Circulation: Genomic and Precision Medicine

ORIGINAL ARTICLE

Coronary Artery Disease Risk of Familial Hypercholesterolemia Genetic Variants Independent of Clinically Observed Longitudinal Cholesterol Exposure

Shoa L. Clarke, MD, PhD; Catherine Tcheandjieu, DVM, PhD; Austin T. Hilliard, PhD; Kyung Min Lee, PhD; Julie Lynch, PhD, RN; Kyong-Mi Chang, MD; Donald Miller, ScD; Joshua W. Knowles, MD, PhD; Christopher O'Donnell, MD; Philip S. Tsao, PhD; Daniel J. Rader, MD; Peter W. Wilson, MD; Yan V. Sun, PhD; J. Michael Gaziano, MD, MPH; Themistocles L. Assimes, MD, PhD; VA Million Veteran Program

BACKGROUND: Familial hypercholesterolemia (FH) genetic variants confer risk for coronary artery disease independent of LDL-C (low-density lipoprotein cholesterol) when considering a single measurement. In real clinical settings, longitudinal LDL-C data are often available through the electronic health record. It is unknown whether genetic testing for FH variants provides additional risk-stratifying information once longitudinal LDL-C is considered.

METHODS: We used the extensive electronic health record data available through the Million Veteran Program to conduct a nested case-control study. The primary outcome was coronary artery disease, derived from electronic health record codes for acute myocardial infarction and coronary revascularization. Incidence density sampling was used to match case/control exposure windows, defined by the date of the first LDL-C measurement to the date of the first coronary artery disease code of the index case. Adjustments for the first, maximum, or mean LDL-C were analyzed. FH variants in *LDLR*, *APOB*, and *PCSK9* (Proprotein convertase subtilisin/kexin type 9) were assessed by custom genotype array.

RESULTS: In a cohort of 23 091 predominantly prevalent cases at enrollment and 230 910 matched controls, FH variant carriers had an increased risk for coronary artery disease (odds ratio [OR], 1.53 [95% CI, 1.24–1.89]). Adjusting for mean LDL-C led to the greatest attenuation of the risk estimate, but significant risk remained (odds ratio, 1.33 [95% CI, 1.08–1.64]). The degree of attenuation was not affected by the number and the spread of LDL-C measures available.

CONCLUSIONS: The risk associated with carrying an FH variant cannot be fully captured by the LDL-C data available in the electronic health record, even when considering multiple LDL-C measurements spanning more than a decade.

Key Words: cholesterol ■ coronary artery disease ■ electronic health record ■ hypercholesterolemia ■ risk

Familial hypercholesterolemia (FH) is a monogenic disorder that causes elevated LDL-C (low-density lipoprotein cholesterol) from birth, leading to increased risk for cardiovascular disease. Early identification and treatment of individuals with FH may significantly improve outcomes.[1,2] However, FH is underdiagnosed and undertreated.[3] Current practice relies on family history, physical exam, and cholesterol screening to identify FH, but many FH variant carriers do not meet criteria for the clinical diagnosis of FH.[4]

Correspondence to: Shoa L. Clarke, MD, PhD, 870 Quarry Rd, Falk Building, Stanford, CA 94306, Email shoa@stanford.edu or Themistocles L. Assimes, MD, PhD, 1070 Arastradero Rd, Suite 300, Palo Alto, CA 94304, Email tassimes@stanford.edu

This article was sent to Ruth McPherson, MD, PhD, Guest Editor, for review by expert referees, editorial decision, and final disposition.
Supplemental Material is available at: https://www.ahajournals.org/doi/suppl/10.1161/CIRCGEN.121.003501.
For Sources of Funding and Disclosures, see page 146.

Nonstandard Abbreviations and Acronyms

CAD	coronary artery disease
EHR	electronic health record
FH	familial hypercholesterolemia
LDL-C	low-density lipoprotein cholesterol
MVP	Million Veteran Program

Prior studies suggest that carrying an FH variant confers independent risk for coronary artery disease (CAD) after adjustment for a single baseline LDL-C measurement.[5,6] These observations have supported efforts to increase clinical genetic testing for FH.[7] However, clinicians often have access to multiple historical LDL-C measurements documented in the medical record. It is unknown whether FH variants continue to confer independent risk after accounting for longitudinal LDL-C exposure.

Estimating the risk among FH variant carriers while accounting for multiple LDL-C measurements over many years is challenging given the relatively small size of most observational cohort studies. However, the maturation of biobanks within large-scale integrated healthcare systems with extensive electronic health records (EHR) provides unprecedented opportunities. We analyzed linked genetic and EHR-derived data for >400 000 participants in the Million Veteran Program (MVP)[8] to test the hypothesis that clinically measured longitudinal LDL-C exposure can account for the CAD risk associated with carrying an FH variant.

METHODS

The Veterans Affairs (VA) Institutional Review Board approved the MVP study protocol in accordance with the principles outlined in the Declaration of Helsinki. Informed consent was obtained from all participants. The individual-level data of veteran participants are only available upon approval from the United States Department of VA Institutional Review Board.

Full methods are now available in the Supplemental Material.

RESULTS

FH variant carriers in the MVP population

We identified 55 FH variants (51 *LDLR*, 2 *APOB*, 2 *PCSK9* [Proprotein convertase subtilisin/kexin type 9]) among 455 734 MVP participants (Table S1 in the Supplemental Material). FH variants were defined by (1) ClinVar annotations of *LDLR*, *APOB*, and *PCSK9*; (2) predicted loss-of-function variants in *LDLR*; and (3) predicted pathogenic missense variants in *LDLR*. Additionally, we assessed two missense variants in *APOB* that were previously found to be associated with severe hypercholesterolemia in MVP[9] but were labeled as "uncertain" or "conflicting evidence" in ClinVar. We found that one of these variants was strongly associated with CAD (Table S2 in the Supplemental Material), and thus we chose to keep it in our analysis as an FH variant. All identified FH variants were directly genotyped. In total, we found 1504 carriers of these variants, for an approximate prevalence of 1 in 303 (Table 1). After excluding individuals with missing demographic data and filtering for relatedness, we were left with 435 946 unrelated individuals, including 1497 FH carriers (Figure 1).

