

1 **The relationship between Lipoprotein A and other lipids with**  
2 **prostate cancer risk: A multivariable Mendelian**  
3 **randomisation study.**

4  
5 Anna Ioannidou<sup>1</sup>, Eleanor L Watts<sup>2</sup>, Aurora Perez-Cornago<sup>2</sup>, Elizabeth A Platz<sup>3-5</sup>, Ian G Mills<sup>6,7,8</sup>,  
6 Timothy J Key<sup>2</sup>, Ruth C Travis<sup>2</sup>, The PRACTICAL consortium, CRUK, BPC3, CAPS,  
7 PEGASUS<sup>†</sup>, Konstantinos K Tsilidis<sup>1,9\*</sup>, Verena Zuber<sup>1\*</sup>

8  
9 <sup>1</sup> Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London,  
10 London, UK.

11 <sup>2</sup> Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford,  
12 Oxford, UK.

13 <sup>3</sup> Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore,  
14 Maryland, USA

15 <sup>4</sup> Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, Maryland, USA

16 <sup>5</sup> Department of Urology and the James Buchanan Brady Urological Institute, Johns Hopkins  
17 University School of Medicine, Baltimore, Maryland, USA

18 <sup>6</sup> Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK

19 <sup>7</sup> Patrick G Johnston Centre for Cancer Research (PGJCCR), Queen's University Belfast, Belfast,  
20 UK

21 <sup>8</sup> Centre for Cancer Biomarkers (CCBIO), University of Bergen, 5021 Bergen, Norway

22 <sup>9</sup> Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina,  
23 Greece.

24

25 \* contributed equally

26 † Members from the PRACTICAL Consortium, CRUK, BPC3, CAPS and PEGASUS are provided  
27 in the Supplement (S1 file).

28

29 Funding: KKT was supported by Cancer Research UK (C18281/A29019). ELW was supported by  
30 the NDPH Early Career Research Fellowship. APC is supported by a Cancer Research UK  
31 Population Research Fellowship (C60192/A28516) and by the World Cancer Research Fund

32 (WCRF UK), as part of the World Cancer Research Fund International grant programme  
33 (2019/1953).  
34

## 35 **Abstract**

### 36 **Background**

37 Numerous epidemiological studies have investigated the role of blood lipids in prostate cancer  
38 (PCa) risk though findings remain inconclusive to date. The ongoing research has mainly involved  
39 observational studies which are often prone to confounding. This study aimed to identify the  
40 relationship between genetically predicted blood lipid concentrations and PCa.

### 41 **Methods and Findings**

42 Data for low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol,  
43 triglycerides (TG), apolipoprotein A (apoA) and B (apoB), lipoprotein A (Lp(a)) and PCa were  
44 acquired from genome-wide association studies in UK Biobank and the PRACTICAL consortium,  
45 respectively. We used a two-sample Mendelian randomisation (MR) approach with both univariable  
46 and multivariable (MVMR) models and utilised a variety of robust methods and sensitivity analyses  
47 to assess the possibility of MR assumptions violation. No association was observed between  
48 genetically predicted concentrations of HDL, TG, apoA and apoB and PCa risk. Genetically  
49 predicted LDL concentration was positively associated with total PCa in the univariable analysis  
50 but adjustment for HDL, TG and Lp(a) led to a null association. Genetically predicted concentration  
51 of Lp(a) was associated with higher total PCa risk in the univariable ( $OR_{\text{weighted median per sd}} = 1.091$ ;  
52  $95\% \text{ CI } 1.028\text{-}1.157$ ;  $P=0.004$ ) and MVMR analyses after adjustment for the other lipid traits  
53 ( $OR_{\text{IVW per sd}} = 1.068$ ;  $95\% \text{ CI } 1.005\text{-}1.134$ ;  $P = 0.034$ ). Genetically predicted Lp(a) was also  
54 associated with advanced (MVMR  $OR_{\text{IVW per sd}} = 1.078$ ;  $95\% \text{ CI } 0.999\text{-}1.163$ ;  $P=0.055$ ) and early  
55 age onset PCa (MVMR  $OR_{\text{IVW per sd}} = 1.150$ ;  $95\% \text{ CI } 1.015,1.303$ ;  $P = 0.028$ ). Although multiple  
56 estimation methods were utilized to minimize the effect of pleiotropic traits, the presence of any  
57 unmeasured pleiotropy cannot be excluded and may limit our findings.

### 58 **Conclusions**

59 We observed that genetically predicted Lp(a) concentrations are associated with an increased PCa  
60 risk. Future studies are required to understand the underlying biological pathways of this finding, as  
61 it may inform PCa prevention through Lp(a)-lowering strategies.

62

## 63 **Introduction**

64 Prostate cancer (PCa) is one of the most frequently diagnosed cancers in men [1] with 1,276,106  
65 incident cases reported globally during 2018 [2]. There is high geographical heterogeneity of PCa  
66 incidence, which is reflected in a 40-fold difference in the age-adjusted incidence rates across the  
67 globe [3]. Several studies have argued that this could be attributed to the increased number of  
68 diagnoses in countries where the prostate specific antigen (PSA) screening is prevalent.  
69 Nevertheless, the basis for this heterogeneity remains poorly understood [1].

70 Given that PCa is also clinically heterogeneous, risk factors identified to date differ by disease  
71 aggressiveness [4]. In particular, established risk factors for total PCa are mainly non-modifiable,  
72 including older age, African-descent and genetics [5], whereas some potential risk factors for  
73 aggressive PCa include smoking, obesity [6], lower vitamin D and higher blood lipid levels [4]  
74 which are modifiable. Lipid-lowering therapies are cheap and well-established for lowering  
75 cardiovascular risk. Yet, there is no conclusive evidence that repurposed lipid-lowering drugs are  
76 effective for the prevention of PCa. It is therefore important to determine whether blood lipids  
77 increase PCa risk, especially lethal disease [7]. A meta-analysis of 14 prospective studies published  
78 in 2015 [10] did not observe significant associations between TC, HDL or LDL concentrations and  
79 risk of total or high-grade PCa, but high between-study heterogeneity was evident for most  
80 associations. Two meta-analyses have examined the role of statin use in PCa risk, and both  
81 observed inverse associations of statins and advanced PCa risk [11][12]. Nonetheless, whether  
82 these associations can be attributed to lower cholesterol itself or some other mechanism is  
83 unknown.

84 Observational studies may suffer from unobserved confounding and reverse-causation [13], which  
85 could explain inconsistent findings among studies. Mendelian randomisation (MR) uses genetic  
86 variants as proxies for the exposures of interest and, if carefully conducted, can complement  
87 observational research [14] and support triangulation of evidence. That is because genetic variants  
88 are randomly allocated to offspring by parents and independently assorted during meiosis, which

89 minimise issues with reverse-causation and confounding [13,15]. In addition, most studies on lipids  
90 and PCa measure lipid levels only once which can lead to measurement error in the findings,  
91 whereas genetically predicted lipid levels capture lifelong expected levels. Previous MR research is  
92 limited to two studies that examined the role of HDL, LDL and TG in PCa risk overall and by  
93 disease stage and grade, and both reported null associations [16,17]. However, neither study  
94 adjusted for multiple lipid traits, which may have limited their findings given that different lipids  
95 are correlated and pleiotropic [18]. In this paper, we aim to identify whether genetically predicted  
96 lipid traits are associated with overall PCa risk, and in particular advanced and early age onset  
97 disease. We incorporated a univariable and multivariable MR (MVMR) framework to adjust for  
98 pleiotropic lipid effects, and examined the role of HDL, LDL and TG, as well as additional lipid-  
99 traits that have been underexamined to date, such as lipoprotein A (Lp(a)), apolipoprotein A (apoA)  
100 and apolipoprotein B (apoB).

