SUPPLEMENTAL MATERIALS

Subtyping Severe Hypercholesterolemia by Genetic Determinant to Stratify Risk of Coronary Artery Disease

Alexander S.F. Berry, PhD[a], Laney K. Jones, PharmD, MPH[b,c], Eric J. Sijbrands, MD, PhD[d], Samuel S. Gidding, MD[b*], and Matthew T. Oetjens, PhD[a*]

a Autism and Developmental Medicine Institute, Geisinger, Lewisburg, PA 17837
b Department of Genomic Health, Geisinger, Danville, PA 17821
c Heart and Vascular Institute, Geisinger, Danville, PA 17821
d Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, PO-box 2040, 3000 CA Rotterdam, The Netherlands.
*Authors jointly supervised this manuscript

Table of Contents

Major Resources Tables

Animals (in vivo studies)

Species	Vendor or Source	Background Strain	Sex	Persistent ID / URL
NA	NA	NA	NA	NA

Genetically Modified Animals

	Species	Vendor or Source	Background Strain	Other Information	Persistent ID / URL
Parent - Male	NA	NA	NA	NA	NA
Parent - Female	NA	NA	NA	NA	NA

Antibodies

Target antigen	Vendor or Source	Catalog #	Working concentration	Lot # (preferred but not required)	Persistent ID / URL
NA	NA	NA	NA	NA	NA

DNA/cDNA Clones

Clone Name	Sequence	Source / Repository	Persistent ID / URL
NA	NA	NA	NA

Cultured Cells

Name	Vendor or Source	Sex (F, M, or unknown)	Persistent ID / URL
NA	NA	NA	NA

Data & Code Availability

Description	Source / Repository	Persistent ID / URL
UK Biobank EHR, self-report, and genetic data	UK Biobank	NA
Data analysis	Upon request to authors	NA

Other

Description	Source / Repository	Persistent ID / URL
NA	NA	NA

Supplemental Methods

Genotype and Exome Sequence Data

Whole exome sequencing data for 200,602 participants was performed by Regeneron Genetics Center using IDT's xGen probe library and the Illumina NovaSeq 6000 sequencing platform. To identify monogenic causes of FH, we examined exome sequence data collected on the 149,326 UK Biobank participants that passed exclusion criteria. Variants were filtered if they failed to meet QC metrics stored in the VCF file: DP ≥ 10, GQ ≥ 20 and AB ≥ 0.2. To assess functional consequences and pathogenicity of variants, we first annotated the exome sequences with the Ensembl Variant Effect Predictor (release 99)[45] and ClinVar (February 2021)[46].

Exome sequences were screened for FH-causing variants in *LDLR* (NM_000527.4), *APOB* (NM_000384.2), and *PCKS9* (NM_174936.3). An FH-causing variant was defined as 1) a putative loss-of-function (pLOF) variant in *LDLR* except those in the first, penultimate or last exon, 2) a variant classified in ClinVar as pathogenic/likely pathogenic (P/LP) with at least two review stars and at least one submission from a clinical laboratory or 3) a variant interpreted as P/LP by an American Board of Genetics and Genomics (AMBGG)-certified clinical geneticist (n = 49 unique variants) reported in previous publication[47].

LDL-C and Lipoprotein(a)

LDL-C was measured by enzymatic protective selection analysis on a Beckman Coulter AU5800. LDL-C measurements from UK Biobank field 30780 were used. Lp(a) was measured during the baseline interview by immunoturbidimetric analysis on a Randox AU5800. Lp(a) measurements from UK Biobank field 30790 were used. For measurements outside of the reportable range, measurements found in UK Biobank Return 2321 were used.

LDL-C and Lp(a) Polygenic Scores

Genotype array data (UK BiLEVE array) available through the UK Biobank was used to calculate the LDL-C and Lp(a) genetic instruments. These data were imputed to the Haplotype Reference Consortium (HRC) reference panel and available for 200,602 participants[17]. We restricted the HRC genotypes to ~15.0 million biallelic variants with an imputation r2 ≥ 0.3, minor allele frequency [MAF] ≥ 0.001.

Polygenic hypercholesterolemia is defined as a high LDL-C in the presence of a hypercholesterolemia polygenic score in the 90th percentile. We calculated a hypercholesterolemia polygenic score on our sample using the software tool PRS-CS[48]. We set the global shrinkage parameter (phi) of PRS-CS to auto, which means that the parameter is learnt from the data using a fully Bayesian approach. For variant weights, we used the summary statistics from a GWAS of quantitative LDL-C levels (inverse-normal transformed) on a European ancestry sample reported by the Global Lipids Genetics Consortium[7]. The GWAS withheld the UK Biobank sample from the analysis and included 842,660 samples in total. The resulting LDL-C polygenic score included 913,492 SNPs. With this polygenic score in hand, we computed the distribution in the UK Biobank using the PLINK2.0 score function[49].

For our study of genetically predicted Lp(a), we used the same genetic instrument for predicted Lp(a) levels created by Burgess et al.[18] and reported in the UK Biobank by

Trinder et al.[5] We recreated the distribution of the Lp(a) genetic predictor in the UK Biobank from the 43 variant weights using the PLINK2.0 score function. In a generalized linear model, measured Lp(a) ≥125nmol/L correlated with genetically predicted Lp(a) ≥125nmol/L with an area-under-the-curve of 0.86.

LDL-C Adjustment

We expect genetic variables associated with LDL-C – the LDL-C polygenic score and rare monogenic FH variants – to have a stronger correlation with untreated LDL-C values as this is more representative of an individual's baseline physiology, produced in part by genetics. On the other hand, by altering LDL-C levels, LLM introduces noise into statistical models and weakens the association with genetic variables. To determine if adjusting LDL-C by dividing direct LDL-C by 0.7 in those using LLM approximates the untreated state, we compared the correlation between genetic variables and LDL-C levels in a subset of the cohort untreated for LLM (n = 111,896). Next, to test if our adjustment of LDL-C values by LLM use is more representative of untreated LDL-C, we compared the association between genetic variables and LDL-C values before and after adjustment by LLM usage. Adjusted LDL-C values were more strongly associated with the LDL-C polygenic score and monogenic FH compared to the unadjusted values. The correlation between genetic variables and adjusted LDL-C was nearly identical to the correlation in the untreated group, supporting our approach to LDL-C adjustment (Tables S4 and S5).

AHA Lifestyle Score

The American Heart Association's (AHA) Lifestyle Score ranks behaviors based on poor, intermediate, and ideal categories for smoking, BMI, and physical activity. These categories were defined based on the American Heart Association's 2020 Strategic Impact Goal guideline[50]. Here, we defined an ideal diet based on the ideal intake of dietary components for cardiovascular health[51]. The AHA Lifestyle Score was poor if the participant had at least 3 poor lifestyle factors, ideal if a participant had at least 3 ideal lifestyle factors, and intermediate if they had any other combination of lifestyle factors. Lifestyle score was considered an ordered factor with three levels: poor, intermediate, and ideal.

Figure S1. Clinical risk factors and comorbidities associated with genetic severe hypercholesterolemia subtypes compared to the non-genetic subtype. Forest plots show the differences in baseline A) clinical risk factors, and B) prevalence of comorbidities in those with each genetic subtype compared to the non-genetic subtype. Continuous variables were converted to z-scores. Coefficients (z-scores for continuous traits and beta for binary traits) for clinical characteristics and odds ratios for comorbidities with 95% confidence intervals are shown. LLM=lipid-lowering medication; T1D=type 1 diabetes; T2D=type 2 diabetes; CAD=coronary artery disease; CVD=cardiovascular disease; PAD=peripheral artery disease; AVS=aortic valve stenosis).