LDL-C Metrics and Association With FH Carrier Status

The majority of participants (418 790 or 96.1%) had at least one LDL-C measurement in the EHR, and the median number of LDL-C measurements per individual was 12 (interquartile range, 6–21). In total, ≈6.3 million LDL-C measurements were used in this study. MVP participants carrying FH variants showed a wide range of LDL-C values (Figure 2A). The prevalence of FH variant carriers among subjects with severe hypercholesterolemia (LDL-C>190 mg/dL) varied dramatically depending on which LDL-C metric was used to define severe hypercholesterolemia (Table 2). In general, however, LDL-C metrics offered only modest discriminatory power for predicting FH carrier status, with mean LDL-C performing better than the other metrics (Figure 2B).

Table 1. Prevalence of FH Variant Carriers in the Million Veteran Program

	Ancestry group					
	All	African	Asian	European	Hispanic	Unclassified
n	455 734	87 163	4553	318 694	34 151	11 173
FH variant carriers	1504	258	11	1095*	111	29
LDLR LoF	165	20	3	130	10	2
LDLR missense	944	222	6	606	91	19
APOB	383	16	2	349	8	8
PCSK9	13	0	0	11	2	0
Prevalence (95% CI)	1:303 (1:288–319)	1:338 (1:301–385)	1:414 (1:260–1010)	1:291 (1:275–309)	1:308 (1:259–378)	1:385 (1:283–605)

FH indicates familial hypercholesterolemia; LoF, loss of function; *PCSK9*, Proprotein convertase subtilisin/kexin type 9.
*One individual was found to be a carrier of both an *LDLR* missense variant and an *APOB* variant.

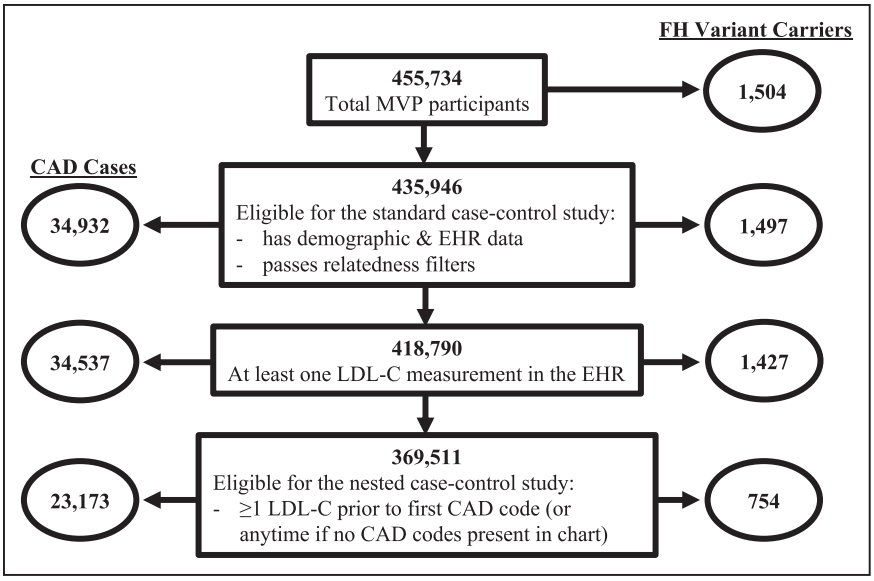

Figure 1. Summary of the study cohort at each stage of analysis. CAD indicates coronary artery disease; EHR, electronic health record; FH, familial hypercholesterolemia; LDL-C, low-density lipoprotein cholesterol; and MVP, Million Veteran Program.

FH Genetic Variants, LDL-C Exposure, and Risk for CAD

We first conducted a standard case-control study of CAD to provide comparison to prior sequencing-based population studies of FH variant carriers.[4,6] We identified 34 932 CAD cases. A majority of cases (29 300; 84%) were prevalent at the time of enrollment with a mean time from first CAD code to enrollment of 7.6±4.9 years. For incident cases, the mean time from enrollment to the date of the first CAD code was 2.0±1.5 years. We compared cases to 291 408 controls defined as having no codes suggestive of CAD documented across the full span of EHR data. All traditional risk factors were more prevalent among cases compared with controls (Table S3 in the Supplemental Material). The odds ratio (OR) for CAD among FH carriers was 1.7 (95% CI, 1.4–2.0). The OR for premature CAD (male <55 and female <65) was 3.0 (95% CI, 1.7–5.0), consistent with other population studies[4,6] (Figure S1 in the Supplemental Material). When adjusting for LDL-C using the first available measurement, the risk attenuated but remained significant for all CAD (OR, 1.4 [95% CI, 1.2–1.6]) and for premature CAD (OR, 2.1 [95% CI, 1.2–3.7]).