## 101 **Methods**

### 102 **Study populations**

#### 103 **Blood Lipids Data**

104 Genome-wide association (GWA) data for HDL, LDL, TG, Lp(a), apoA and apoB were available  
105 from UK Biobank, with information on over 13.7 million single-nucleotide polymorphisms (SNPs)  
106 [19]. Model adjustments in this UK Biobank GWAS included age, age<sup>2</sup>, inferred\_sex, interaction  
107 terms for age\*inferred\_sex and for age<sup>2</sup>\*inferred\_sex, and the first 20 principal components. All  
108 measured serum biomarkers were approximately normally distributed except Lp(a), which was  
109 positively skewed. For consistency purposes, inverse rank-normalised data were used for all  
110 biomarkers. The heritability estimation from summary statistics (HESS) was applied in each of  
111 these lipids in a GWAS of Sinnott-Armstrong et al. [20] and heritability was reported as follows;  
112 HDL 36%, LDL 29%, TG 29%, Lp(a) 24%, apoA 31% and apoB 32% indicating strong genetic  
113 regulation of all lipid traits considered as exposures (Table S1). For the purpose of this research and

114 to match with the PCa GWAS, only European ancestry male participants were included  
115 (N=167,020)

## 116 **PCa Data**

117 Summary association statistics for PCa risk were acquired from the PRACTICAL consortium and  
118 are based on Schumacher et al. [21]. The genotyping was performed using a custom array, namely  
119 the OncoArray. A two-stage imputation was performed, using the SHAPEIT20 and IMPUTEv221.  
120 For our analysis, we used total, advanced (metastatic or Gleason score (GS)  $\geq 8$  or PSA $>100$   
121 ng/mL or PCa death) and early age onset (PCa age  $\leq 55$ ) PCa. Study participants for total PCa  
122 make up to a total of 79,166 cases and 61,106 controls, advanced PCa cases include 15,167  
123 participants and 58,308 controls, whereas early age onset PCa includes 6,988 cases and 44,256  
124 controls. All participants were of European ancestry.

## 125 **Assumptions**

126 The following assumptions were made for all our MR analyses and are described in combination for  
127 both the univariable and MVMR approaches [22].

128 **1. Relevance:** Genetic variants are associated with the exposure of interest in the case of  
129 univariable MR, whereas for MVMR, they are associated with at least one of the exposures.

130 **2. Exchangeability:** Genetic variants are independent of all confounders of the exposure-  
131 outcome association for the univariable MR, whereas in the MVMR, variants are  
132 independent of all confounders of each of the exposure-outcome associations.

133 **3. Exclusion-restriction:** Genetic variants are independent of the outcome given the  
134 exposure/es and all the confounders.

## 135 **Main MR analyses**

136 All MR analyses were performed in R version 4.0.0. Due to the availability of exposure (blood  
137 lipids) and outcome (PCa) data from two different sources, we used a two-sample MR study design.  
138 In the univariable MR, SNPs that satisfied genome-wide significance ( $P < 5 \times 10^{-8}$ ) were selected  
139 for each trait. After removing any inconsistencies with respect to the effect and non-effect alleles,

140 we harmonised the data so that the exposure and outcome datasets would have the same effect  
141 allele. We used the TwoSampleMR package version 0.5.4 to clump the data using a threshold of  
142  $r^2 < 0.001$  to identify and remove any SNPs in high linkage disequilibrium (LD). All SNPs left after  
143 clumping were considered as the instrumental variables (IVs). We firstly ran the univariable  
144 analysis on all blood lipids for both total and advanced PCa, whilst for the early age onset PCa we  
145 focused our univariable analysis primarily on Lp(a). For the main estimation methods of the  
146 univariable analyses, we performed the inverse variance weighting (IVW) [23,24] and weighed  
147 median [13] approaches, and we additionally applied the MR-Egger [25] approach.  
148 To adjust for different lipid traits in our models, we performed a MVMR analysis. We chose to  
149 exclude apoA and apoB to avoid multicollinearity issues due to their high correlation with HDL and  
150 LDL, respectively ( $r_{apoA,HDL} = 0.978$ ;  $P < 2.2 \times 10^{-6}$  /  $r_{apoB,LDL} = 0.984$ ;  $P < 2.2 \times 10^{-6}$ ). The minimum  
151  $p$ -value across the remaining lipids was computed and selection of SNPs was based on those that  
152 satisfied genome-wide significance through the minimum  $p$ -value ( $P < 5 \times 10^{-8}$ ). After  
153 harmonisation was performed, we clumped the data based on a threshold of  $r^2 < 0.001$ . The main  
154 estimation method performed was the IVW, while we additionally implemented the MR-Egger  
155 estimate [26] to control for any remaining unmeasured pleiotropy.

## 156 **Sensitivity MR analyses**

157 As we observed a positive finding for Lp(a) and PCa outcomes, we performed the following  
158 sensitivity analyses in our univariable MR considering only Lp(a) as exposure for total, advanced  
159 and early age onset of PCa.

160 **1. Sensitivity analysis 1:** As an attempt to increase the statistical power of the univariable MR,  
161 we used an eased clumping threshold of  $r^2 < 0.01$  and re-fitted the models based on a larger  
162 set of IVs.

163 **2. Sensitivity analysis 2:** Variants that were used as IVs for Lp(a) in the Burgess et al. paper  
164 [27] based on a clumping threshold of  $r^2 < 0.4$ , were separately fitted to the univariable  
165 models to validate findings on a different IV set. 43 IVs were initially proposed in the paper,

166 but only 28 of those were available in our current datasets and satisfied genome-wide  
167 significance. The univariable models were re-fitted based on these 28 IVs.

168 **3. Sensitivity analysis 3:** As the *LPA* gene (chromosome 6: 160,531,482-160,664,275) is the  
169 main gene associated with Lp(a) concentrations and explains about 70-90% of its variability  
170 [28], we selected variants located in the *LPA* gene based on a clumping threshold of  
171  $r^2 < 0.001$  to represent strong biological instruments and potentially support the effect of  
172 Lp(a). Four such variants were identified and subsequently utilised as IVs.

173 **4. Sensitivity analysis 4:** Additional robust estimation methods were utilised as part of our  
174 sensitivity analyses to control and/or test for horizontal pleiotropy. These included the MR-  
175 PRESSO [29] and contamination mixture [30].

176 As obesity may be considered a probable confounder for lipids and PCa [16], we also performed an  
177 additional adjustment for body mass index (BMI) in all the MVMR models, using genetic  
178 association data for BMI from UK Biobank [19]. Lp(a) is assembled in the liver [31], whereas liver  
179 function/disease has been proposed to influence PCa detection and outcomes [31-32]. Genetic  
180 associations for aspartate aminotransferase (AST) were thereby adjusted in a MVMR model  
181 including Lp(a) and total PCa. In addition, as kidney disease has been suggested to affect Lp(a)  
182 concentrations [34], and creatinine was previously associated with PCa risk [35], we performed  
183 another MVMR analysis using Lp(a), creatinine and total PCa to control for kidney function. Both  
184 genetic associations for AST and creatinine were acquired from UK Biobank [19]. Throughout our  
185 analyses, we considered significant estimates based on the 95% confidence level. We additionally  
186 estimated a Bonferroni-corrected p-value for the main univariable analyses on total and advanced  
187 PCa, to adjust for the multiple tests performed on each outcome. The total number of tests is  
188 reflected upon the number of different lipids we considered for each PCa outcome. Throughout the  
189 results section, nominally significant p-values are reported.