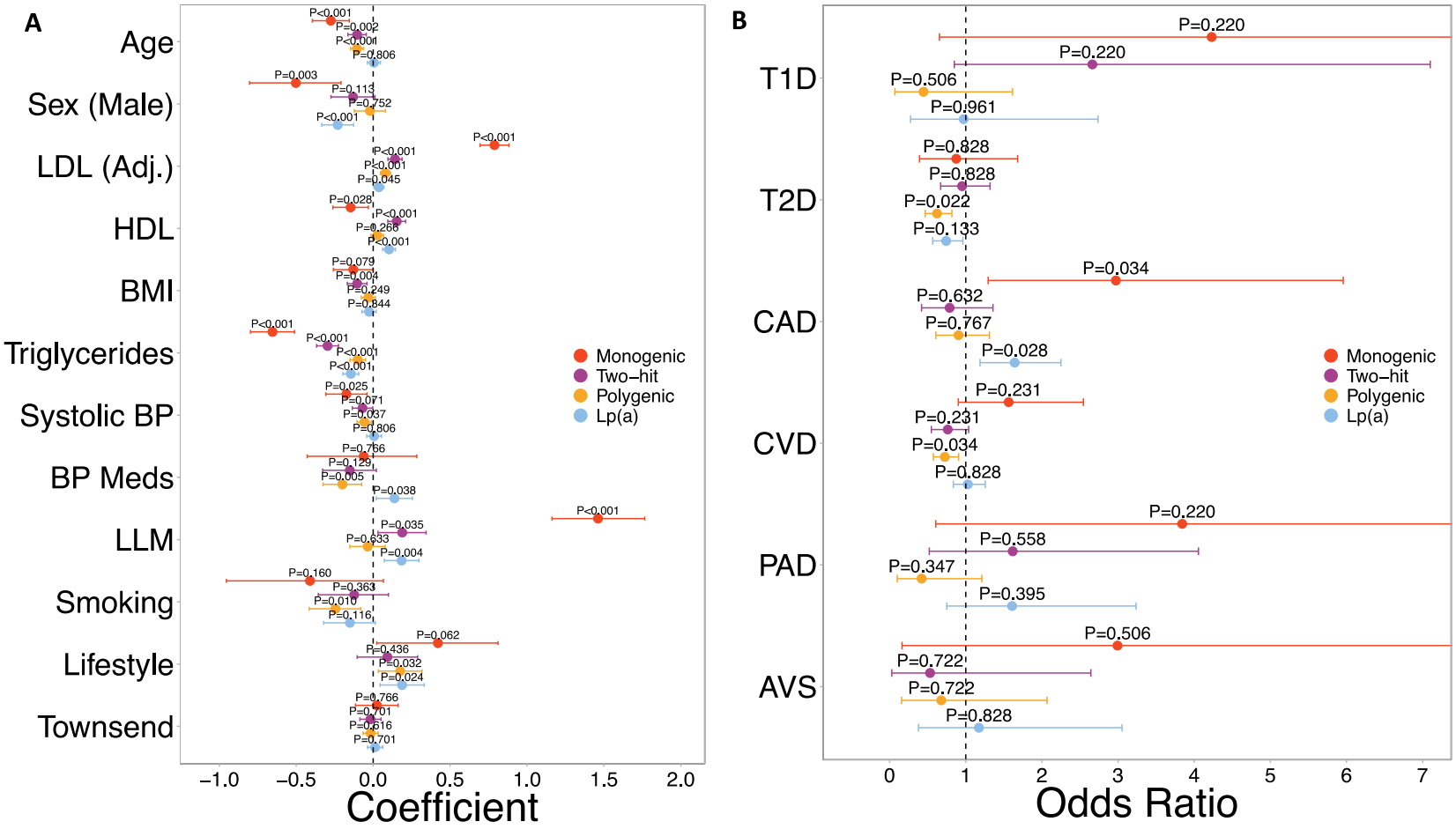

Figure S2. LLM use by FH subtype. LLM use varies by FH subtype. LLM use was stratified by the medication used and is shown for each FH subtype and controls. “Unspecified” includes self-reported LLM use without documentation for a specific LLM.

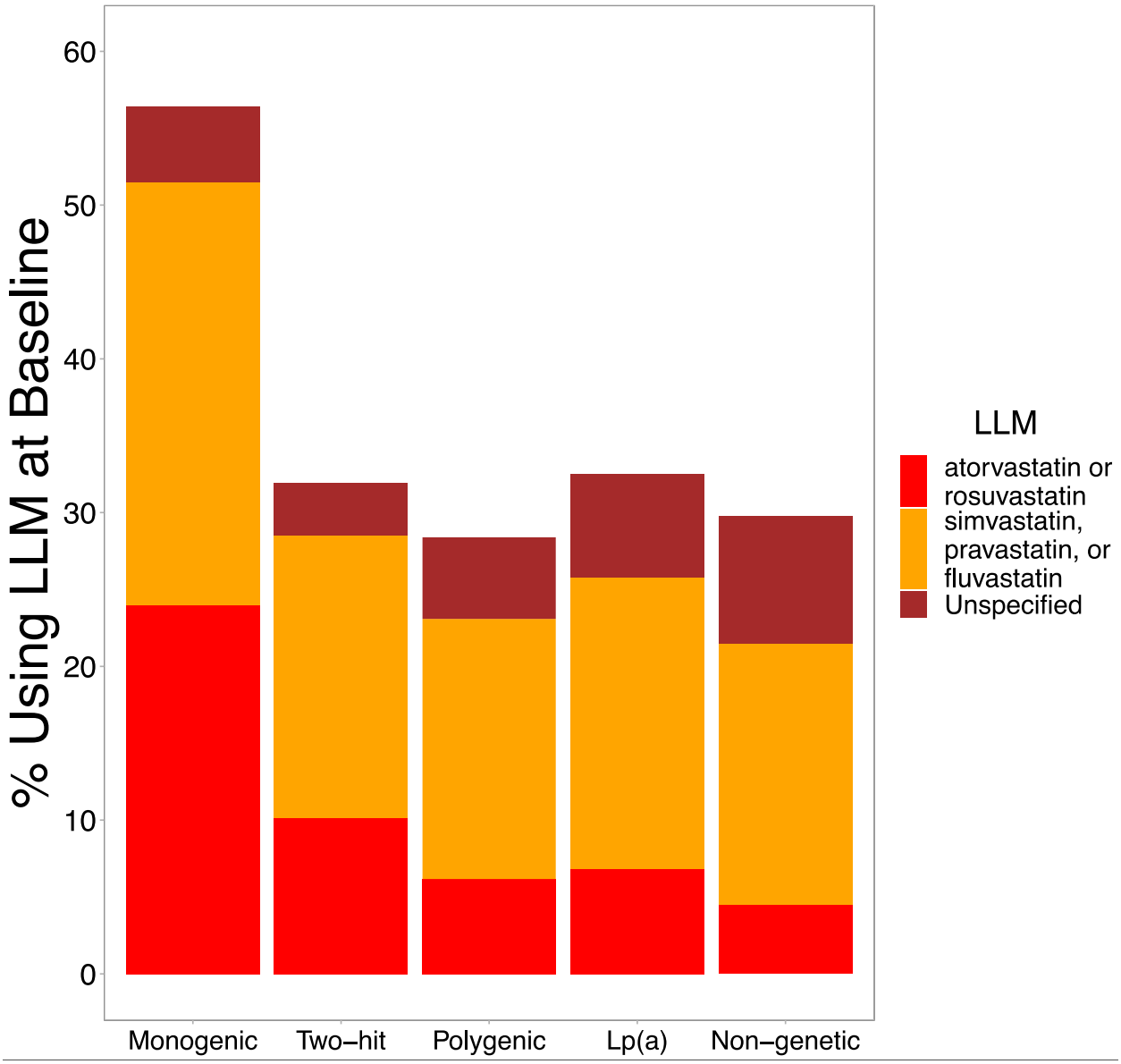

Figure S3. Family history of heart disease. Forest plots show odds ratios for the likelihood that UK Biobank participants with an FH subtype reported having a father, mother, or sibling with heart disease compared to the general population. Family history represents any combination of father, mother, and sibling.