We next conducted a nested case-control study[10] designed to measure the risk of CAD while adjusting for longitudinal LDL-C exposure. Cases were restricted to those with ≥1 LDL-C measurement before the first

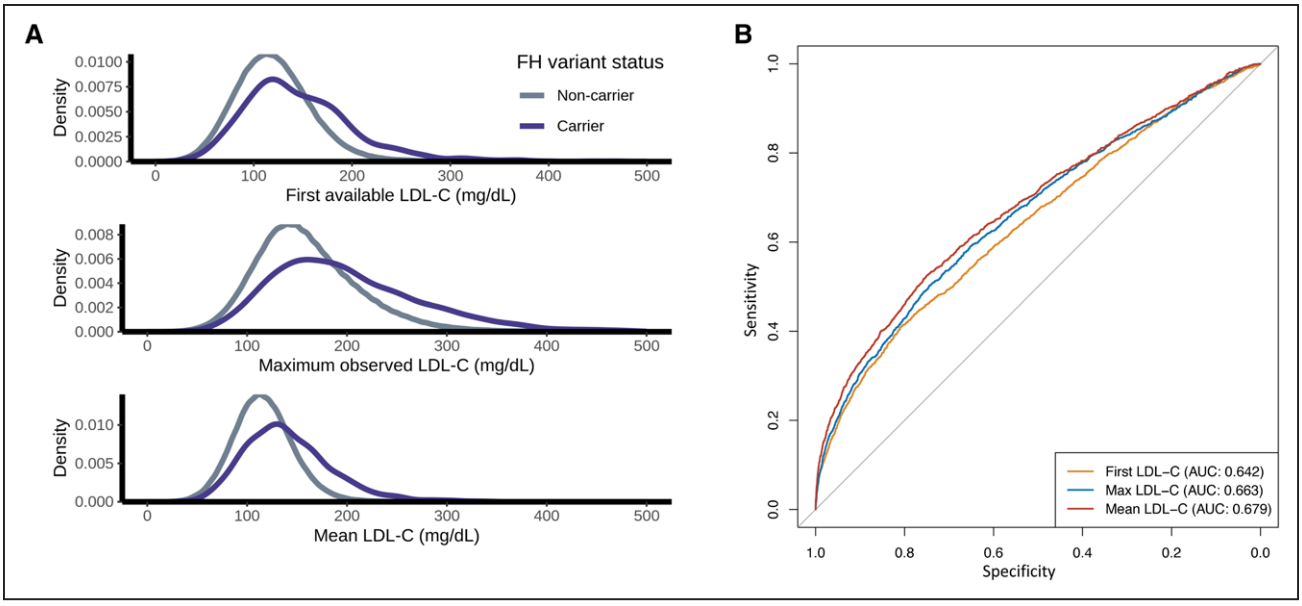

Figure 2. Familial hypercholesterolemia (FH) variant carrier status and LDL-C (low-density lipoprotein cholesterol) metrics. **A**, Density distributions of the first, maximum, and mean LDL-C measurements observed in the electronic health record for individuals with and without FH genetic variants. **B**, Receiver operating characteristic curve for predicting FH variant carrier status using each LDL-C metric with adjustment for age at measurement. For mean LDL-C, the age at each measurement was used to calculate a mean age across all measurements. To convert LDL-C values from mg/dL to mmol/L, divide by 38.67. AUC indicates area under the curve.

Table 2. Prevalence of FH Variant Carriers by LDL-C Level, Defined by the First Available, the Maximum Observed, or the Mean of All Measures

LDL-C, mg/dL	n	FH variant carriers, %
First		
≤130	264 734	640 (0.2)
131–190	135 800	535 (0.4)
>190	18 256	252 (1.4)
>250	1816	70 (3.9)
Maximum		
≤130	124 964	236 (0.2)
131–190	191 581	500 (0.3)
>190	102 245	691 (0.7)
>250	24 089	321 (1.3)
Mean		
≤130	293 585	642 (0.2)
131–190	119 689	607 (0.5)
>190	5516	178 (3.2)
>250	244	28 (11.5)

To convert LDL-C values from mg/dL to mmol/L, divide by 38.67. FH indicates familial hypercholesterolemia; and LDL-C, low-density lipoprotein cholesterol.

diagnosis of CAD (n=23 173). The median number of prior measurements was 6 (interquartile range, 2–12), and the median span of prior measurements was 49 months (interquartile range, 12–100). Both FH variant carriers and noncarriers had similarly extensive prior LDL-C data (Figure S2 in the Supplemental Material). For each case, we matched 10 random controls, matching on date of first LDL-C, sex, year of birth, and ancestry. We used the principle of incidence density sampling to allow measurement of LDL-C exposure over matched etiologic exposure windows for all subjects in a given set (Figure 3). Three LDL-C metrics over the exposure window were considered: first (earliest available measurement), max (highest observed measurement during the exposure window), and mean (average of all LDL-C observed during the exposure window). In total, 23 091 cases (99.6%) were successfully matched to 10 controls (Table 3). The OR for CAD among FH variant carriers was 1.53 (95% CI, 1.24–1.89). When adding an adjustment for the first, the maximum observed, or the mean LDL-C before the index date, the OR progressively attenuated, but the risk among FH variant carriers remained significant (Figure 4, Table S4 in the Supplemental Material). We observed the same pattern of incomplete attenuation when analyzing the subset of matched sets restricted to incident cases occurring after enrollment (Figure S3 in the Supplemental Material). In an additional sensitivity analysis, we assessed the impact of using alternative approaches to statin correction (see Methods in the Supplemental Material). Our results were robust across each approach, which included no statin correction, a more aggressive statin correction, a less aggressive statin correction, and a variable statin correction based on LDL-C level (Table S5 in the Supplemental Material).

We next tested for modifiers of the CAD risk associated with carrying an FH variant. We found a significant interaction between sex and carrier status (P=0.03). The interaction remained significant with adjustments for LDL-C (Figure 4). Stratified analyses showed an OR for CAD of 3.65 (CI, 1.51–8.84) among female FH variant carriers and 1.46 (CI, 1.17–1.82) among male carriers (Figure 4). Importantly, female subjects were younger than male subjects on average. We also found that female FH carriers tend to have higher LDL-C than male FH carriers, whereas female and male noncarriers have relatively similar LDL-C. Statin use and CAD risk factors are less prevalent among female subjects compared with males (Table S6 in the Supplemental Material). We did not find a statistically significant interaction between ancestry and FH carrier status. Although, we saw a trend towards significance for African ancestry, and stratified analysis showed a higher risk estimate within the African ancestry group (Figure 4). Notably, MVP subjects with African ancestry tended to be younger than those with European ancestry (Table S7 in the Supplemental Material).