## 190 **Results**

191 Descriptive statistics for lipid measurements were available from UK Biobank [19] and can be seen  
192 in S2 Table. Throughout this section we report results based solely on the IVW and weighted  
193 median methods. Results from the additional methods we used, including MR-Egger (Table S3-S6),  
194 MR-PRESSO and contamination mixture estimates (Table S3), and MVMR analyses adjusting for  
195 BMI, AST and creatinine (Tables S6-9), can be found in the supplement and were in general in  
196 agreement with the main analyses presented in the text below. The marginal associations of the  
197 genetic instruments with exposures, outcomes and confounders/mediators are shown in Tables S10-  
198 15.

### 199 **Univariable MR**

#### 200 **Total PCa**

201 The univariable MR analysis showed that genetically predicted HDL ( $OR_{IVW} = 0.994$ ; 95% CI  
202  $= [0.942, 1.051]$ ;  $P = 0.825$ ), TG ( $OR_{IVW} = 1.026$ ; 95% CI  $= [0.961, 1.105]$ ;  $P = 0.449$ ), apoA  
203 ( $OR_{IVW} = 1.025$ ; 95% CI  $= [0.970, 1.083]$ ;  $P = 0.372$ ) and apoB ( $OR_{IVW} = 1.026$ ; 95% CI  $=$   
204  $0.961, 1.094]$ ;  $P = 0.411$ ) concentrations were not associated with total PCa risk (Table S4). In  
205 contrast, the odds ratio of total PCa was 1.088 per standard deviation (sd) increase in genetically  
206 predicted LDL (95% CI  $= [1.010, 1.162]$ ;  $P = 0.016$ ). This association was however not supported by  
207 the weighted median approach ( $OR = 1.016$ ; 95% CI  $= [0.942, 1.094]$ ;  $P = 0.669$ ). This raised concerns  
208 for potential pleiotropic effects present in our model and results were further assessed in the  
209 multivariable model.

210 Genetically predicted Lp(a) had an insignificant association on total PCa as estimated from the IVW  
211 ( $OR_{IVW} = 1.066$ ; 95% CI  $= [0.909, 1.249]$ ;  $P = 0.431$ ) method (Table 1), but the odds ratio of total  
212 PCa in the weighted median approach was 1.091 per sd increase in genetically predicted Lp(a)  
213 (95% CI  $= [1.028, 1.157]$ ;  $P = 0.004$ ). Alteration of the clumping threshold in sensitivity analysis 1  
214 resulted in a higher number of IVs fitted to our model, which supported a significant effect estimate

215 for Lp(a) in both the IVW ( $OR_{IVW} = 1.076$ ; 95% CI = [1.016,1.114];  $P=0.012$ ) and weighted median  
216 approaches ( $OR_{weighted\ median} = 1.066$ ; 95% CI = [1.012,1.123];  $P=0.016$ ). Sensitivity analysis 2,  
217 which included IVs according to the Burgess et al. paper [27], also supported a relationship between  
218 genetically elevated Lp(a) and total PCa ( $OR_{IVW}=1.037$ ; 95% CI = [1.009,1.066];  $P=0.010$ ,  
219  $OR_{weighted\ median} = 1.044$ ; 95% CI = [1.026,1.061];  $P=6.58 \times 10^{-7}$ ). Sensitivity analysis 3, which  
220 involved variants located in the *LPA* gene only, supported an even stronger OR  
221 ( $OR_{weighted\ median} = 1.439$ ; 95% CI = [1.280,1.619];  $P=1.80 \times 10^{-9}$ ) for total PCa per sd increase in  
222 genetically predicted Lp(a).

### 223 **Advanced PCa**

224 The univariable MR analysis did not reveal any significant association between blood lipids and  
225 advanced PCa risk (HDL;  $OR_{IVW} = 0.977$ ; 95% CI = [0.905,1.051];  $P = 0.552$ , LDL;  $OR_{IVW}=1.067$ ;  
226 95% CI = [0.970,1.174]);  $P=0.191$ , TG;  $OR_{IVW} = 1.004$ ; 95% CI = [0.923,1.094];  $P = 0.921$ , Lp(a) ;  
227  $OR_{IVW} = 1.064$ ; 95% CI = [0.910,1.245];  $P = 0.435$ , ApoA;  $OR_{IVW} = 1.001$ ; 95% CI = [0.932,1.073],  
228  $P = 0.991$ , ApoB;  $OR_{IVW} = 0.992$ ; 95% CI = [0.914,1.073],  $P = 0.837$ ) (Table 1, Table S5).  
229 However, sensitivity analysis 2 for Lp(a), which included IVs of the Burgess et al. paper [27],  
230 supported an association between genetically elevated Lp(a) ( $OR_{weighted\ median} = 1.033$ ; 95% CI=  
231 [1.001,1.065];  $P=0.046$ ) and advanced PCa. In addition, sensitivity analysis 3, which restricted to  
232 variants in the *LPA* gene, supported an association between genetically elevated Lp(a) and advanced  
233 PCa ( $OR_{weighted\ median} = 1.388$ ; 95% CI = [1.213,1.590];  $P=2.14 \times 10^{-6}$ ).

### 234 **Early age onset of PCa**

235 Genetically predicted Lp(a) was associated with an increased risk of early age onset PCa in the  
236 main univariable analysis ( $OR_{weighted\ median} = 1.257$ ; 95% CI = [1.107,1.426];  $P=4.00 \times 10^{-4}$ )  
237 (Table 1). All univariable sensitivity analyses performed, confirmed a significant relationship  
238 between genetically elevated Lp(a) and early age onset of PCa. [(Sensitivity analysis 1;  $OR_{IVW} =$   
239  $1.215$ ; 95% CI = [1.096, 1.349];  $P=2.03 \times 10^{-4}$ ,  $OR_{weighted\ median} = 1.217$ ; 95% CI = [1.084,1.365];

240  $P=0.001$ ), (Sensitivity analysis 2;  $OR_{IVW} = 1.076$ ; 95% CI = [1.027,1.126];  $p$ -value=0.002,  
 241  $OR_{weighted\ median} = 1.079$ ; 95% CI =[1.036,1.124];  $P=3.20 \times 10^{-4}$ ), (Sensitivity analysis 3;  
 242  $OR_{weighted\ median} = 1.502$ ; 95% CI = [1.276,1.770];  $P=1.04 \times 10^{-6}$ )].  
 243

