Lp(a) <3.8 nmol/L; very high Lp(a): Lp(a) >189 nmol/L; Reference: Lp(a)3.8-189 nmol/L. BMI: body mass index, CI: confidence interval, HDL-C: high density lipoprotein cholesterol, HR: hazard ratio, LDL-C: low-density lipoprotein cholesterol.

Table 3. Multivariate logistic regression of Lp(a) concentration on non-alcoholic liver disease

Lp(a) concentration	Model	Covariates	HR	95% CI	*p*	Lp(a) concentration	Model	HR	95% CI	*p*
Full cohort (participants not on lipid lowering medication)										
very low Lp(a)	Model 1	Age, Sex	1.56	1.44 - 1.69	<0.001	very high Lp(a)	Model 1	0.77	0.67 - 0.89	<0.001
	Model 2	Model 1 + BMI and hypertension	1.49	1.37 - 1.61	<0.001		Model 2	0.76	0.66 - 0.88	<0.001
	Model 3	Model 2 + LDL-C + HDL-C	1.43	1.31 - 1.55	<0.001		Model 3	0.84	0.72 - 0.97	0.015
	Model 4	Model 3 + triglycerides	1.39	1.27 - 1.51	<0.001		Model 4	0.86	0.75 - 1.00	0.046
Full cohort (participants on lipid lowering medication)										
very low Lp(a)	Model 1	Age, Sex	1.46	1.29 - 1.65	<0.001	very high Lp(a)	Model 1	0.82	0.71 - 0.95	0.009
	Model 2	Model 1 + BMI and hypertension	1.37	1.20 - 1.55	<0.001		Model 2	0.86	0.74 - 1.00	0.046
	Model 3	Model 2 + LDL-C + HDL-C	1.33	1.17 - 1.52	<0.001		Model 3	0.88	0.76 - 1.03	0.124
	Model 4	Model 3 + triglycerides	1.29	1.13 - 1.48	<0.001		Model 4	0.91	0.78 - 1.06	0.228

Very low Lp(a): Lp(a) <3.8 nmol/L; very high Lp(a): Lp(a) >189 nmol/L; Reference: Lp(a)3.8-189 nmol/L. BMI: body mass index, CI: confidence interval, HDL-C: high density lipoprotein cholesterol, HR: hazard ratio, LDL-C: low-density lipoprotein cholesterol.

Figure 1. Forest plot illustrating the unadjusted and adjusted hazard ratios of very low Lp(a) concentrations at baseline with the development of T2D, NAFLD and CAD.

Hazard ratios for diseases in **(a)** participants not using lipid-lowering medications and **(b)** participants using lipid-lowering medications. Model 1: adjusted for age and sex; Model 2: Model 1 + body mass index and hypertension; Model 3: Model 2 + low-density lipoprotein cholesterol and high-density lipoprotein cholesterol; Model 4: Model 3 + triglycerides. Hazard ratios (HR) with 95% confidence intervals (CI) of Model 4 are shown for new-onset T2D, NAFLD and CAD. N: Number of participants, $N_{event:}$ number of events.