Lastly, we sought to determine if the incomplete attenuation pattern we observed in this study was primarily driven by subjects with the limited historical

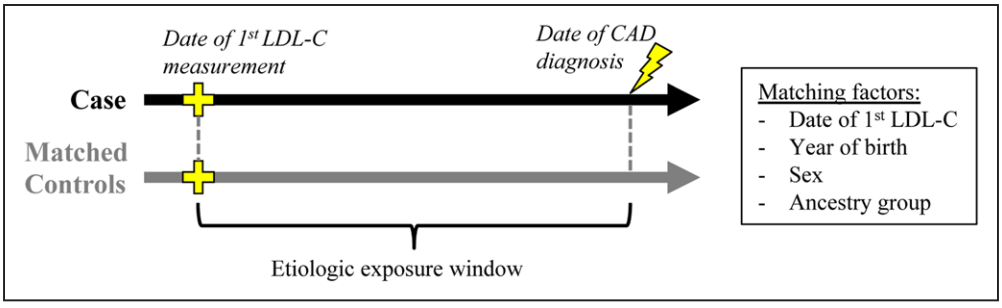

Figure 3. Illustration of case-control sets with matched etiologic exposure windows.
Incidence density sampling was used to generate matched sets for the nested case-control study. For each case, the index date was set to the date of the first coronary artery disease (CAD) code. Any subject with no CAD codes before or within 1 mo after the index date was eligible to serve as a control, and 10 random controls were selected, matching on the date of the first LDL-C (low-density lipoprotein cholesterol) measurement, the year of birth, sex, and ancestry.

Table 3. Characteristic of the Nested Case-Control Cohort

Characteristic	CAD cases	Matched controls
Demographics		
n	23 091	230 910
Male	22 497 (97.4)	224 970 (97.4)
Age at enrollment, y	66.3±9.1	66.3±9.0
Ancestry group		
African	3620 (15.7)	36 200 (15.7)
Asian	144 (0.6)	1440 (0.6)
European	17 553 (76.0)	175 530 (76.0)
Hispanic	1434 (6.2)	14 340 (6.2)
Unclassified	340 (1.5)	3400 (1.5)
Lipid data		
Age at first LDL-C, y	57.3±9.0	57.2±9.0
First LDL-C to index date, y	5.7±4.6	5.7±4.5
LDL-C, mg/dL		
First	131.6±42.5	125.2±38.5
Maximum before index date	164.0±54.8	151.0±48.5
Mean before index date	130.7±36.5	124.1±32.9
Medical history		
Hypertension		
before first LDL-C	11 447 (49.6)	93 759 (40.6)
before index date	17 938 (77.7)	148 201 (64.2)
Diabetes		
before first LDL-C	5634 (24.4)	33 379 (14.5)
before index date	9438 (40.9)	61 175 (26.5)
Tobacco		
before first LDL-C	4115 (17.8)	32 561 (14.1)
before index date	8320 (36.0)	62 993 (27.3)
Statin use		
before first LDL-C	3882 (16.8)	29 076 (12.6)
before index date	14 644 (63.4)	113 023 (48.9)
FH variant carrier	103 (0.4)	651 (0.3)
Case type		
Prevalent cases	17 642 (76.4)	NA
Index date to enrollment, y	5.7 (4.1)	NA
Incident cases	5449 (23.6)	NA
Enrollment to index date, y	2.0 (1.5)	NA

Values are n (%) or mean±SD. To convert LDL-C values from mg/dL to mmol/L, divide by 38.67. CAD indicates coronary artery disease; FH, familial hypercholesterolemia; LDL-C, low-density lipoprotein cholesterol; and NA, not applicable.

LDL-C data. We, therefore, generated matched sets of subjects with extensive LDL-C data. In a matched cohort requiring ≥5 LDL-C measures spanning ≥5 years before the index date (9786 cases, 97 860 controls) and in a matched cohort requiring ≥10 LDL-C measures spanning ≥10 years (3615 cases, 36 150 controls), we did not observe any notable differences in the degree of attenuation of the risk for CAD (Tables S8 and S9 in the Supplemental Material).

DISCUSSION

In this study, we aimed to determine if the longitudinal LDL-C exposure observed in medical records can account for the increased CAD risk among carriers of FH genetic variants. We adopted a nested case-control design and carefully matched the etiologic exposure window of case-control sets using the principle of incidence density sampling. We showed that adjusting for longitudinal LDL-C exposure using multiple measurements does not fully attenuate the CAD risk associated with carrying an FH variant, even when extensive LDL-C records are available.

We found evidence of a modification of effect of FH variant carrier status by sex. Among female subjects, the CAD risk was higher with and without LDL-C adjustment. This difference may be due to less survival bias among the female participants, who were younger than the male participants and had fewer risk factors. Other sex differences may also contribute. For example, across childhood and adolescence, untreated girls with FH demonstrate consistently higher LDL-C levels than untreated boys,[11] and adult women with FH may be undertreated compared with men.[12] We observed patterns in MVP consistent with these prior findings, but additional studies are needed to better understand sex differences while accounting for several potential confounders.

A strength of MVP is the genetic diversity, which is more reflective of the US population than European biobanks. To our knowledge, this study is the largest to date to estimate the CAD risk associated with FH variant carrier status among persons with significant African ancestry. We found that carrying an FH variant conferred greater CAD risk among this group compared with subjects of European ancestry. This difference may reflect selection biases that occur with stratification. However, racial disparities in the treatment of FH may contribute. For example, in an analysis of self-reported race and ethnicity in the CASCADE-FH registry (Cascade Screening for Awareness and Detection of Familial Hypercholesterolemia), US Blacks were more likely to be undertreated compared with White patients.[12] In our cohort, statin use among FH carriers of African and European ancestry was similar (Table S7 in the Supplemental Material), but additional work is needed to assess timing and adequacy of treatment.