		<i>Method</i>	<i>OR</i>	<i>95% CI</i>	<i>P</i>	<i>Padjusted†</i>
<i>Total PCa</i>	<b>Main Analysis</b>	IVW	1.066	[0.909,1.249]	0.431	1
		<b>Weighted Median</b>	<b>1.091</b>	<b>[1.028,1.157]</b>	<b>0.004**</b>	<b>0.024*</b>
	<b>Sensitivity Analysis 1</b>	IVW	<b>1.076</b>	<b>[1.016,1.114]</b>	<b>0.012*</b>	-
		<b>Weighted Median</b>	<b>1.066</b>	<b>[1.012,1.123]</b>	<b>0.016*</b>	-
	<b>Sensitivity Analysis 2</b>	IVW	<b>1.037</b>	<b>[1.009,1.066]</b>	<b>0.010**</b>	-
		<b>Weighted Median</b>	<b>1.044</b>	<b>[1.026,1.061]</b>	<b>6.582x10<sup>-7</sup>***</b>	-
	<b>Sensitivity Analysis 3</b>	IVW	1.228	[0.960,1.570]	0.104	-
		<b>Weighted Median</b>	<b>1.439</b>	<b>[1.280,1.619]</b>	<b>1.799x10<sup>-9</sup>***</b>	-
	<i>Advanced PCa</i>	<b>Main Analysis</b>	IVW	1.064	[0.910,1.245]	0.435
<b>Weighted Median</b>			1.071	[0.973,1.179]	0.158	0.948
<b>Sensitivity Analysis 1</b>		IVW	1.051	[0.981,1.127]	0.158	-
		<b>Weighted Median</b>	1.051	[0.974,1.135]	0.197	-
<b>Sensitivity Analysis 2</b>		IVW	1.024	[0.992,1.058]	0.139	-
		<b>Weighted Median</b>	<b>1.033</b>	<b>[1.001,1.065]</b>	<b>0.046*</b>	-
<b>Sensitivity Analysis 3</b>		IVW	1.226	[0.957,1.570]	0.107	-
		<b>Weighted Median</b>	<b>1.388</b>	<b>[1.213,1.590]</b>	<b>2.138x10<sup>-6</sup>***</b>	-
<i>Early age onset PCa</i>		<b>Main Analysis</b>	IVW	1.169	[0.927,1.473]	0.188
	<b>Weighted Median</b>		<b>1.257</b>	<b>[1.107,1.426]</b>	<b>4.000x10<sup>-4</sup>***</b>	-
	<b>Sensitivity Analysis 1</b>	IVW	<b>1.215</b>	<b>[1.096,1.349]</b>	<b>2.028x10<sup>-4</sup>***</b>	-
		<b>Weighted Median</b>	<b>1.217</b>	<b>[1.084,1.365]</b>	<b>0.001***</b>	-
	<b>Sensitivity Analysis 2</b>	IVW	<b>1.076</b>	<b>[1.027,1.126]</b>	<b>0.002**</b>	-
		<b>Weighted Median</b>	<b>1.079</b>	<b>[1.036,1.124]</b>	<b>3.201x10<sup>-4</sup>***</b>	-
	<b>Sensitivity Analysis 3</b>	IVW	1.481	[0.907,2.418]	0.116	-
		<b>Weighted Median</b>	<b>1.502</b>	<b>[1.276,1.770]</b>	<b>1.038x10<sup>-6</sup>***</b>	-

244 **Table 1. Univariable estimates of genetically predicted Lp(a) on each PCa outcome.**

## 245 **Multivariable MR**

246 As an attempt to control for pleiotropic pathways that could arise from the relationship between  
247 different lipid traits, we incorporated a MVMR model including Lp(a), HDL, LDL and TG jointly  
248 as exposures for each PCa outcome. The significant association observed between genetically  
249 predicted LDL and total PCa in univariable MR attenuated in the multivariable MR model and was  
250 no longer significant (OR = 1.052; 95% CI = [0.973,1.134]; P=0.183) (Table 2). However, after  
251 adjusting for HDL, LDL and TG, genetically predicted Lp(a) remained significantly and positively  
252 associated with total PCa (OR = 1.068; 95% CI = [1.005,1.134]; P=0.034). Additional adjustment  
253 for BMI led to an almost unaltered OR for total PCa risk per sd increase in genetically predicted  
254 Lp(a) (OR = 1.066; 95% CI = [1.008,1.129]; P=0.026) (Table S7). Genetically predicted Lp(a) was  
255 associated at borderline significance with advanced PCa after adjusting for multiple lipid traits  
256 (OR= 1.078; 95% CI = [0.999,1.163]; P=0.055). Additional adjustment for BMI led to an OR of  
257 1.075 (95% CI: [1,1.155]; P=0.050). Genetically elevated Lp(a) remained significantly associated  
258 with early age onset of PCa (OR= 1.150; 95% CI = [1.015,1.303]; P = 0.028), in agreement with the  
259 univariable MR analysis. Adjustment for BMI yielded a similar effect of 1.155 (95% CI =  
260 [1.029,1.297]; P = 0.015). IVW estimates for all lipids from the MVMR can be seen in Table 2

Results are based on the inverse variance weighting (IVW) and weighted median estimation methods. ORs for each PCa outcome are reported per sd increase in genetically predicted Lp(a). The main analysis included IVs based on a clumping threshold of 0.001, sensitivity analysis 1 is based on an eased clumping threshold of 0.01, sensitivity analysis 2 is based on a different IV set from another paper and finally sensitivity analysis 3 is based upon variants located in the *LPA* gene. Associations of P<0.05 are shown in bold. .P ∈ (0.05,0.1], \* P ∈ (0.01,0.05], \*\* P ∈ (0.001,0.01], \*\*\* P ∈ (0,0.001].

† Padjusted is the Bonferroni corrected p-value, taking into account the total number of tests performed in the main analysis for total and advanced PCa. This reflects a total of 5 univariable analyses performed on each outcome (one for each lipid). For the early age onset of PCa only Lp(a) was fitted to the model and thus no correction was applied to the p-value.

261 below, whereas the IVW BMI-adjusted results can be found in Table S7. We compared the  
262 multivariable Lp(a) estimates from all the analyses performed on total, advanced and early age

263 onset PCa with the univariable estimates and additional sensitivity analyses through a panel of three  
 264 distinct forest plots (Fig 1). IVs according to variants in the *LPA* gene (Sensitivity Analysis 3),  
 265 supported the strongest effect between genetically predicted Lp(a) concentrations and each PCa  
 266 outcome.

