Polygenic Hyperlipidemias and Coronary Artery Disease Risk

Ripatti, Polygenic Hyperlipidemias and CAD Risk

Pietari Ripatti[1], MD; Joel T Rämö[1], MD; Nina J Mars[1], MD, PhD; Sanni Söderlund[2,3], MD, PhD; Christian Benner[1], PhD; Ida Surakka[1,4], PhD; Tuomo Kiiskinen[1], MD; Aki S Havulinna[1,5], PhD; Priit Palta[1], MD, PhD; Nelson B Freimer[6], MD, PhD; Veikko Salomaa[5], MD, PhD; Matti Pirinen[1,7,8], PhD; FinnGen, Aarno Palotie[1,9-13], MD, PhD; Marja-Riitta Taskinen[3,14¶], MD, PhD; and Samuli Ripatti[1,7,15¶*], PhD.

[1]Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Helsinki, Finland

[2]Department of Internal Medicine, Helsinki University Hospital, Helsinki, Finland

[3]Research Programs Unit, Diabetes & Obesity, University of Helsinki, Helsinki, Finland

[4]Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA

[5]Department of Public Health Solutions, National Institute for Health and Welfare, Helsinki, Finland

[6]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, California, USA

[7]Department of Public Health, Clinicum, Faculty of Medicine, University of Helsinki, Helsinki, Finland

[8]Department of Mathematics and Statistics, Faculty of Science, University of Helsinki, Helsinki, Finland

[9]Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

[10]The Stanley Center for Psychiatric Research, The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

[11]Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA

[12]Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA

[13]Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts, USA

[14]Clinical Research Institute HUCH Ltd., Helsinki, Finland

[15]The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

[¶]These authors contributed equally to this work.

*Corresponding author

Street address: Tukholmankatu 8, FI-00290 Helsinki, Finland

Correspondence: P.O. Box 20, FI-00014 University of Helsinki, Helsinki, Finland

Tel: +358 40 567 0826

E-mail: samuli.ripatti@helsinki.fi

Word count: 4605.

Journal Subject Terms:

Biomarkers, Lipids and Cholesterol, Epidemiology, Risk Factors, Genetics, Coronary Artery Disease

Abstract

Background

Hyperlipidemia is a highly heritable risk factor for coronary artery disease (CAD). Monogenic familial hypercholesterolemia associates with higher increase in CAD risk than expected from a single LDL-C measurement, likely due to lifelong cumulative exposure to high LDL-C. It remains unclear to what extent a high polygenic load of LDL-C or TG-increasing variants associates with increased CAD risk.

Methods and Results

We derived polygenic risk scores (PRS) with ~6M variants for LDL-C and TG with weights from a UK biobank-based genome-wide association study with ~500K samples. We evaluated the impact of polygenic hypercholesterolemia and hypertriglyceridemia to lipid levels in 27 039 individuals from the FINRISK cohort, and to CAD risk in 135 300 individuals (13 695 CAD cases) from the FinnGen project.

In FINRISK, LDL-C ranged from 2.83 (95% CI 2.79-2.89) to 3.80 (3.72-3.88) and TG from 0.99 (0.95-1.01) to 1.52 (1.48-1.58) mmol/l between the lowest and highest 5% of the respective PRS distributions. The corresponding CAD prevalences ranged from 8.2% to 12.7% for the LDL-C PRS and from 8.2% to 12.1% for the TG PRS in FinnGen. Furthermore, CAD risk was 1.36-fold higher (OR, 95% CI 1.24-1.49) for the LDL-C PRS and 1.31-fold higher (1.20-1.44) for the TG PRS for those with the PRS >95th percentile vs those without. These estimates were only slightly attenuated when adjusting for a CAD PRS (OR 1.26 [95% CI 1.15-1.39] for LDL-C and 1.21 [1.10-1.32] for TG PRS).

Conclusions

The CAD risk associated with a high polygenic load for lipid-increasing variants was proportional to their impact on lipid levels and mostly independent of a CAD PRS. In contrast with a PRS for CAD, the lipid PRSs point to known and directly modifiable risk factors providing more direct guidance for clinical translation.