In sum, our observations support the notion that genetic testing adds important predictive value to standard clinical assessment, even when longitudinal LDL-C measures are considered. This finding is consistent with a recently proposed framework that recommends both LDL-C measurement and genetic assessment to identify the highest risk patients.[13] Our study suggests that among adults, typical LDL-C monitoring does not optimally stratify subjects by their lifelong exposure to LDL-C. The cholesterol exposure pattern of FH carriers versus noncarriers is most distinct during childhood.[14] We hypothesize that much of the excess risk associated

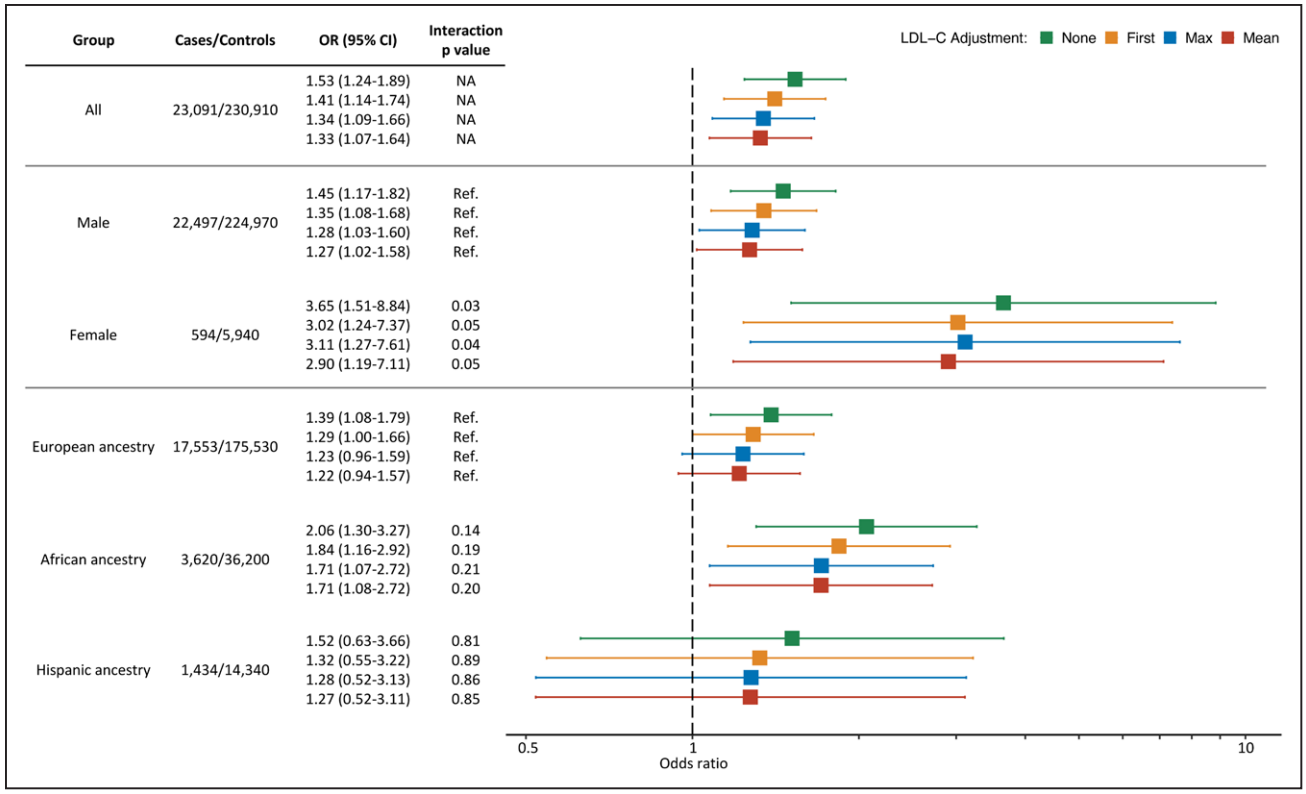

Figure 4. Association between familial hypercholesterolemia (FH) variants and coronary artery disease (CAD) with adjustments for historical LDL-C (low-density lipoprotein cholesterol) exposure.
Risk of CAD associated with FH genetic variants in the full cohort (**top** segment) and with stratification by sex (**middle** segment) and ancestry (**bottom** segment). Interaction P values are listed where appropriate, and "Ref." denotes the reference group. Odds ratios (OR) were estimated using logistic regression, adjusting for the indicated LDL-C (low-density lipoprotein cholesterol) metric in addition to the nested case-control matching factors, tobacco use, hypertension, diabetes, statin prescription, and number of LDL-C measurements during the exposure window. NA indicates not applicable.

with FH variants accumulate during childhood and early adulthood, a time when a majority are not treated. Thus, adult FH carriers and noncarriers who demonstrate similar patterns of LDL-C may have already separated their risk trajectories in the decades before LDL-C monitoring.

Pediatric guidelines recommend screening LDL-C in children to identify FH early in life.[15] It is possible that if childhood LDL-C data were available, adjustment for LDL-C exposure over a greater fraction of one's lifetime may supplant the predictive power of FH variant carrier status. However, evaluation of lifelong LDL-C measurements is not currently feasible in most clinical settings, whereas genetic testing is rapidly becoming widely available.

The cost-effectiveness of genetic testing for FH remains a debate. Cascade screening is one cost-effective strategy,[16] but it is underutilized in the United States.[17] Universal screening may ultimately prove cost-effective when considering the possibility of simultaneously testing for actionable genetic variants across multiple syndromes. For example, ≈1% of UK Biobank subjects harbor pathogenic variants for FH, hereditary breast or ovarian cancer syndrome, or Lynch syndrome.[18] As genetic testing becomes more informative for a wider spectrum of diseases, and as the cost continues to decline, we expect genetic risk assessment to become an integral part of primary prevention. The existence of effective, safe, and inexpensive primary prevention strategies such as lifestyle counseling and statins affords CAD a major advantage in this respect. Efforts are underway within MVP to implement return of actionable results to research participants, and the presence of an FH variant is one such actionable result being explored.