267

268

269

270

271

272 **Table 2.**

**MVMR**

273 **results for**

**each PCa**

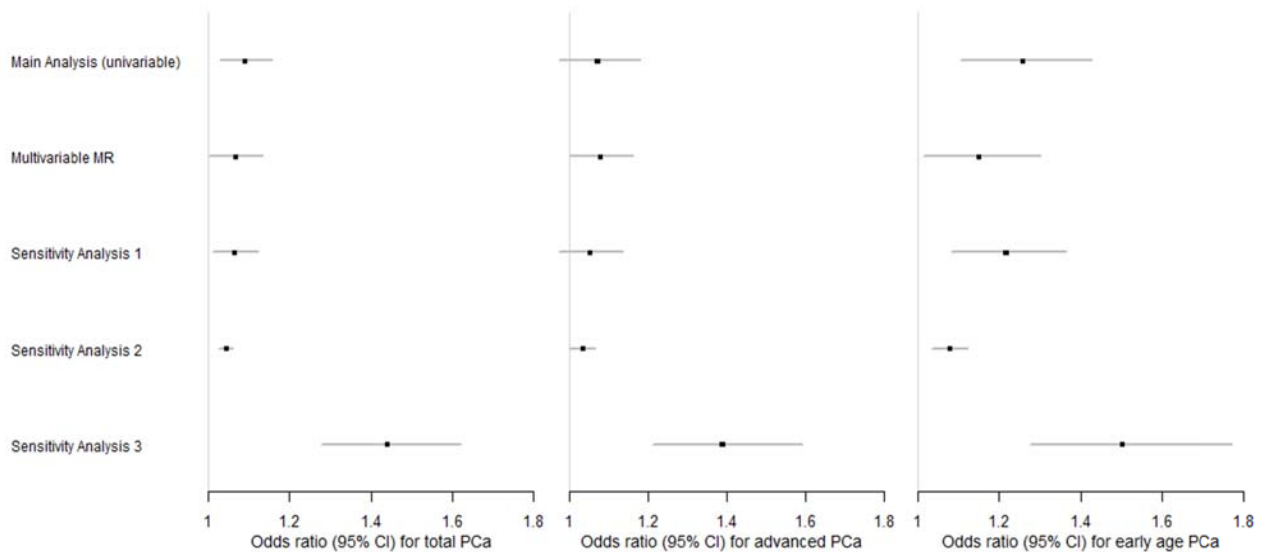
274 **outcome.**

	<i>Biomarker</i>	<i>OR</i>	<i>95% CI</i>	<i>P</i>
<i>Total PCa</i>	HDL	1.009	[0.946,1.077]	0.775
	LDL	1.052	[0.973,1.134]	0.183
	TG	1.026	[0.953,1.106]	0.485
	<b>Lp(a)</b>	<b>1.068</b>	<b>[1.005,1.134]</b>	<b>0.034*</b>
<i>Advanced PCa</i>	HDL			0.872
		0.993	[0.917,1.077]	
	LDL	1.007	[0.916,1.106]	0.892
	TG	1.003	[0.914,1.101]	0.952
	Lp(a)	1.078	[0.999,1.163]	0.055.
<i>Early age onset PCa</i>	HDL			0.732
		1.024	[0.894,1.174]	
	LDL	1.112	[0.948,1.305]	0.192
	TG	1.062	[0.908,1.242]	0.450
	<b>Lp(a)</b>	<b>1.150</b>	<b>[1.015,1.303]</b>	<b>0.028*</b>

284

Each estimate (OR) is based on the multivariable IVW method and represents the direct effect of the risk factor on the respective outcome after controlling for the other three biomarkers in MVMR. ORs are reported per sd increase in the respective biomarker. Genetically elevated Lp(a) is significantly associated with total and early age onset PCa, whereas it is associated also at borderline significance with advanced PCa. Associations of  $P < 0.05$  are shown in bold.  $.P \in (0.05, 0.1]$ ,  $* P \in (0.01, 0.05]$ ,  $** P \in (0.001, 0.01]$ ,  $*** P \in (0, 0.001]$ .

285



286

**Fig 1. Forest plots of the Lp(a) effects observed in different analyses based on each PCa type.**

287

The main and sensitivity analyses estimates are based on the weighted median approach, whereas MVMR includes the IVW estimates. Sensitivity analyses 1-3 refer to the univariable models.

288

Sensitivity analysis 1 is based on an eased clumping threshold of 0.01, sensitivity analysis 2 includes an IV set based on another paper and finally sensitivity analysis 3 is based upon variants located in the *LPA* gene. Each square represents the OR for each PCa outcome, reported per sd increase in the

289

biomarker, with the 95% CI represented by the error bars.

290

## 291 **Discussion**

292 Our MR analyses provide evidence that genetically predicted Lp(a) concentration is associated with  
293 risk of total and early age onset PCa and suggestive evidence for an association with advanced PCa.

294 The majority of the different methods we used for our univariable analyses agreed for total and  
295 early age onset PCa. Lp(a) was only significantly associated with advanced PCa in two of the  
296 univariable sensitivity analyses. IVs located in the *LPA* gene supported the strongest and most  
297 significant Lp(a) associations for total, advanced and early age onset PCa. Given the strong  
298 regulation of Lp(a) levels by the *LPA* gene region [28], the latter findings are based on strong  
299 instruments with a clear biological function. Adjustment for multiple lipid traits and BMI in the  
300 MVMR models, further supported a significant association of genetically predicted Lp(a) on total,  
301 advanced and early age onset PCa.

302 The null associations observed for genetically predicted HDL on total PCa agree with findings from  
303 two previous MR analyses [16][17], and the null findings for TG are also supported in the Bull et al.  
304 [16] paper. In our analysis, there was some evidence for a significant LDL association with total  
305 PCa risk, though this was likely a false indication due to pleiotropy, as suggested by the MVMR  
306 model which indicated no association with LDL. As Lp(a) includes an LDL component [36], the  
307 attenuation of LDL to the null in the MVMR could be attributed to independent actions of Lp(a)  
308 itself, as we did not observe any association between other Lp(a) components and PCa risk.

309 Alternative explanations are that Lp(a) concentrations are less affected by statins compared to LDL  
310 [37], thus genetically predicted Lp(a) may be more accurate for current actual levels than  
311 genetically predicted LDL, or that the association for Lp(a) dominates over LDL due to the high  
312 between-person variability of Lp(a) [20]. The authors of the Bull et al. paper [16] suggested a  
313 potential role of LDL and TG in advanced/high-grade PCa; our findings for LDL and TG in  
314 advanced PCa risk are not in agreement with theirs. Our analyses included adjustment for multiple  
315 lipid traits in contrast with the previously mentioned papers, which we believe plays a vital role in  
316 MR analysis modelling blood lipids, given the high correlation between them. As far as we are

317 aware, no previous MR study has investigated the role of apoA and apoB in PCa risk. Our null  
318 results are nonetheless in agreement with observational studies by Katzke et al [8], which involved  
319 the prospective EPIC–Heidelberg cohort and Borgquist et al [38], which was based on the  
320 prospective Malmö Diet and Cancer Study (MDCS).

321 To the best of our knowledge, no previous MR study has examined the role of Lp(a) in PCa risk.  
322 The positive association observed for Lp(a) in total PCa was supported by the observational study  
323 of Katzke et al. [8]. Results showed that top vs. bottom quartile levels of Lp(a) were associated with  
324 a 47% higher risk of PCa (OR=1.47; 95% CI = 1.06-2.04). Wang F et al. [39], another observational  
325 study, examined the role of Lp(a) in high-risk PCa via a multivariable regression adjusted for age,  
326 BMI, hypertension, diabetes, coronary artery disease, and lipid-lowering drugs. They observed that  
327 high Lp(a) levels were positively associated (OR<sub>Q4 vs. Q1</sub>=2.687; 95% CI = 1.113-6.491; P = 0.028)  
328 with high-risk PCa, which agrees with our findings for advanced PCa in the MVMR analysis. In  
329 addition, a recent large prospective cohort among 211,754 men in UK Biobank [40] observed a  
330 suggestive positive association between Lp(a) and PCa risk (hazard ratio<sub>per SD</sub> =1.02; 95% CI:  
331 [0.99,1.06]). Our literature review did not reveal any studies investigating the role of Lp(a) in early  
332 age onset PCa.

333 A range of different biological mechanisms have been proposed to explain pro-cancer effects of  
334 cholesterol at the cellular level, including cell proliferation, inflammation, membrane organization,  
335 and steroidogenesis [41]. It is unclear whether total cholesterol or any lipoprotein particle is the  
336 causal factor, and the potential pathophysiological mechanisms of Lp(a) have not been well-studied.  
337 However, emerging evidence from the cardiovascular literature supports pleiotropic functions of  
338 Lp(a) and complex mediation pathways with other lipid particles [42]. Lp(a) is highly heritable  
339 (heritability=24%) [20], with the majority of individuals having low Lp(a) levels. However, African  
340 Americans which are known to have the highest risk for PCa, tend to also have higher circulating  
341 Lp(a) levels [28]. It has been previously observed that mean Lp(a) concentrations for African  
342 Americans are 106 (60-180) nmol/l, whereas Caucasians such as non-Hispanic whites have mean



343 Lp(a) concentrations of 24 (7.2-79.2) nmol/l [43]. Although the exact explanation behind ethnic  
344 discrepancies in PCa is currently unknown, it has been hypothesised that access to health care may  
345 play a partial role in this. Yet, given that disparities in PCa risk are apparent regardless of cancer  
346 detection issues, it is likely that biological factors are key drivers of this phenomenon [44]. Two  
347 recent papers have further provided evidence of a different immune response [45] and inflammatory  
348 signalling [46] for African Americans vs. Caucasians, which can be linked to their poorer PCa  
349 prognosis. Considering Lp(a) as a modifier of the immune/inflammatory response [47], the  
350 increased Lp(a) concentrations in African Americans and our observed association between  
351 genetically elevated Lp(a) and PCa, we hypothesise that Lp(a) may partially account for some of the  
352 observed discrepancies in PCa risk by ethnicity. Future large-scale genomic studies in African  
353 ancestry populations [48] would be required to evaluate the hypothesis that Lp(a) can explain  
354 discrepancies in PCa risk by race.