Key Words

coronary artery disease; hypercholesterolemia; hypertriglyceridemia; polygenic risk score; polygenic hyperlipidemia

Introduction

Hypercholesterolemia, particularly high LDL-cholesterol (LDL-C), is an established, heritable, and treatable risk factor for coronary artery disease (CAD).[1, 2] Additionally, accumulating evidence suggests that increased triglycerides (TG; hypertriglyceridemia) are causally linked to CAD.[3-5]

Increased levels of both LDL-C and TGs result from a combination of genetic and non-genetic factors.[6, 7] Genetic factors include rare highly penetrant variants and a long tail of common variants with smaller effect sizes. While high impact variants in the *LDLR*, *PCSK9,* and *APOB* genes cause familial hypercholesterolemia, it has also been suggested that similarly high LDL-C levels could result from a high polygenic burden of LDL-C-increasing variants.[8, 9] While monogenic FH with an identified mutation associates with a higher CAD risk than expected on the basis of a single LDL-C measurement, the contribution of an accumulation of a large number of LDL-C-increasing alleles to CAD risk is unclear.[10]

Similarly to hypercholesterolemia, both polygenic burden and highly penetrant variants contribute to hypertriglyceridemia.[6] However, highly penetrant variants underlying hypertriglyceridemia are much fewer and very rare (estimated population prevalence 1:1 000 000).[6] On the other hand, many individuals with hypertriglyceridemia have a high polygenic burden of TG-increasing variants.[6] Unlike LDL-C, it is unknown whether genetically increased TGs confer higher CAD risk than non-genetic hypertriglyceridemia. Genetics supporting a causal link between hypertriglyceridemia and CAD, and the evidence for beneficial therapeutic reducing of TGs to reduce CVD risk, however, highlight the potential also for association between polygenic load of TG elevating alleles and CAD risk.[3-5, 11, 12]

In this cohort study of 27 039 individuals from the Finnish FINRISK population cohort with lipid measurements, and 135 300 individuals including 13 695 CAD cases from the FinnGen project, we evaluated the impact of high polygenic LDL-C and TG to CAD risk. We developed separate genome-wide PRSs for both LDL-C and TG to define polygenic hypercholesterolemia and hypertriglyceridemia. First, we tested to what extent PRSs for LDL-C and TG associate with measured lipid levels. Second, we tested to what degree polygenic hypercholesterolemia and polygenic hypertriglyceridemia associate with increased risk for CAD.

Methods

Ethics Statement

All samples were collected in accordance with the Declaration of Helsinki. For the Finnish Institute of Health and Welfare (THL) driven FinnGen preparatory project (here called FinnGen), all patients and control subjects provided informed consent for biobank research, based on the Finnish Biobank Act. Alternatively, older cohorts were based on study-specific consents and later transferred to the THL Biobank after approval by Valvira, the National Supervisory Authority for Welfare and Health. Recruitment protocols followed the biobank protocols approved by Valvira. The FinnGen project was additionally approved by THL (approval numbers THL/2031/6.02.00/2017, and amendments THL/341/6.02.00/2018, THL/2222/6.02.00/2018, and THL/283/6.02.00/2019). Written informed consent was obtained from all participants except the 1992 FINRISK survey, for which verbal informed consent was obtained as required by legislation and ethics committees at the time. Earlier FINRISK surveys were approved by various ethics committees.[13] The Coordinating Ethics Committee of the Helsinki and Uusimaa Hospital District approved the FinnGen project (number HUS/990/2017) and the 2007 and 2012 FINRISK surveys. The North West Multi-Centre Research Ethics Committee approved the UKBB study.

Subjects and Measurements

The National FINRISK Study is a Finnish population survey conducted every 5 years since 1972 with independent, random, and representative samples across the country.[13] We used 27 039 individuals from the 1992 to 2012 collections. Circulating biochemical markers were measured from venous blood samples drawn after a minimum of 4-h fast using standard methods.[13] The effect of lipid-lowering therapy in those using medication was adjusted for by dividing LDL-C by 0.7 as utilized previously.[10] LDL-C was calculated using the Friedewald formula.[14] Non-HDL-C was calculated as total cholesterol (TC) - HDL-C and remnant cholesterol (remnant-C) as TC - HDL-C - LDL-C.