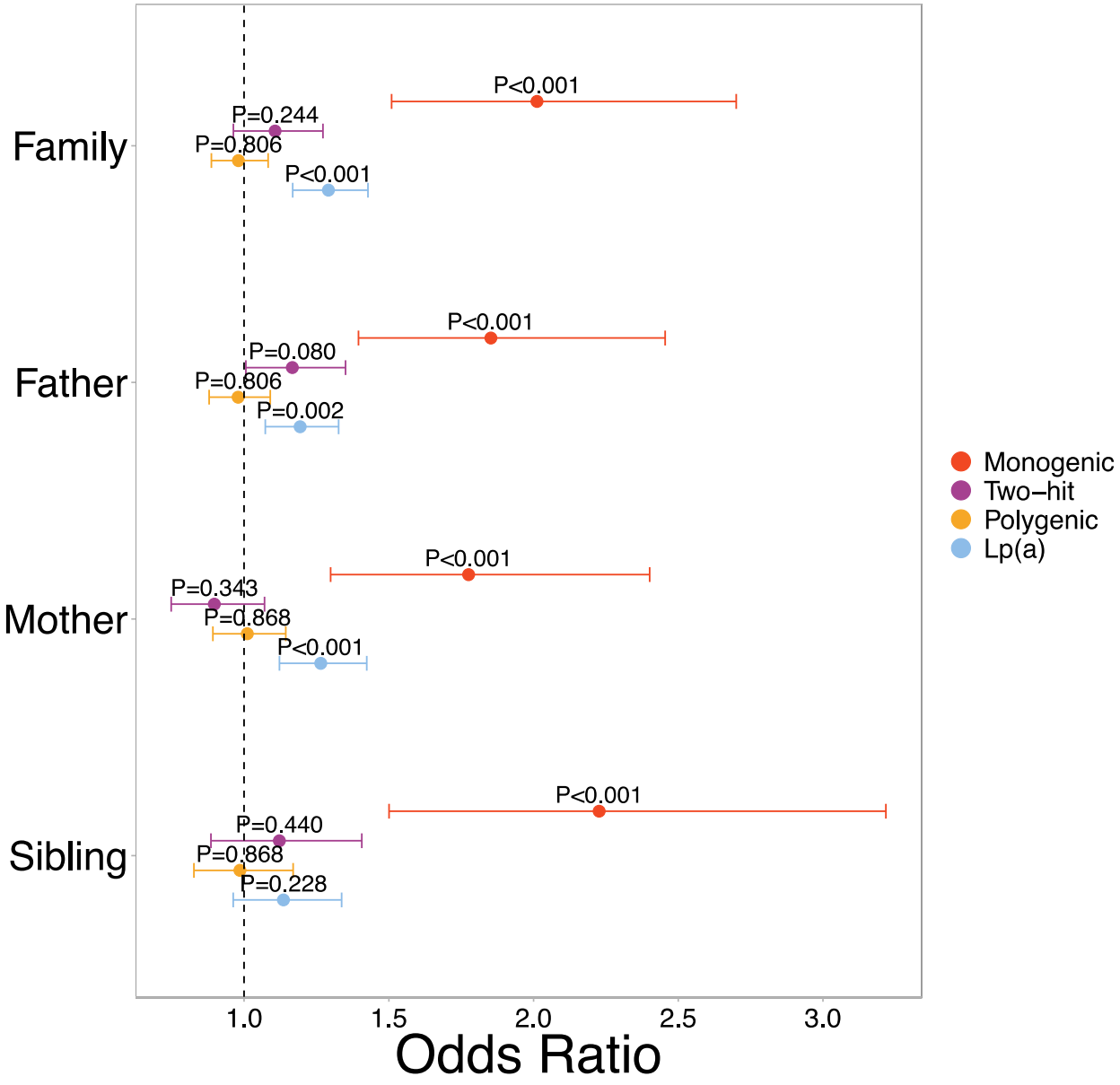

Figure S4. 10-year incident CAD risk by monogenic FH gene. Kaplan-Meier plots show 10-year incident ASCVD risk among UK Biobank participants with *PCSK9* (n=6), *APOB* (n=65), *LDLR* Moderate (n=99), and *LDLR* predicted loss-of-function (n=17) monogenic FH variants relative to the non-genetic hypercholesterolemia subtype. Cox proportional-hazard ratios and 95% confidence intervals are shown.
[1]Cox proportional hazard ratio cannot be calculated when no events occur.