355 We note several limitations to our research. Currently, there is no direct way to prove that the  
356 second and third MR assumptions hold and as such, violations would result in biased MR estimates.  
357 A large number of robust methods and sensitivity analyses were used to probe into potential  
358 violations mainly due to horizontal pleiotropy, but its presence cannot be excluded. The samples  
359 analysed for our main MR analyses were restricted to Europeans to avoid issues with heterogeneity,  
360 which is required for a two-sample MR [49]. However, this may affect generalisability of the  
361 results, which are restricted to those of European ancestry. The number of variants associated with  
362 Lp(a) was limited in comparison to other lipids. Initially, 5894 variants were identified to be  
363 associated with Lp(a) at GWAS significance, whereas all other lipids had more than 10,000  
364 associated variants. This then resulted in a final sample size of 10 variants due to LD clumping in  
365 the main univariable analysis, which may have decreased our statistical power. However, after  
366 relaxing the LD clumping threshold in our sensitivity analyses, we included more variants, the  
367 findings of which corroborated the main results. In addition, some previous observational studies  
368 have suggested potential threshold effects for cholesterol concentrations and PCa [50,51], which

369 cannot be studied in two-sample MR with summary-level data, and future one-sample MR studies  
370 are warranted.

371 Apart from the caveats in our study, there are also several strengths that should be noted. We used  
372 an MR study design, in which the outcome of interest is compared between genotypes, analogous to  
373 that between treatment and placebo groups in a randomised controlled trial. However, inference  
374 should be made with great caution as alterations of genetically predicted risk factors are not  
375 identical to those due to a drug or dietary intervention [52]. Secondly, as lipids are dependent on  
376 each other for their main functionalities [18], it is important to control for pleiotropic pathways that  
377 may arise from these dependencies. One method to do so, is via the use of MVMR which allows to  
378 include genetic information on exposures that may correlate with each other into a joint  
379 multivariable model [53] and our study forms the first such MVMR conducted to investigate the  
380 relationship between various lipid traits and PCa risk. Thirdly, the use of UK Biobank data allowed  
381 us to include information on underexamined lipid traits such as Lp(a), apoA and apoB, in  
382 comparison to previous PCa studies which mainly considered HDL, LDL and TG. In addition, we  
383 have sex-specific genetic associations and this allowed us to work with male-specific data which are  
384 more relevant to PCa. Finally, our analyses are based on large sample sizes which were acquired  
385 from UK Biobank [19] and the PRACTICAL consortium [21].

386 In summary, findings from this study point towards a positive association between genetically  
387 predicted Lp(a) concentrations and risk of total, advanced and early age onset PCa. Screening for  
388 high Lp(a) concentrations could possibly be investigated in the future to identify high-risk groups  
389 for PCa. Given that Lp(a) concentrations depend significantly on genetics [54], modification of  
390 Lp(a) levels may be achieved by developing Lp(a)-lowering drugs [55] that might be on the  
391 horizon. A personalised approach in repurposing lipid drugs that target Lp(a) directly for high-risk  
392 individuals could consequently be considered, upon replication of our findings, to study their  
393 effectiveness against PCa prevention. The mechanisms behind the observed association remain  
394 however unclear given the uncertainty underlining the pleiotropic physiological functions of the

395 *LPA* gene itself, which controls about 70-90% of the Lp(a) variability [36,54]. Further research into  
396 this complex gene such as colocalization analysis would be required to understand more of its  
397 functionality and consequently its role in PCa risk.

## 398 **References**

- 399 1. Rawla P. Epidemiology of Prostate Cancer. *World J Oncol.* 2019;10(2):63–89.
- 400 2. Cancer R. Global cancer observatory [Internet]. [cited 2020 Jun 26]. Available from:  
401 <https://gco.iarc.fr/>
- 402 3. Loda M, Mucci LA, Mittelstadt ML, Van Hemelrijck M, Cotter MB. Pathology and  
403 epidemiology of cancer. In: *Pathology and Epidemiology of Cancer.* 2016. p. 1–670.
- 404 4. Pernar CH, Ebot EM, Wilson KM, Mucci LA. The Epidemiology of Prostate Cancer. *Cold*  
405 *Spring Harb Perspect Med.* 2018;8(12).
- 406 5. Tse LA, Ho WM, Wang F, He YH, Ng CF. Environmental risk factors of prostate cancer: a  
407 case-control study. *Hong Kong Med J = Xianggang yi xue za zhi.* 2018;24(4):30–3.
- 408 6. Moyad MA. Preventing aggressive prostate cancer with proven cardiovascular disease  
409 preventive methods. *Asian J Androl.* 2015;17(6):874–7.
- 410 7. Hurwitz LM, Agalliu I, Albanes D, Barry KH, Berndt SI, et al. Recommended Definitions of  
411 Aggressive Prostate Cancer for Etiologic Epidemiologic Research. *JNCI J Natl Cancer Inst.*  
412 2021;113(6):727–34.
- 413 8. Katzke VA, Sookthai D, Johnson T, Kühn T, Kaaks R. Blood lipids and lipoproteins in  
414 relation to incidence and mortality risks for CVD and cancer in the prospective EPIC-  
415 Heidelberg cohort. *BMC Med.* 2017;15(1).
- 416 9. Jamnagerwalla J, Howard LE, Allott EH, Vidal AC, Moreira DM, Castro-Santamaria R, et al.  
417 Serum cholesterol and risk of high-grade prostate cancer: Results from the REDUCE study.  
418 *Prostate Cancer Prostatic Dis.* 2018;21(2):252–9.
- 419 10. Yu Peng L, Yu Xue Z, Fei LP, Cheng C, Ya Shuang Z, Da Peng L, et al. Cholesterol levels