The FinnGen preparatory phase aggregates Finnish biobank samples and currently comprises 135 300 participants.[15] The samples have been linked with national hospital discharge and causes-of-death registries. Clinical CAD event endpoints were constructed from major adverse coronary events (MACE)

defined as either myocardial infarction (MI) (International Classification of Diseases [ICD]-10 codes I20.0 or I21-22, ICD-9 410 or 411.0, or ICD-8 410 or 411.0 for hospital discharge; or ICD-10 I21-25, I46, R96, or R98, ICD-9 410-414 or 798 [excluding 7980A], or ICD-8 410-414 or 798 for main cause-of-death) or coronary revascularization (coronary angioplasty [PCI] or coronary artery bypass grafting [CABG]).[16]

The UK Biobank comprises extensive phenotypic data on some 500 000 individuals of the general UK population between 40 and 69 years.[17] All participants were interviewed, answered standardised questionnaires, and had physical measurements taken at baseline. The UKBB cohort was linked with national Hospital Episode Statistics, cancer, and death registry data.[17] Circulating biochemical markers were measured from serum samples drawn after a mean fasting time of 3.8 hours. LDL was measured using enzymatic selective protection and TG using enzymatic methods (http://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/serum_biochemistry.pdf).

Genotyping and Polygenic Risk Score Calculation

Samples were genotyped and imputed using standard methods as described in S1 Text.

The LDL-C, TG, and CAD PRSs were calculated as the sum of the risk allele dosages weighted by their effect sizes using LDpred.[18] The recent LDpred method is a Bayesian approach to calculate a posterior mean effect size for each variant based on a prior of effect size and linkage disequilibrium (LD; a measure of how much a variant correlates with other variants).[18] Whole-genome sequences from 503 European samples from the 1000 Genomes project phase 3 served as the LD reference population for LDpred.[19] We utilised the infinitesimal prior on the fraction of causal variants in a given phenotype.

The weights for the lipid PRSs were based on a genome-wide association study (GWAS) of 468 732 samples with LDL-C measurements and 469 240 with TG measurements from the UKBB. As part of quality control, related subjects and subjects taking lipid-lowering medicine were excluded. We performed the lipid GWAS using BOLT-LMM and adjusted for sex, age and the first 15 principal components.[20] The weights for the CAD PRS were based on summary statistics obtained from a GWAS of ischemic heart disease (IHD) in the UKBB (PheCode 411) performed using SAIGE.[21] The PRSs were calculated using PLINK 2.0 Alpha 1.[22] The final PRSs included 5 707 489 variants for LDL-C and TG and 5 709 394 variants for CAD.

Statistical Analysis

Variation explained by PRSs was estimated as adjusted r^2 from linear regression, with residual lipid measurements after adjusting for age and sex as the response. TG measurements were additionally log-transformed. Bootstrapping with percentile CIs of served to estimate median lipid levels in PRS bins. Binomial logistic regression served to estimate odds ratios (OR) for CAD outcomes. The logistic regression models were adjusted for age, sex, first ten principal components, and genotyping batch. All tests were two-sided. Statistical analyses were performed using R (version 3.6.1).[23]

Results

Polygenic Hyperlipidemias and Lipid Levels

We first defined PRSs for LDL-C and TG using an approach of reweighting the effects of genome-wide variants using GWAS summary statistics and the LD structure of a reference population implemented in the software package LDpred.[18] As the largest freely available population-wide dataset of lipid measures and genetic markers, we drew the summary statistics from a GWAS of ~500 000 individuals from the UKBB with lipid measures and tested the association between PRSs and lipid levels in the Finnish FINRISK study, independent of the original GWAS dataset. The FINRISK study comprised 27 039 individuals randomly drawn from the Finnish population (Table 1). Median LDL-C was 3.39 mmol/l and TG 1.19 mmol/l in the whole cohort with slightly lower values in the more recent collections (S1 Figure).

[Table 1 placeholder]

The PRSs consisted of six million markers and explained 5.3% (adjusted r^2) of variation in LDL-C and 4.9% in TG. In FINRISK, median LDL-C was 2.83 (95% CI 2.79-2.89) mmol/l in the lowest and 3.80 (3.72-3.88) mmol/l in the highest 5% of the LDL-C PRS distribution (Figure 1 a). Similarly, median TG was 0.99 (95% CI 0.95-1.01) mmol/l in the lowest and 1.52 (1.48-1.58) mmol/l in the highest 5% of the TG PRS distribution (Figure 1 b). The correlation between the LDL-C PRS and the TG PRS was low ($r = 0.13$). All in all, the LDL-C and TG PRSs were specific to and had clear impact on their respective lipid levels.