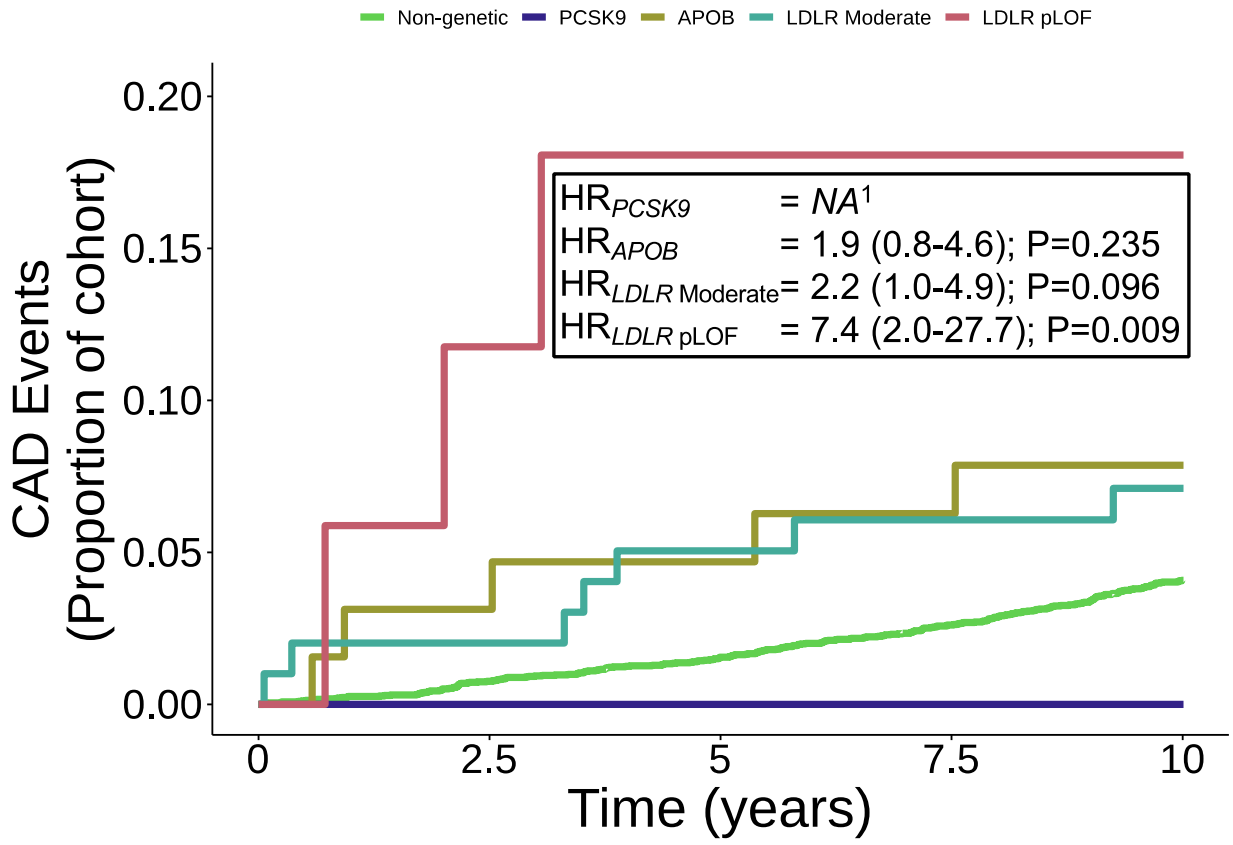

Figure S5. Severe hypercholesterolemia stratified by genetic subtype using genetically predicted Lp(a). Kaplan-Meier plots show 10-year incident ASCVD risk among UK Biobank participants with severe hypercholesterolemia. Participants were binned by severe hypercholesterolemia subtype and compared to the non-genetic hypercholesterolemia subtype. The two-hit and elevated Lp(a) subtypes were defined by having genetically predicted Lp(a) ≥ 125nmol/L. Cox proportional-hazard ratios and 95% confidence intervals adjusted for CAD risk factors and principal components of ancestry are shown.

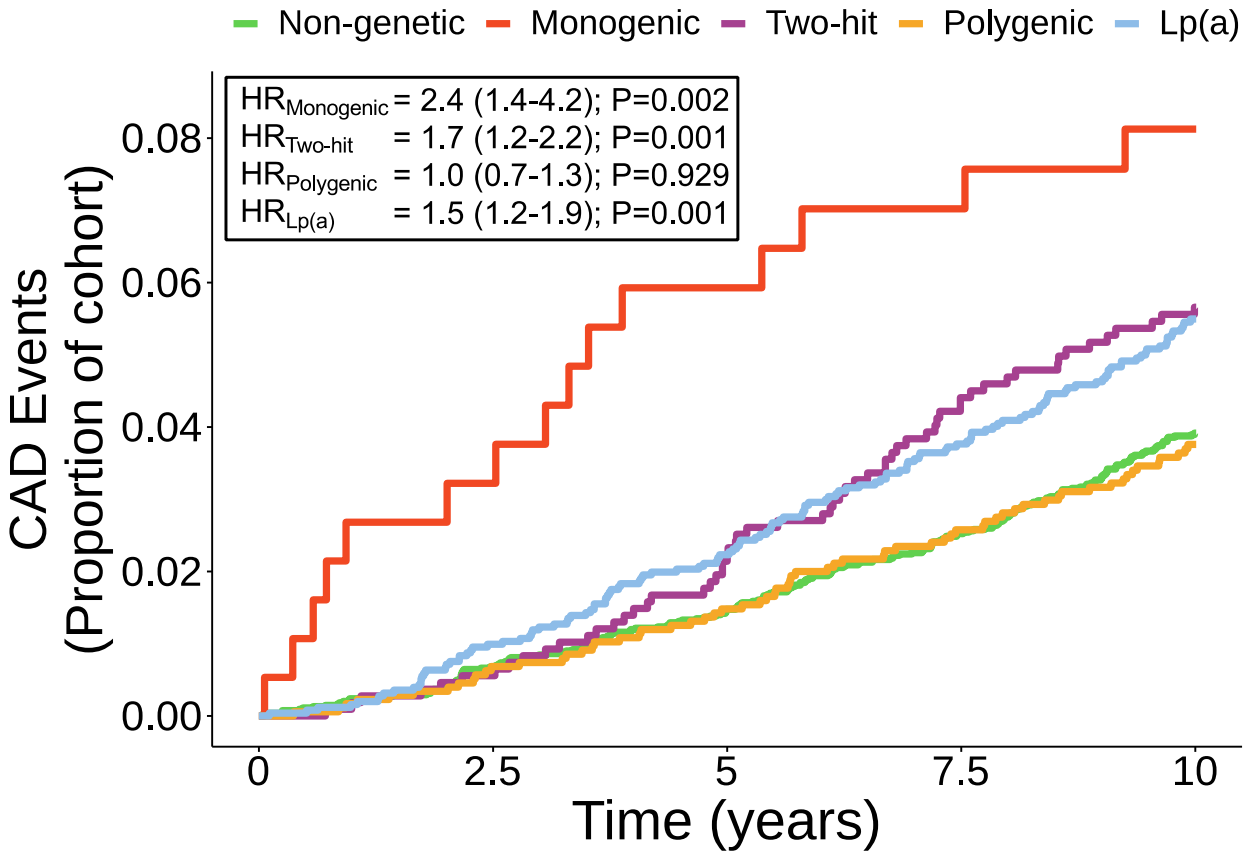

Figure S6. Severe hypercholesterolemia stratified by LDL-C and genetic subtype in males and females. Kaplan-Meier plots show 10-year incident ASCVD risk among UK Biobank participants with severe hypercholesterolemia. A) Male and C) female participants were binned by LLM-adjusted LDL-C and compared to those with LDL-C ≥190mg/dL and <210mg/dL. B) Male and D) female participants were binned by severe hypercholesterolemia subtype and compared to the non-genetic hypercholesterolemia subtype. Cox proportional-hazard ratios and 95% confidence intervals adjusted for CAD risk factors and principal components of ancestry are shown.