- 420 in blood and the risk of prostate cancer: A meta-analysis of 14 prospective studies. *Cancer*  
421 *Epidemiol Biomarkers Prev.* 2015;24(7):1086–93.
- 422 11. Bansal D, Undela K, D’Cruz S, Schifano F. Statin Use and Risk of Prostate Cancer: A Meta-  
423 Analysis of Observational Studies. *PLoS One.* 2012;7(10).
- 424 12. Tan P, Wei S, Tang Z, Gao L, Zhang C, Nie P, et al. LDL-lowering therapy and the risk of  
425 prostate cancer: A meta-analysis of 6 randomized controlled trials and 36 observational  
426 studies. *Sci Rep.* 2016;6.
- 427 13. Yarmolinsky J, Wade KH, Richmond RC, Langdon RJ, Bull CJ, Tilling KM, et al. Causal  
428 inference in cancer epidemiology: What is the role of mendelian randomization? *Cancer*  
429 *Epidemiol Biomarkers Prev.* 2018;27(9):995–1010.
- 430 14. Davies NM, Holmes M V., Davey Smith G. Reading Mendelian randomisation studies: A  
431 guide, glossary, and checklist for clinicians. *BMJ.* 2018;362.
- 432 15. Zheng J, Baird D, Borges M-C, Bowden J, Hemani G, Haycock P, et al. Recent  
433 Developments in Mendelian Randomization Studies. *Curr Epidemiol Reports.*  
434 2017;4(4):330–45.
- 435 16. Bull CJ, Bonilla C, Holly JMP, Perks CM, Davies N, Haycock P, et al. Blood lipids and  
436 prostate cancer: a Mendelian randomization analysis. *Cancer Med.* 2016;5(6):1125–36.
- 437 17. Adams CD, Richmond R, Santos Ferreira DL, Spiller W, Tan V, Zheng J, et al. Circulating  
438 metabolic biomarkers of screen-detected prostate cancer in the ProtecT study. *Cancer*  
439 *Epidemiol Biomarkers Prev.* 2019;28(1):208–16.
- 440 18. Dominiczak MH, Caslake MJ. Apolipoproteins: Metabolic role and clinical biochemistry  
441 applications. *Ann Clin Biochem.* 2011;48(6):498–515.
- 442 19. Lab N. GWAS of UK Biobank biomarker measurements - Neale lab [Internet]. 2020 [cited  
443 2020 Aug 16]. Available from: [http://www.nealelab.is/blog/2019/9/16/biomarkers-gwas-](http://www.nealelab.is/blog/2019/9/16/biomarkers-gwas-results)  
444 [results](http://www.nealelab.is/blog/2019/9/16/biomarkers-gwas-results)
- 445 20. Sinnott-Armstrong N, Tanigawa Y, Amar D, Mars N, Benner C, Aguirre M, et al. Genetics

- 446 of 35 blood and urine biomarkers in the UK Biobank. *Nat Genet.* 2021;
- 447 21. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, et al.  
448 Association analyses of more than 140,000 men identify 63 new prostate cancer  
449 susceptibility loci. *Nat Genet.* 2018;50(7):928–36.
- 450 22. Labrecque J, Swanson SA. Understanding the Assumptions Underlying Instrumental  
451 Variable Analyses: a Brief Review of Falsification Strategies and Related Tools. *Curr*  
452 *Epidemiol Reports.* 2018;5(3):214–20.
- 453 23. Rees JMB, Wood AM, Dudbridge F, Burgess S. Robust methods in Mendelian  
454 randomization via penalization of heterogeneous causal estimates. *PLoS One.* 2019;14(9).
- 455 24. Burgess S, Butterworth A TS. Mendelian randomization analysis with multiple genetic  
456 variants using summarized data. *Genet Epidemiol.* 2013;37(7):658–65.
- 457 25. Bowden J, Fabiola Del Greco M, Minelli C, Smith GD, Sheehan NA, Thompson JR.  
458 Assessing the suitability of summary data for two-sample mendelian randomization analyses  
459 using MR-Egger regression: The role of the I<sup>2</sup> statistic. *Int J Epidemiol.* 2016;45(6):1961–  
460 74.
- 461 26. Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable  
462 Mendelian randomization to correct for both measured and unmeasured pleiotropy. *Stat Med.*  
463 2017;36(29):4705–18.
- 464 27. Burgess S, Ference BA, Staley JR, Freitag DF, Mason AM, Nielsen SF, et al. Association of  
465 LPA variants with risk of coronary disease and the implications for lipoprotein(a)-lowering  
466 therapies: A mendelian randomization analysis. *JAMA Cardiol.* 2018;3(7):619–27.
- 467 28. Kronenberg F, Utermann G. Lipoprotein(a): Resurrected by genetics. *J Intern Med.*  
468 2013;273(1):6–30.
- 469 29. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in  
470 causal relationships inferred from Mendelian randomization between complex traits and  
471 diseases. *Nat Genet.* 2018;50(5):693–8.

- 472 30. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM. A robust and efficient method for  
473 Mendelian randomization with hundreds of genetic variants. *Nat Commun.* 2020;11(1).
- 474 31. Maranhão RC, Carvalho PO, Strunz CC, Pileggi F. Lipoprotein (a): Structure,  
475 pathophysiology and clinical implications. *Arq Bras Cardiol.* 2014;103(1):76–84.
- 476 32. Wang A, Lazo M, Ballentine Carter H, Groopman JD, Nelson WG, Platz EA. Association  
477 between Liver Fibrosis and Serum PSA among U.S. Men: National Health and Nutrition  
478 Examination Survey (NHANES), 2001–2010. *Cancer Epidemiol Biomarkers Prev.*  
479 2019;28(8):1331–8.
- 480 33. Bañez LL, Loftis RM, Freedland SJ, Presti JC, Aronson WJ, Amling CL, et al. The influence  
481 of hepatic function on prostate cancer outcomes after radical prostatectomy. *Prostate Cancer*  
482 *Prostatic Dis.* 2010;13(2):173–7.
- 483 34. Hopewell JC, Haynes R, Baigent C. The role of lipoprotein (a) in chronic kidney disease. *J*  
484 *Lipid Res.* 2018;59(4):577–85.
- 485 35. Weinstein SJ, Mackrain K, Stolzenberg-Solomon RZ, Selhub J, Virtamo J, Albanes D.  
486 Serum creatinine and prostate cancer risk in a prospective study. *Cancer Epidemiol*  
487 *Biomarkers Prev.* 2009;18(10):2643–9.
- 488 36. Schmidt K, Noureen A, Kronenberg F, Utermann G. Structure, function, and genetics of  
489 lipoprotein (a). *J Lipid Res.* 2016;57(8):1339–59.
- 490 37. Van Capelleveen JC, Van Der Valk FM, Stroes ESG. Current therapies for lowering  
491 lipoprotein (a). *J Lipid Res.* 2016;57(9):1612–8.
- 492 38. Borgquist S, Butt T, Almgren P, Shiffman D, Stocks T, Orho-Melander M, et al.  
493 Apolipoproteins, lipids and risk of cancer. *Int J Cancer.* 2016;138(11):2648–56.
- 494 39. Wang FM, Zhang Y. High Lipoprotein(a) Level Is Independently Associated with Adverse  
495 Clinicopathological Features in Patients with Prostate Cancer. *Dis Markers.* 2019;2019.
- 496 40. Perez-Cornago A, Fensom GK, Andrews C, Watts EL, Allen NE, Martin RM, et al.  
497 Examination of potential novel biochemical factors in relation to prostate cancer incidence

- 498 and mortality in UK Biobank. *Br J Cancer*. 2020;123(12):1808–17.
- 499 41. Murtola TJ, Syväälä H, Pennanen P, Bläuer M, Solakivi T, Ylikomi T, et al. The importance  
500 of LDL and Cholesterol metabolism for prostate epithelial cell growth. *PLoS One*. 2012;7(6).
- 501 42. Boffa MB, Koschinsky ML. Oxidized phospholipids as a unifying theory for lipoprotein(a)  
502 and cardiovascular disease. *Nat Rev Cardiol*. 2019;16(5):305–18.
- 503 43. Enkhmaa B, Anuurad E, Berglund L. Lipoprotein (a): Impact by ethnicity and environmental  
504 and medical conditions. *J Lipid Res*. 2016;57(7):1111–25.
- 505 44. Rebbeck TR. Prostate cancer disparities by race and ethnicity: From nucleotide to  
506 neighborhood. *Cold Spring Harb Perspect Med*. 2018;8(9).
- 507 45. Thomas JK, Mir H, Kapur N, Singh S. Racial differences in immunological landscape  
508 modifiers contributing to disparity in prostate cancer. *Cancers (Basel)*. 2019;11(12).
- 509 46. Tang W, Wallace TA, Yi M, Magi-Galluzzi C, Dorsey TH, Onabajo OO, et al. IFNL4-ΔG  
510 allele is associated with an interferon signature in tumors and survival of African-American  
511 men with prostate cancer. *Clin Cancer Res*. 2018;24(21):5471–81.
- 512 47. Orsó E, Schmitz G. Lipoprotein(a) and its role in inflammation, atherosclerosis and  
513 malignancies. *Clin Res Cardiol Suppl*. 2017;12:31–7.
- 514 48. Fatumo S. The opportunity in African genome resource for precision medicine.  
515 *EBioMedicine*. 2020;54.
- 516 49. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Smith GD. Best (but oft-forgotten)  
517 practices: The design, analysis, and interpretation of Mendelian randomization studies. *Am J*  
518 *Clin Nutr*. 2016;103(4):965–78.
- 519 50. Platz EA, Till C, Goodman PJ, Parnes HL, Figg WD, Albanes D, et al. Men with low serum  
520 cholesterol have a lower risk of high-grade prostate cancer in the placebo arm of the prostate  
521 cancer prevention trial. *Cancer Epidemiol Biomarkers Prev*. 2009;18(11):2807–13.
- 522 51. Platz EA, Clinton SK, Giovannucci E. Association between plasma cholesterol and prostate  
523 cancer in the PSA era. *Int J Cancer*. 2008;123(7):1693–8.