Study Limitations

We note several limitations of our study. First, a majority of the CAD cases are prevalent, occurring up to 20 years before enrollment. Although we implemented a prospective analysis, our risk estimates still suffer from survivor bias because only prevalent cases that survived to enroll in MVP could be observed. Moreover, MVP participants tend to be older at enrollment and have more CAD risk factors when compared with other biobanks, further enhancing survivor bias. Thus, our study likely underestimates the risk of FH variants. However, underestimating the risk of FH is not expected to alter our main conclusion regarding patterns of risk attenuation.

A second limitation of our study is the use of a genotyping array rather than gene sequencing to identify FH variants. Although the MVP array is designed to detect rare protein-altering variants and known disease-causing variants, we expect to miss some variants that would be identified through sequencing. In particular, we were not able to evaluate for copy number variants, which likely account for 5-10% of FH variants at the *LDLR* locus.[19,20] Based on prior US data[4] as well as a recent global meta-analysis,[21] we may reasonably estimate the expected prevalence of FH variant carriers in our cohort to be no more than 1 in ≈250 to 300. We observed a prevalence of 1 in 303 in this study. Thus, we expect the number of missed carriers to be quite small and to have minimal impact on our analysis. Corroborating this supposition, we found that our risk estimates for CAD are consistent with other population studies that identified FH carriers through sequencing (Figure S1 in the Supplemental Material).

A third limitation of our analysis is that it does not capture care provided outside of the VA. Lab measurements, prescriptions, and diagnoses that only occurred in non-VA settings may be missed. However, we do not expect such missing data to be substantial or to alter our basic conclusions.

A fourth limitation of our study is that we used extensive prescription data to account for statin use, but we did not account for nonstatin LDL-lowering medications. The best approach for adjusting longitudinal LDL-C data for different classes and combinations of medications is unknown and will require future research efforts. Importantly, PCSK9 inhibitors were not available or prescribed in the VA healthcare system for nearly all of the study period.

Lastly, the MVP cohort is predominantly male. Our risk estimates are less precise in women due to a small sample size. Larger studies of FH among women are needed to confirm our findings and to better understand potential sex differences.

In conclusion, FH genetic variants confer significant risk for CAD that is independent of LDL-C exposure as defined by longitudinal measurements in the EHR. We believe that the residual risk associated with FH variants reflects the limitations of clinical phenotyping for capturing genetic risk. Whereas FH variants impact LDL-C exposure continuously throughout life, clinical measurements of LDL-C can only sample a fraction of this exposure.

ARTICLE INFORMATION

Received June 16, 2021; accepted January 17, 2022.

Affiliations

VA Palo Alto Health Care system, CA (S.L.C., C.T., A.T.H., P.S.T., T.L.A.). Division of Cardiovascular Medicine, Department of Medicine, Stanford University School of Medicine, CA (S.L.C., C.T., A.T.H., J.W.K., P.S.T., T.L.A.). VA Informatics & Computing Infrastructure, VA Salt Lake City Health Care System, Salt Lake City, UT (K.M.L., J.L.). College of Nursing & Health Sciences, University of Massachusetts, Boston (J.L.). Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA (K.-M.C.). Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia (K.-M.C., D.J.R.). Edith Nourse Rogers Memorial VA Hospital, Bedford, MA (D.M.). Center for Population Health, University of Massachusetts, Lowell, MA (D.M.). Diabetes Research Center (J.W.K.) and Cardiovascular Institute (J.W.K., P.S.T., T.L.A.), Stanford University School of Medicine, CA. VA Boston Healthcare System, Boston, MA (C.O., J.M.G.). Department of Medicine, Harvard Medical School, Boston, MA (C.O.). Atlanta VA Medical Center, Decatur, GA (P.W.W., Y.V.S.). Department of Medicine, Emory University School of Medicine, Atlanta, GA (P.W.W.). Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA (P.W.W., Y.V.S.).

Acknowledgments

We greatly appreciate the participation of the Million Veterans Program participants and the support of the Million Veterans Program staff.

Sources of Funding

This work was supported by funding from the Department of Veterans Affairs Office of Research and Development, Million Veteran Program Grants 2I01BX003362 and 1I01BX004821. This publication does not represent the views of the Department of Veteran Affairs or the United States Government.

Disclosures

None.

Supplemental Materials

Supplemental Methods
Tables S1–S11
Figures S1–S3
References[22–43]