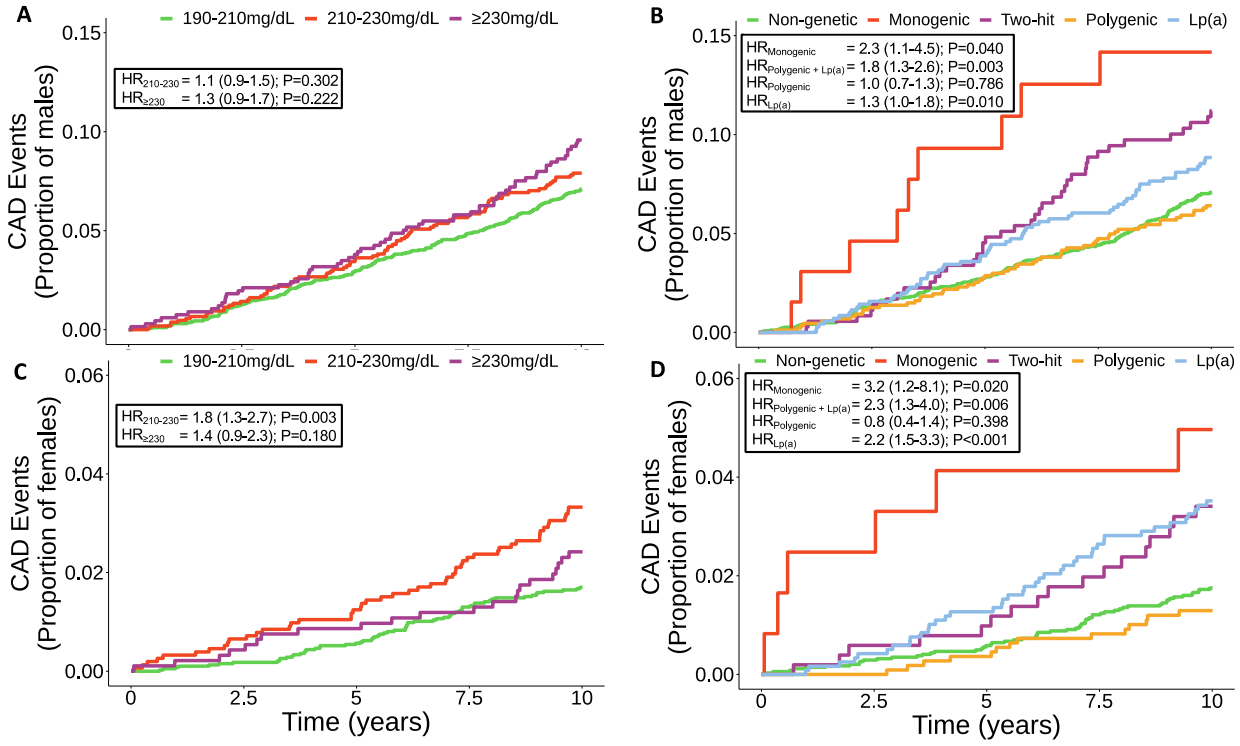

Figure S7. LLM use by monogenic FH variant. Among participants with the monogenic FH subtype and without prevalent CVD, A) the percent taking lipid-lowering medication at baseline is shown for each variant.

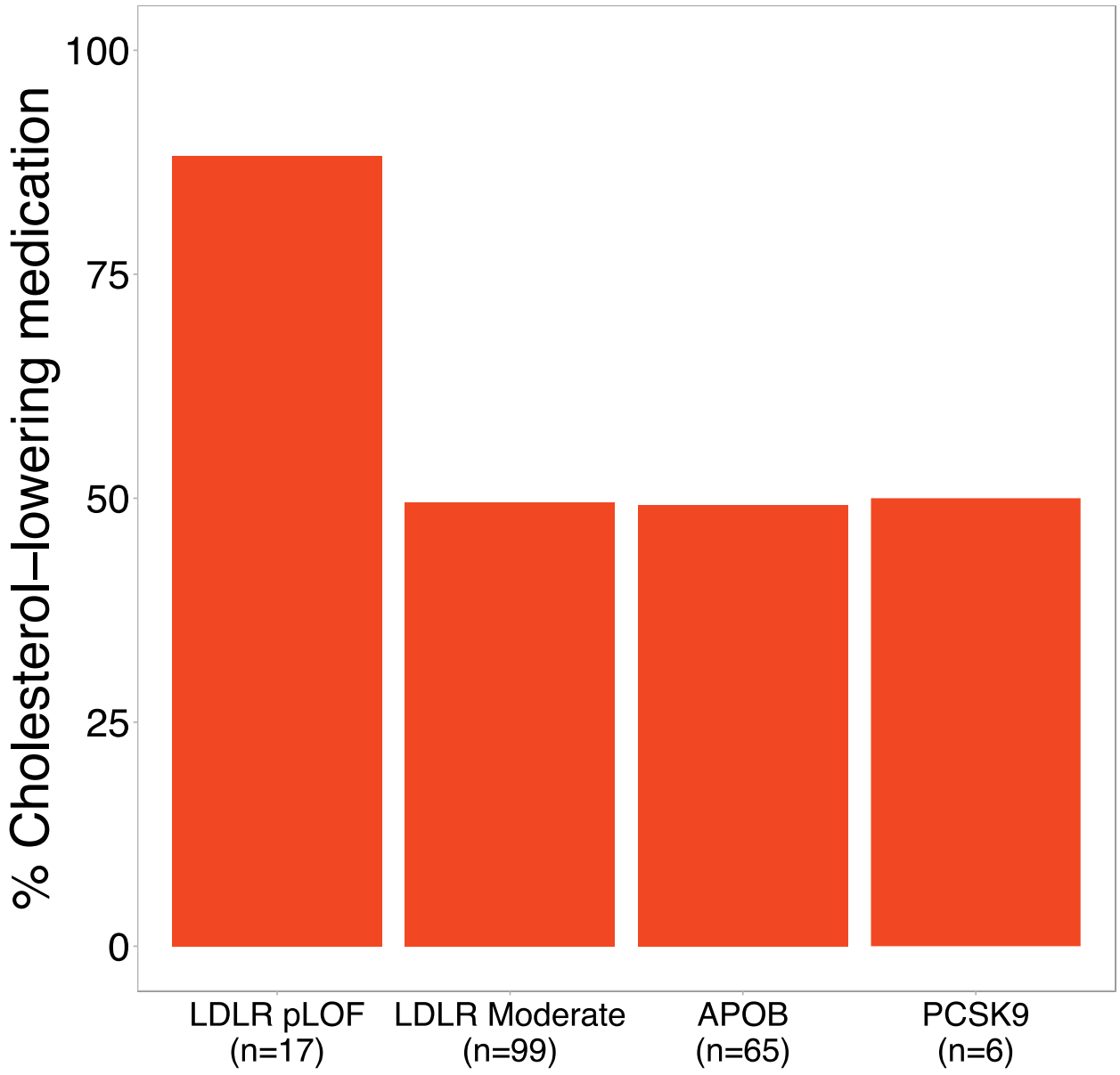

Table S1. Monogenic FH variant list.

Group	Chr	Pos (hg38)	Ref	Alt	HGVSc	Consequence
APOB	2	21006288	C	T	NM_000384.2:c.10580G>A	missense_variant
APOB	2	21006289	G	A	NM_000384.2:c.10579C>T	missense_variant
LDLR_MODERATE	19	11100291	T	G	NM_000527.4:c.136T>G	missense_variant
LDLR_MODERATE	19	11102714	C	T	NM_000527.4:c.241C>T	missense_variant
LDLR_MODERATE	19	11102732	T	G	NM_000527.4:c.259T>G	missense_variant
LDLR_MODERATE	19	11102739	G	A	NM_000527.4:c.266G>A	missense_variant
LDLR_MODERATE	19	11102774	G	A	NM_000527.4:c.301G>A	missense_variant
LDLR_MODERATE	19	11105243	G	A	NM_000527.4:c.337G>A	missense_variant
LDLR_MODERATE	19	11105324	G	A	NM_000527.4:c.418G>A	missense_variant
LDLR_MODERATE	19	11105408	G	A	NM_000527.4:c.502G>A	missense_variant
LDLR_MODERATE	19	11105556	ATGG	A	NM_000527.4:c.654_656del	inframe_deletion
LDLR_MODERATE	19	11105567	G	A	NM_000527.4:c.661G>A	missense_variant
LDLR_MODERATE	19	11105568	A	G	NM_000527.4:c.662A>G	missense_variant
LDLR_MODERATE	19	11105587	C	G	NM_000527.4:c.681C>G	missense_variant
LDLR_MODERATE	19	11105588	G	C	NM_000527.4:c.682G>C	missense_variant
LDLR_MODERATE	19	11106588	G	A	NM_000527.4:c.718G>A	missense_variant
LDLR_MODERATE	19	11106631	A	C	NM_000527.4:c.761A>C	missense_variant
LDLR_MODERATE	19	11107436	G	A	NM_000527.4:c.862G>A	missense_variant
LDLR_MODERATE	19	11107461	G	A	NM_000527.4:c.887G>A	missense_variant
LDLR_MODERATE	19	11107484	G	A	NM_000527.4:c.910G>A	missense_variant