- 524 52. Burgess S, Butterworth A, Malarstig A, Thompson SG. Use of Mendelian randomisation to  
525 assess potential benefit of clinical intervention. *BMJ*. 2012;345.
- 526 53. Zuber V, Colijn JM, Klaver C, Burgess S. Selecting causal risk factors from high-throughput  
527 experiments using multivariable mendelian randomization. *bioRxiv*. 2018;
- 528 54. Kronenberg F. Human Genetics and the Causal Role of Lipoprotein(a) for Various Diseases.  
529 *Cardiovasc Drugs Ther*. 2016;30(1):87–100.
- 530 55. Viney NJ, van Capelleveen JC, Geary RS, Xia S, Tami JA, Yu RZ, et al. Antisense  
531 oligonucleotides targeting apolipoprotein(a) in people with raised lipoprotein(a): two  
532 randomised, double-blind, placebo-controlled, dose-ranging trials. *Lancet*.  
533 2016;388(10057):2239–53.

## 534 **Supporting information**

535 **S1 Table. Heritability of each blood lipid as estimated in a GWAS of Sinnott-Armstrong et al**  
536 **[20]**. The estimates represent total heritability and the methodology followed was the Heritability  
537 estimation summary statistics (HESS). The number of SNPS fitted in each model for the main  
538 univariable analysis is also reported.

539 **S2 Table. Descriptive statistics of each blood lipid.** Lp(a) is positively skewed, whereas the rest  
540 of the lipids are approximately normally distributed. Measurements are based on all samples (both  
541 sexes) in UK Biobank [19].

542 **S3 Table. MR-Egger, MR-Presso and Contamination Mixture estimates for Lp(a) on each**  
543 **PCa outcome.** The contamination mixture method may indicate two distinct CIs associated with a  
544 single estimate. Associations of  $P < 0.05$  are shown in bold.  $.P \in (0.05, 0.1]$ ,  $* P \in (0.01, 0.05]$ ,  $** P \in$   
545  $(0.001, 0.01]$ ,  $*** P \in (0, 0.001]$ .

546 **S4 Table. Univariable MR estimates for each lipid on total PCa.** Genetically elevated LDL is  
547 significantly associated with total PCa only through the IVW approach, whereas the pleiotropy-



548 robust methods do not support this association. Associations of  $P < 0.05$  are shown in bold.  $P \in$   
549  $(0.05, 0.1]$ , \*  $P \in (0.01, 0.05]$ , \*\*  $P \in (0.001, 0.01]$ , \*\*\*  $P \in (0, 0.001]$ .

550 **S5 Table. Univariable MR estimates for each lipid on advanced PCa.** None of these lipids are  
551 associated with advanced PCa.

552 **S6 Table. MR-Egger estimates for Lp(a) and each PCa outcome.** Sensitivity Analyses refer to  
553 the univariable MR and are as follows; Sensitivity Analysis 1 includes variants according to an  
554 eased clumping threshold of  $r^2 < 0.01$ , sensitivity analysis 2 includes SNPs for Lp(a) according to a  
555 different IV set, whereas sensitivity analysis 3 is based on variants included in the LPA gene.

556 Associations of  $P < 0.05$  are shown in bold.  $P \in (0.05, 0.1]$ , \*  $P \in (0.01, 0.05]$ , \*\*  $P \in (0.001, 0.01]$ ,  
557 \*\*\*  $P \in (0, 0.001]$ .

558 **S7 Table. IVW estimates from the MVMR model adjusted for BMI.** Genetically elevated Lp(a)  
559 is associated with overall, advanced and early onset of PCa in these models. Associations of  $P < 0.05$   
560 are shown in bold.  $P \in (0.05, 0.1]$ , \*  $P \in (0.01, 0.05]$ , \*\*  $P \in (0.001, 0.01]$ , \*\*\*  $P \in (0, 0.001]$ .

561 **S8 Table. IVW and MR-Egger estimates from the total PCa MVMR model adjusted for Lp(a)  
562 and AST.** Genetically elevated Lp(a) is significantly associated with total PCa through the IVW  
563 method. Associations of  $P < 0.05$  are shown in bold.  $P \in (0.05, 0.1]$ , \*  $P \in (0.01, 0.05]$ , \*\*  $P \in$   
564  $(0.001, 0.01]$ , \*\*\*  $P \in (0, 0.001]$ .

565 **S9 Table. IVW and MR-Egger estimates from the total PCa MVMR model adjusted for Lp(a)  
566 and creatinine.** Genetically elevated Lp(a) is significantly associated with total PCa through the  
567 IVW method and associated at borderline significance with total PCa with the MR-Egger approach.  
568 Associations of  $P < 0.05$  are shown in bold.  $P \in (0.05, 0.1]$ , \*  $P \in (0.01, 0.05]$ , \*\*  $P \in (0.001, 0.01]$ ,  
569 \*\*\*  $P \in (0, 0.001]$ .

570 **S10 Table. Association between SNPs and each lipid biomarker and BMI.** The following list  
571 includes IVs used in the main analysis, MVMR (adjusted for HDL, LDL, TG, Lp(a) and BMI) and  
572 sensitivity analyses 1-4.

573 **S11 Table. Association between SNPs and total PCa.** The following list includes IVs used in the  
574 main analysis, MVMR (adjusted for HDL,LDL,TG,Lp(a) and BMI) and sensitivity analyses 1-4 .

575 **S12 Table. Association between SNPs and advanced PCa.** The following list includes IVs used  
576 in the main analysis, MVMR (adjusted for HDL,LDL,TG,Lp(a) and BMI) and sensitivity analyses  
577 1-4 .

578 **S13 Table. Association between SNPs and early age onset PCa.** The following SNPs include IVs  
579 used in the main analysis, MVMR (adjusted for HDL,LDL,TG,Lp(a) and BMI) and sensitivity  
580 analyses 1-4 .

581 **S14 Table. Association between SNPs, Lp(a), AST and total PCa.** The following SNPs include  
582 IVs used in the additional MVMR analysis on Lp(a), AST and total PCa only.

583 **S15 Table. Association between SNPs, Lp(a), creatinine and total PCa.** The following SNPs  
584 include IVs used in the additional MVMR analysis on Lp(a), creatinine and total PCa only.

585 **S1 File. Members from the PRACTICAL Consortium, CRUK, BPC3, CAPS and PEGASUS**  
586