REFERENCES

1. Versmissen J, Oosterveer DM, Yazdanpanah M, Defesche JC, Basart DC, Liem AH, Heeringa J, Witteman JC, Lansberg PJ, Kastelein JJ, et al. Efficacy of statins in familial hypercholesterolaemia: a long term cohort study. *BMJ*. 2008;337:a2423. doi: 10.1136/bmj.a2423
2. Luirink IK, Wiegman A, Kusters DM, Hof MH, Groothoff JW, de Groot E, Kastelein JJP, Hutten BA. 20-year follow-up of statins in children with familial hypercholesterolemia. *N Engl J Med*. 2019;381:1547–1556. doi: 10.1056/NEJMoa1816454
3. Nordestgaard BG, Chapman MJ, Humphries SE, Ginsberg HN, Masana L, Descamps OS, Wiklund O, Hegele RA, Raal FJ, Defesche JC, et al; European Atherosclerosis Society Consensus Panel. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society. *Eur Heart J*. 2013;34:3478–390a. doi: 10.1093/eurheartj/eht273
4. Abul-Husn NS, Manickam K, Jones LK, Wright EA, Hartzel DN, Gonzaga-Jauregui C, O'Dushlaine C, Leader JB, Lester Kirchner H, Lindbuchler DM, et al. Genetic identification of familial hypercholesterolemia within a single U.S. health care system. *Science*. 2016;354:aaf7000. doi: 10.1126/science.aaf7000
5. Khera AV, Won HH, Peloso GM, Lawson KS, Bartz TM, Deng X, van Leeuwen EM, Natarajan P, Emdin CA, Bick AG, et al. Diagnostic yield and clinical utility of sequencing familial hypercholesterolemia genes in patients with severe hypercholesterolemia. *J Am Coll Cardiol*. 2016;67:2578–2589. doi: 10.1016/j.jacc.2016.03.520
6. Trinder M, Francis GA, Brunham LR. Association of monogenic vs polygenic hypercholesterolemia with risk of atherosclerotic cardiovascular disease. *JAMA Cardiol*. 2020;5:390–399. doi: 10.1001/jamacardio.2019.5954.
7. Sturm AC, Knowles JW, Gidding SS, Ahmad ZS, Ahmed CD, Ballantyne CM, Baum SJ, Bourbon M, Carrié A, Cuchel M, et al; Convened by the Familial Hypercholesterolemia Foundation. Clinical genetic testing for familial hypercholesterolemia: JACC scientific expert panel. *J Am Coll Cardiol*. 2018;72:662–680. doi: 10.1016/j.jacc.2018.05.044
8. Gaziano JM, Concato J, Brophy M, Fiore L, Pyarajan S, Breeling J, Whitbourne S, Deen J, Shannon C, Humphries D, et al. Million veteran program: a mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol*. 2016;70:214–223. doi: 10.1016/j.jclinepi.2015.09.016
9. Sun YV, Damrauer SM, Hui Q, Assimes TL, Ho YL, Natarajan P, Klarin D, Huang J, Lynch J, DuVall SL, et al. Effects of genetic variants