LDLR_MODERATE	19	11107486	C	G	NM_000527.4:c.912C>G	missense_variant
LDLR_MODERATE	19	11107512	G	A	NM_000527.4:c.938G>A	missense_variant&splice_region_variant
LDLR_MODERATE	19	11110714	G	A	NM_000527.4:c.1003G>A	missense_variant
LDLR_MODERATE	19	11111514	A	G	NM_000527.4:c.1061A>G	missense_variant&splice_region_variant
LDLR_MODERATE	19	11111571	G	A	NM_000527.4:c.1118G>A	missense_variant
LDLR_MODERATE	19	11113307	C	T	NM_000527.4:c.1216C>T	missense_variant
LDLR_MODERATE	19	11113313	G	A	NM_000527.4:c.1222G>A	missense_variant
LDLR_MODERATE	19	11113322	A	G	NM_000527.4:c.1231A>G	missense_variant
LDLR_MODERATE	19	11113337	C	T	NM_000527.4:c.1246C>T	missense_variant
LDLR_MODERATE	19	11113343	G	A	NM_000527.4:c.1252G>A	missense_variant
LDLR_MODERATE	19	11113376	G	A	NM_000527.4:c.1285G>A	missense_variant
LDLR_MODERATE	19	11113590	G	T	NM_000527.4:c.1414G>T	missense_variant
LDLR_MODERATE	19	11113600	C	T	NM_000527.4:c.1424C>T	missense_variant
LDLR_MODERATE	19	11113612	T	C	NM_000527.4:c.1436T>C	missense_variant
LDLR_MODERATE	19	11113620	G	A	NM_000527.4:c.1444G>A	missense_variant
LDLR_MODERATE	19	11113650	G	A	NM_000527.4:c.1474G>A	missense_variant
LDLR_MODERATE	19	11113678	C	T	NM_000527.4:c.1502C>T	missense_variant
LDLR_MODERATE	19	11113743	G	A	NM_000527.4:c.1567G>A	missense_variant
LDLR_MODERATE	19	11116125	G	A	NM_000527.4:c.1618G>A	missense_variant
LDLR_MODERATE	19	11116141	G	A	NM_000527.4:c.1634G>A	missense_variant
LDLR_MODERATE	19	11116198	A	G	NM_000527.4:c.1691A>G	missense_variant
LDLR_MODERATE	19	11116873	C	T	NM_000527.4:c.1720C>T	missense_variant

LDLR_MODERATE	19	11116898	T	C	NM_000527.4:c.1745T>C	missense_variant
LDLR_MODERATE	19	11116900	C	T	NM_000527.4:c.1747C>T	missense_variant
LDLR_MODERATE	19	11116901	A	G	NM_000527.4:c.1748A>G	missense_variant
LDLR_MODERATE	19	11116928	G	A	NM_000527.4:c.1775G>A	missense_variant
LDLR_MODERATE	19	11116936	C	T	NM_000527.4:c.1783C>T	missense_variant
LDLR_MODERATE	19	11116976	C	G	NM_000527.4:c.1823C>G	missense_variant
LDLR_MODERATE	19	11120106	G	T	NM_000527.4:c.1860G>T	missense_variant
LDLR_MODERATE	19	11120110	G	A	NM_000527.4:c.1864G>A	missense_variant
LDLR_MODERATE	19	11120143	C	T	NM_000527.4:c.1897C>T	missense_variant
LDLR_MODERATE	19	11120144	G	A	NM_000527.4:c.1898G>A	missense_variant
LDLR_MODERATE	19	11120411	T	C	NM_000527.4:c.2029T>C	missense_variant
LDLR_MODERATE	19	11120436	C	T	NM_000527.4:c.2054C>T	missense_variant
LDLR_pLOF	19	11100272	CA	C	NM_000527.4:c.118del	frameshift_variant
LDLR_pLOF	19	11102662	A	G	NM_000527.4:c.191-2A>G	splice_acceptor_variant
LDLR_pLOF	19	11102683	CG	C	NM_000527.4:c.214del	frameshift_variant
LDLR_pLOF	19	11102785	TCG	T	NM_000527.4:c.313_313+1del	splice_donor_variant&coding_sequence_variant
LDLR_pLOF	19	11102787	G	GT	NM_000527.4:c.313+2dup	splice_donor_variant
LDLR_pLOF	19	11102787	G	A	NM_000527.4:c.313+1G>A	splice_donor_variant
LDLR_pLOF	19	11102787	G	C	NM_000527.4:c.313+1G>C	splice_donor_variant
LDLR_pLOF	19	11105339	GTGCTCACCT	G	NM_000527.4:c.435_457del	frameshift_variant

			GTGGTCCCCGCCAGC			
LDLR_pLOF	19	11105407	C	A	NM_000527.4:c.501C>A	stop_gained
LDLR_pLOF	19	11105415	AC	A	NM_000527.4:c.513del	frameshift_variant
LDLR_pLOF	19	11105470	C	G	NM_000527.4:c.564C>G	stop_gained
LDLR_pLOF	19	11105579	A	T	NM_000527.4:c.673A>T	stop_gained
LDLR_pLOF	19	11105585	GAC	G	NM_000527.4:c.680_681del	frameshift_variant
LDLR_pLOF	19	11105588	G	T	NM_000527.4:c.682G>T	stop_gained
LDLR_pLOF	19	11105589	AG	A	NM_000527.4:c.684del	frameshift_variant
LDLR_pLOF	19	11107504	CA	C	NM_000527.4:c.933del	frameshift_variant
LDLR_pLOF	19	11110759	C	T	NM_000527.4:c.1048C>T	stop_gained
LDLR_pLOF	19	11110764	ATGCGAAGG	A	NM_000527.4:c.1056_1060+3d	splice_donor_variant&coding_sequence_variant
LDLR_pLOF	19	11113268	G	A	NM_000527.4:c.1187-10G>A	intron_variant
LDLR_pLOF	19	11113295	TTC	T	NM_000527.4:c.1206_1207del	frameshift_variant
LDLR_pLOF	19	11113348	C	G	NM_000527.4:c.1257C>G	stop_gained
LDLR_pLOF	19	11113554	CA	C	NM_000527.4:c.1379del	frameshift_variant
LDLR_pLOF	19	11113559	CG	C	NM_000527.4:c.1384del	frameshift_variant
LDLR_pLOF	19	11113767	G	A	NM_000527.4:c.1586+5G>A	splice_region_variant&intron_variant

LDLR_pLOF	19	11116114	G	A	NM_000527.4:c.1607G>A	stop_gained
LDLR_pLOF	19	11116139	AG	A	NM_000527.4:c.1637del	frameshift_variant
LDLR_pLOF	19	11116951	GAGGATGA	G	NM_000527.4:c.1800_1806del	frameshift_variant
LDLR_pLOF	19	11120091	G	A	NM_000527.4:c.1846-1G>A	splice_acceptor_variant
LDLR_pLOF	19	11120110	GAT	G	NM_000527.4:c.1867_1868del	frameshift_variant
LDLR_pLOF	19	11120419	T	A	NM_000527.4:c.2037T>A	stop_gained
LDLR_pLOF	19	11120442	T	TC	NM_000527.4:c.2061dup	frameshift_variant
LDLR_pLOF	19	11123200	G	T	NM_000527.4:c.2167G>T	stop_gained
LDLR_pLOF	19	11123324	TA	T	NM_000527.4:c.2292del	frameshift_variant
LDLR_pLOF	19	11128027	C	CA	NM_000527.4:c.2332dup	frameshift_variant
PCSK9	1	55039931	G	A	NM_174936.3:c.94G>A	missense_variant
PCSK9	1	55057454	G	T	NM_174936.3:c.1120G>T	missense_variant
PCSK9	1	55058630	C	T	NM_174936.3:c.1486C>T	missense_variant

Table S2. UK Biobank codes and measurements used for identification of secondary exclusions of FH. Having any codes in this table was considered a secondary cause of FH.

Diagnosis	EHR-documentation		Self-reported		
	ICD-9	ICD-10	Non-cancer illness code (20002)	Treatment/ Medication Code (20003)	UK Biobank field ID
Hypothyroidism	243	E03	1226	1141191044	
	244			1140884516	
				1140874852	
				1140910814	
				1141178036	
				1140884512	
				1140910520	
				1140910518	
Obstructive Liver Disease	5762	K83.1	1159		
			1160		
Nephrotic Syndrome	581	N04	1609		
Alkaline Phosphatase > 200 U/L					30610
Triglycerides > 400 mg/dL					30870

Table S3. UK Biobank codes used for identification of lipid-lowering medication use.

	Medication for cholesterol (6153)	Medication for cholesterol (6177)	Treatment/ Medication Code (20003)
Any LLM	1	1	All codes below
Simvastatin			1140861958
			1140910652
			1140881748
			1141200040
			1141188146
			1140910654
Pravastatin			1140888648
			1140910632
			1140861970
Fluvastatin			1140888594
			1140864592
Atorvastatin			1141146234
			1141146138
Rosuvastatin			1141192410
			1141187780
			1141192414

Table S4. LDL Polygenic score correlation with adjusted and unadjusted LDL-C. Pearson correlation coefficients are shown for the LLM-untreated cohort and the total cohort when using LDL-C without adjustment for LLM and LDL-C adjusted for LLM.

Cohort	LDL	r	t	df	p
Untreated	Unadjusted	0.41	145.9	108,258	<0.001
Total	Unadjusted	0.31	115.6	130,089	<0.001
Total	Adjusted	0.39	151.8	130,089	<0.001

Table S5. Monogenic variant correlation with adjusted and unadjusted LDL-C. Coefficients (with 95% confidence intervals) and p-values for the generalized linear models are shown for the LLM-untreated cohort and the total cohort when using LDL-C without adjustment for LLM and LDL-C adjusted for LLM.

Cohort	LDL	Estimate (95% CI)	p-value
Untreated	Unadjusted	59.4 (54.6 – 64.1)	4.4E-133
Total	Unadjusted	31.6 (28.1 – 35.0)	1.9E-72
Total	Adjusted	58.3 (55.0 – 61.7)	1.3E-254

Table S6. Determinant and subtype definitions.

Determinant of severe hypercholesterolemia	Determinant N	Determinant definition	Subtype	Subtype N[1]
FH Variant	357	P/LP variant in an FH gene	Monogenic FH	204
LDL-C PRS top decile	9,657	PRS in the top decile, Lp(a) < 125nmol/L, no FH variant	Polygenic hypercholesterolemia	2,070
Lp(a)≥125nmol/L	17,940	Lp(a) ≥ 125nmol/L, PRS below the top decile, no FH variant	Elevated Lp(a) hypercholesterolemia	2,023
LDL-C PRS top decile & Lp(a)≥125nmol/L	3,325	Lp(a) ≥ 125nmol/L, PRS in the top decile, no FH variant	Two-hit hypercholesterolemia	913
Control	98,812	Lp(a) < 125nmol/L, PRS below the top decile, no FH variant	Non-genetic hypercholesterolemia	6,528

[1]Subtype N was calculated before excluding prevalent CVD.

Table S7. UK Biobank codes used for identification of CAD and comorbidities

Diagnosis	EHR-documentation			Self-reported		
	ICD-9	ICD-10	OPCS-4	Non-cancer illness code (20002)	Operation code (20004)	Vascular/heart problems diagnosed by doctor (6150)
CAD	410	I21	K40.1-4	1075: Heart attack	1070: Coronary angioplasty	1: Heart attack
	411	I22	K41.1-4		1095: Coronary artery bypass graft	
	412	I23	K45.1-5			
	429.79	I24.1	K49.1-2			
		I25.2	K49.8-9			
			K50.2			
			K75.1-4			
			K75.8-9			
CVD	410	G45	K40	1074: Angina	1070: Coronary angioplasty	1: Heart attack
	411	I20	K41	1075: Heart attack	1071: Other arterial surgery/revascularization procedures	2: Angina
	412	I21	K42	1082: Transient ischaemic attack	1095: Coronary artery bypass graft	3: Stroke
	413	I22	K43	1583: Ischaemic stroke	1109: Carotid artery angioplasty	
	414	I23	K44		1514: Coronary angiogram	
	434	I24	K45			
	436	I25	K46			
		I63	K47.1			
		I64	K49			
			K50			

			K75			
PAD	440	I70	L50			
	443.8-9	I73.8-9	L51			
	444	I74	L52			
			L54			
			L58			
			L59			
			L60			
			L63			
			X09			
AVS		I350				
		I352				

Table S8. Familial hypercholesterolemia variant counts.

Variant	Sample Size (n)
LDLR pLOF	35
LDLR missense	191
APOB missense	114
PCSK9 missense	17

Table S9. Prevalence in the cohort, odds ratios of having severe hypercholesterolemia, population attributable fraction of severe hypercholesterolemia, and attributable risk proportion of severe hypercholesterolemia for each genetic determinant of severe hypercholesterolemia.

Determinant	Prevalence (%)	OR (95% CI)	PAF (%) (95% CI)	ARP (%) (95% CI)
FH Variant	0.3	13.9 (11.2-17.2)	1.5 (1.3-1.7)	84.4 (83.0-85.9)
PRS top decile	7.4	3.2 (3.0-3.4)	11.0 (10.3-11.7)	62.5 (61.0-64.1)
Lp(a) ≥ 125nmol/L	13.8	1.3 (1.2-1.4)	4.0 (3.2-4.7)	23.2 (19.7-26.7)
PRS + Lp(a)	2.6	4.1 (3.7-4.4)	5.4 (4.9-5.8)	68.9 (67.1-70.7)

All values are relative to the population after removing those with the determinant-of-interest, as opposed to relative to controls without a determinant.
Abbreviations: OR = Odds Ratio; PAF = Population Attributable Fraction; ARP = Attributable Risk Proportion

Table S10. Absolute 10-year incident CAD risk.