associated with familial hypercholesterolemia on low-density lipoprotein-cholesterol levels and cardiovascular outcomes in the million veteran program. *Circ Genom Precis Med.* 2018;11:e002192. doi: 10.1161/CIRCGEN.118.002192
10. Essebag V, Genest J Jr, Suissa S, Pilote L. The nested case-control study in cardiology. *Am Heart J.* 2003;146:581–590. doi: 10.1016/S0002-8703(03)00512-X
11. Holven KB, Narverud I, van Lennep JR, Versmissen J, Øyri LKL, Galema-Boers A, Langslet G, Ulven SM, Veierød MB, Retterstøl K, et al. Sex differences in cholesterol levels from birth to 19 years of age may lead to increased cholesterol burden in females with FH. *J Clin Lipidol.* 2018;12:748–755.e2. doi: 10.1016/j.jacl.2018.02.021
12. Amrock SM, Duell PB, Knickelbine T, Martin SS, O'Brien EC, Watson KE, Mitri J, Kindt I, Shrader P, Baum SJ, et al. Health disparities among adult patients with a phenotypic diagnosis of familial hypercholesterolemia in the CASCADE-FH™ patient registry. *Atherosclerosis.* 2017;267:19–26. doi: 10.1016/j.atherosclerosis.2017.10.006
13. Khera AV, Hegele RA. What is familial hypercholesterolemia, and why does it matter? *Circulation.* 2020;141:1760–1763. doi: 10.1161/CIRCULATIONAHA.120.046961
14. Wald DS, Bestwick JP, Wald NJ. Child-parent screening for familial hypercholesterolaemia: screening strategy based on a meta-analysis. *BMJ.* 2007;335:599. doi: 10.1136/bmj.39300.616076.55
15. Expert Panel on Integrated Guidelines for Cardiovascular Health and Risk Reduction in Children and Adolescents, National Heart, Lung, and Blood Institute. Expert panel on integrated guidelines for cardiovascular health and risk reduction in children and adolescents: summary report. *Pediatrics.* 2011;128(Suppl 5):S213–S256. doi: 10.1542/peds.2009-2107C
16. Kerr M, Pears R, Miedzybrodzka Z, Haralambos K, Cather M, Watson M, Humphries SE. Cost effectiveness of cascade testing for familial hypercholesterolaemia, based on data from familial hypercholesterolaemia services in the UK. *Eur Heart J.* 2017;38:1832–1839. doi: 10.1093/eurheartj/ehx111
17. Ahmad ZS, Andersen RL, Andersen LH, O'Brien EC, Kindt I, Shrader P, Vasandani C, Newman CB, deGoma EM, Baum SJ, et al. US physician practices for diagnosing familial hypercholesterolemia: data from the CASCADE-FH registry. *J Clin Lipidol.* 2016;10:1223–1229. doi: 10.1016/j.jacl.2016.07.011
18. Patel AP, Wang M, Fahed AC, Mason-Suares H, Brockman D, Pelletier R, Amr S, Machini K, Hawley M, Witkowski L, et al. Association of rare pathogenic DNA variants for familial hypercholesterolemia, hereditary breast and ovarian cancer syndrome, and lynch syndrome with disease risk in adults according to family history. *JAMA Netw Open.* 2020;3:e203959. doi: 10.1001/jamanetworkopen.2020.3959
19. Wang J, Ban MR, Hegele RA. Multiplex ligation-dependent probe amplification of LDLR enhances molecular diagnosis of familial hypercholesterolemia. *J Lipid Res.* 2005;46:366–372. doi: 10.1194/jlr.D400030-JLR200
20. Iacocca MA, Wang J, Dron JS, Robinson JF, McIntyre AD, Cao H, Hegele RA. Use of next-generation sequencing to detect LDLR gene copy number variation in familial hypercholesterolemia. *J Lipid Res.* 2017;58:2202–2209. doi: 10.1194/jlr.D079301
21. Beheshti SO, Madsen CM, Varbo A, Nordestgaard BG. Worldwide prevalence of familial hypercholesterolemia: meta-analyses of 11 million subjects. *J Am Coll Cardiol.* 2020;75:2553–2566. doi: 10.1016/j.jacc.2020.03.057
22. Justice AC, Lasky E, McGinnis KA, Skanderson M, Conigliaro J, Fultz SL, Crothers K, Rabeneck L, Rodriguez-Barradas M, Weissman SB, et al; VACS 3 Project Team. Medical disease and alcohol use among veterans with human immunodeficiency infection: a comparison of disease measurement strategies. *Med Care.* 2006;44(8 Suppl 2):S52–S60. doi: 10.1097/01.mlr.0000228003.08925.8c
23. Goulet JL, Erdos J, Kancir S, Levin FL, Wright SM, Daniels SM, Nilan L, Justice AC. Measuring performance directly using the veterans health administration electronic medical record: a comparison with external peer review. *Med Care.* 2007;45:73–79. doi: 10.1097/01.mlr.0000244510.09001.e5
24. McGinnis KA, Brandt CA, Skanderson M, Justice AC, Shahrir S, Butt AA, Brown ST, Freiberg MS, Gibert CL, Goetz MB, et al. Validating smoking data from the Veteran's Affairs Health Factors dataset, an electronic data source. *Nicotine Tob Res.* 2011;13:1233–1239. doi: 10.1093/ntr/ntr206
25. Goulet JL, Brandt C, Crystal S, Fiellin DA, Gibert C, Gordon AJ, Kerns RD, Maisto S, Justice AC. Agreement between electronic medical record-based and self-administered pain numeric rating scale: clinical and research implications. *Med Care.* 2013;51:245–250. doi: 10.1097/MLR.0b013e318277f1ad
26. McGinnis KA, Tate JP, Williams EC, Skanderson M, Bryant KJ, Gordon AJ, Kraemer KL, Maisto SA, Crystal S, Fiellin DA, et al. Comparison of AUDIT-C collected via electronic medical record and self-administered research survey in HIV infected and uninfected patients. *Drug Alcohol Depend.* 2016;168:196–202. doi: 10.1016/j.drugalcdep.2016.09.015
27. Calhoun PS, Wilson SM, Hertzberg JS, Kirby AC, McDonald SD, Dennis PA, Bastian LA, Dedert EA, Beckham JC; VA Mid-Atlantic MIRECC Workgroup. Validation of veterans affairs electronic medical record smoking data among Iraq- and Afghanistan-Era Veterans. *J Gen Intern Med.* 2017;32:1228–1234. doi: 10.1007/s11606-017-4144-5
28. Klarin D, Lynch J, Aragam K, Chaffin M, Assimes TL, Huang J, Lee KM, Shao Q, Huffman JE, Natarajan P, et al; VA Million Veteran Program. Genome-wide association study of peripheral artery disease in the Million Veteran Program. *Nat Med.* 2019;25:1274–1279. doi: 10.1038/s41591-019-0492-5
29. Golden SE, Hooker ER, Shull S, Howard M, Crothers K, Thompson RF, Slatore CG. Validity of Veterans Health Administration structured data to determine accurate smoking status. *Health Informatics J.* 2020;26:1507–1515. doi: 10.1177/1460458219882259
30. McGinnis KA, Justice AC, Bailin S, Wellons M, Freiberg M, Koethe JR. High concordance between chart review adjudication and electronic medical record data to identify prevalent and incident diabetes mellitus among persons with and without HIV. *Pharmacoepidemiol Drug Saf.* 2020;29:1432–1439. doi: 10.1002/pds.5111
31. Rodriguez-Barradas MC, McGinnis KA, Akgün K, Tate JP, Brown ST, Butt AA, Fine M, Goetz MB, Graber CJ, Huang L, et al. Validation for using electronic health records to identify community acquired pneumonia hospitalization among people with and without HIV. *Pneumonia (Nathan).* 2020;12:6. doi: 10.1186/s41479-020-00068-1
32. Hsu F, Kent WJ, Clawson H, Kuhn RM, Diekhans M, Haussler D. The UCSC known genes. *Bioinformatics.* 2006;22:1036–1046. doi: 10.1093/bioinformatics/btl048
33. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4:7. doi: 10.1186/s13742-015-0047-8
34. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR; Genomes Project Consortium. A global reference for human genetic variation. *Nature.* 2015;526:68–74. doi: 10.1038/nature15393
35. Do R, Stitziel NO, Won HH, Jørgensen AB, Duga S, Angelica Merlini P, Kiezun A, Farrall M, Goel A, Zuk O, et al; NHLBI Exome Sequencing Project. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature.* 2015;518:102–106. doi: 10.1038/nature13917
36. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, Maglott DR. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* 2014;42(Database issue):D980–D985. doi: 10.1093/nar/gkt1113
37. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38:e164. doi: 10.1093/nar/gkq603
38. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: a one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Hum Mutat.* 2016;37:235–241. doi: 10.1002/humu.22932
39. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, et al; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17:405–424. doi: 10.1038/gim.2015.30
40. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics.* 2010;26:2867–2873. doi: 10.1093/bioinformatics/btq559
41. Fang H, Hui Q, Lynch J, Honerlaw J, Assimes TL, Huang J, Vujkovic M, Damrauer SM, Pyarajan S, Gaziano JM, et al; VA Million Veteran Program. Harmonizing Genetic Ancestry and Self-identified Race/Ethnicity in Genome-wide Association Studies. *Am J Hum Genet.* 2019;105:763–772. doi: 10.1016/j.ajhg.2019.08.012
42. Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, Kirby A, Sourjina T, Peto R, Collins R, et al; Cholesterol Treatment Trialists' (CTT) Collaborators. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet.* 2005;366:1267–1278. doi: 10.1016/S0140-6736(05)67394-1
43. Pearce N. Analysis of matched case-control studies. *BMJ.* 2016;352:i969. doi: 10.1136/bmj.i969