Comparison	Absolute risk difference (%) (95% CI)
LDL 210-230 vs LDL 190-210	1.2 (0.4-2.5)
LDL ≥230 vs LDL 190-210	1.2 (0.2-2.6)
Monogenic vs non-genetic	4.8 (0.8-8.1)
Two-hit vs non-genetic	3.3 (1.5-5.5)
Elevated Lp(a) vs non-genetic	2.0 (1.0-3.4)
Monogenic vs non-genetic treated	9.2 (0.7-18.3)
Two-hit vs non-genetic treated	3.7 (0.3-7.2)
Monogenic vs non-genetic untreated	1.6 (-2.8-6.4)
Two-hit vs non-genetic untreated	3.1 (0.9-4.8)
Elevated Lp(a) vs non-genetic untreated	2.4 (1.2-3.7)

Table S11. 10-year incident ASCVD risk re-defining polygenic FH. Cox proportional-hazard ratios for 10-year incident ASCVD risk in severe hypercholesterolemic UK Biobank participants with each FH subtype relative to the non-genetic subtype are shown among all participants. The first model is the result of the main analysis. The second model limits the non-genetic subtype to participants with LDL-C PRS below the 50th percentile. The third model limits the non-genetic subtype to participants with LDL-C PRS below the 50th percentile, and re-defines the two-hit and polygenic subtypes to include participants with LDL-C PRS above the 75th percentile with or without elevated Lp(a), respectively. 95% confidence intervals are shown in parentheses.

	FH subtype vs non-genetic	**HR (95% CI)**	**P-value**
Original	Monogenic	2.3 (1.4-4.0)	2.7 x 10^{-3}
	Two-hit	1.9 (1.4-2.6)	6.5 x 10^{-5}
	Polygenic	0.9 (0.7-1.2)	0.51
	Lp(a)	1.5 (1.2-2.0)	5.1 x 10^{-4}
• **Non-genetic includes only those with LDL-C PRS < 50th percentile**	Monogenic	2.3 (1.3-4.1)	1.2 x 10^{-2}
	Two-hit	1.8 (1.3-2.6)	3.5 x 10^{-3}
	Polygenic	0.9 (0.6-1.2)	0.43
	Lp(a)	1.5(1.1-2.0)	1.6 x 10^{-2}
• **Non-genetic includes only those with LDL-C PRS < 50th percentile** • **Two-hit and Polygenic include LDL-C PRS > 75th percentile**	Monogenic	2.4 (1.3-4.2)	7.1 x 10^{-3}
	Two-hit	1.6 (1.2-2.2)	7.1 x 10^{-3}
	Polygenic	0.9 (0.7-1.2)	0.68
	Lp(a)	1.6 (1.1-2.2)	7.1 x 10^{-3}

Table S12. Clinical characteristics and comorbidities of LLM-treated vs untreated. Differences in baseline clinical risk factors and prevalence of comorbidities in those taking LLM at baseline were compared to those untreated with LLM. Standard deviations are shown for continuous traits and odds ratios for binary traits with 95% confidence intervals. (T1D=type 1 diabetes; T2D=type 2 diabetes; CAD=coronary artery disease; CVD=cardiovascular disease; PAD=peripheral artery disease; AVS=aortic valve stenosis).

	Treated	Untreated	SD/OR (95% CI)	Adj. P value
Sample Size (N)	3594	8144	NA	NA
Age (Mean (SD))	60.54 (6.28)	57.83 (7.17)	0.41 (0.38-0.44)	P < 0.001
Sex (% Female)	44	62.3	2.62 (2.41-2.85)	P < 0.001
LDL (mg/dL) ((Mean (SD))	154.59 (21.78)	207.6 (16.77)	-1.57 (-1.6--1.55)	P < 0.001
LDL (est. untreated) (mg/dL) ((Mean (SD))	220.84 (31.11)	207.6 (16.77)	0.43 (0.4-0.45)	P < 0.001
BMI (Mean (SD))	29.11 (4.65)	27.58 (4.08)	0.37 (0.33-0.41)	P < 0.001
Systolic BP (mmHg) (Mean (SD))	144.44 (18.22)	141.69 (18.61)	0 (-0.04-0.03)	0.804
HDL (mg/dL) (Mean (SD))	53.72 (12.72)	59.43 (13.1)	-0.32 (-0.35--0.28)	P < 0.001
Triglycerides (mg/dL) (Mean (SD))	191.81 (77.63)	187.47 (73.12)	0.03 (-0.01-0.07)	0.143
AHA Lifestyle Score[1] (Mean (SD))	0.96 (0.44)	1.03 (0.43)	-0.47 (-0.59--0.36)	P < 0.001

Townsend Deprivation Index[2] (Mean (SD))	-1.28 (3.02)	-1.71 (2.82)	0.19 (0.15-0.23)	P < 0.001
BP meds (%)	51.3	10.6	8.37 (7.57-9.26)	P < 0.001
Smoking (%)	12.3	9.9	1.45 (1.27-1.65)	P < 0.001
Prevalent T1D (%)	23/3594 (0.6)	4/8144 (0)	19.27 (6.53-56.87)	P < 0.001
Prevalent T2D (%)	372/3594 (10.4)	116/8144 (1.4)	8.62 (6.92-10.74)	P < 0.001
Prevalent CAD (%)	6.5	0.1	38 (20.05-72)	P < 0.001
Prevalent CVD (%)	16.7	1.6	10.4 (8.51-12.7)	P < 0.001
Prevalent PAD (%)	1.1	0.1	9.58 (4.18-21.97)	P < 0.001
Prevalent AVS (%)	0.6	0	8.54 (2.84-25.71)	P < 0.001

[1]AHA Lifestyle Score: A higher score indicates a more ideal lifestyle (0 = poor, 1 = intermediate, 2 = ideal).
[2]Townsend Deprivation Index: A higher (less negative) score corresponds to a lower socioeconomic status.

Table S13. 10-year incident CAD risk in treated vs. untreated UK Biobank participants with severe hypercholesterolemia. Cox proportional-hazard ratios for 10-year incident CAD risk in severe hypercholesterolemic UK Biobank participants treated with LLM at baseline compared to severe hypercholesterolemic participants untreated at baseline are shown. The first model is adjusted for principal components of ancestry only. The second model is adjusted for all clinical risk differences that were significantly different between LLM-treated and untreated from Table S11, except for unadjusted LDL-C. 95% confidence intervals are shown in parentheses.

	HR (95% CI)	**P-value**
Controlling for PCs only	1.6 (1.3-1.9)	7.3×10^{-7}
Full model controlling for CAD risk factors	0.7 (0.6-0.9)	1.8×10^{-2}

Table S14. 10-year incident ASCVD risk controlling for body mass index and C reactive protein. Cox proportional-hazard ratios for 10-year incident ASCVD risk in severe hypercholesterolemic UK Biobank participants with each FH subtype relative to the non-genetic subtype are shown among all participants, participants treated with LLM at baseline, and participants untreated with LLM at baseline. Every model is adjusted for all CAD risk factors from the primary analysis in addition to BMI and C-reactive protein. The model of all participants adjusts for baseline LLM use. 95% confidence intervals are shown in parentheses.

	FH subtype vs non-genetic	HR (95% CI)	Adj. P-value
All participants	Monogenic	2.4 (1.4-4.1)	2.1×10^{-3}
	Two-hit	1.9 (1.4-2.6)	4.8×10^{-5}
	Polygenic	0.9 (0.7-1.2)	0.52
	Lp(a)	1.5 (1.2-2.0)	5.2×10^{-4}
Treated	Monogenic	3.1 (1.6-6.0)	3.0×10^{-3}
	Two-hit	1.8 (1.1-3.0)	0.032
	Polygenic	0.9 (0.6-1.4)	0.59
	Lp(a)	1.2 (0.8-1.8)	0.59
Untreated	Monogenic	1.6 (0.6-4.4)	0.49
	Two-hit	2.0 (1.4-2.8)	7.3×10^{-4}
	Polygenic	0.9 (0.6-1.3)	0.54
	Lp(a)	1.7 (1.3-2.3)	5.7×10^{-4}