[Figure 1 placeholder]

Polygenic Hyperlipidemias and CAD Risk

To assess how polygenic hyperlipidemia associates with CAD risk, we analysed 135 300 individuals including 13 695 registry-based CAD cases from the Finnish FinnGen project (Table 1). For polygenic hypercholesterolemia, CAD risk was 1.3-fold (OR 1.30 [95% CI 1.21-1.39]) higher for those in the highest 10% and 1.4-fold (OR 1.36 [95% CI 1.24-1.49]) higher for those in the highest 5% of the LDL-C PRS, compared to the remainder of the population (Figure 2 a). CAD prevalence was accordingly 54% higher (12.7% vs 8.2%) between the highest and lowest 5% of the LDL-C PRS distribution (Figure 3 a).

For polygenic hypertriglyceridemia CAD risk was 1.3-fold (OR 1.28 [95% CI 1.20-1.37]) higher for those in the highest 10% and also 1.3-fold (OR 1.31 [95% CI 1.20-1.44]) higher for those in the highest 5% of the TG PRS, compared to the remainder of the population (Figure 2 b). CAD prevalence was 47% higher (12.1% vs 8.2%) between the highest and lowest 5% of the TG PRS distribution (Figure 3 b).

[Figure 2 placeholder]

[Figure 3 placeholder]

We tested if the lipid PRSs improve CAD risk prediction beyond a similarly derived CAD PRS. We calculated a genome-wide CAD PRS with LDpred-based weights from a GWAS of UKBB IHD diagnoses.[21] Comparing the highest 5% to the remainder of the population, the effects of the lipid PRSs to CAD risk were attenuated only modestly when adjusted for the CAD PRS (LDL-C PRS OR 1.26 [95% CI 1.15-1.39] and TG PRS OR 1.21 [1.10-1.32]; Figure 4).

[Table 2 placeholder]

[Figure 4 placeholder]

Discussion

By developing genome-wide PRSs for LDL-C and TG, we evaluated the impact of high genetic risk for these established and causal risk factors of CAD. We showed that high polygenic burden for both LDL-C or TG associated with considerably increased LDL-C and TG levels, respectively. Similarly, polygenic hypercholesterolemia and -triglyceridemia associated with significantly increased CAD risk. Furthermore, PRSs for LDL-C and TG were mostly independent of a PRS for CAD.

Polygenic hypercholesterolemia, in our study, demonstrated 0.43 mmol/l higher LDL-C levels and 36% higher CAD risk in the highest 5% of the LDL-C PRS compared to the remainder of the population. This is considerably lower than previously reported CAD risk effects of high-impact *LDLR* FH mutations.[24] While the established high-impact *LDLR* FH mutations directly disrupt LDL receptor function causing lifelong high LDL-C levels, the effect sizes of the individual variants contributing to polygenic hypercholesterolemia are small, and they likely increase LDL-C via multiple indirect biological pathways. Whereas monogenic FH is a severe disease with very high CAD risk, polygenic hypercholesterolemia, as captured by the current PRSs, seems to have a smaller effect on LDL-C levels and CAD risk. Furthermore, the benefit of lipid-lowering therapies in individuals with polygenic hypercholesterolemia remains unknown.

In addition to hypercholesterolemia, our results show that TG levels were 0.21 mmol/l lower in the lowest 5% of the TG PRS compared to the remainder of the population, and this translated into 25% lower CAD risk. This relationship is in line with the effect of TG-lowering loss-of-function mutations in the *APOC3* and *ANGPTL4* genes that reduce TG levels by ~0.7-0.8 mmol/l and CAD risk by ~40-50%..[6, 25, 26] Our consistent results support the hypothesis that TG is causal factor for CAD. Pharmacologic TG-lowering shows promise and the benefit of TG-lowering drugs remains to be tested in individuals with polygenic hypertriglyceridemia.[11]

In our study, both LDL-C and TG PRSs associated with CAD risk also when adjusted for a CAD PRS. The key difference between intermediate biomarker PRSs (such as the lipid PRSs) and disease endpoint PRSs (such as a CAD PRS) is that for the biomarker PRSs, the mechanism of effect on clinical outcomes is more direct. The CAD PRS was based on a case-control setting of individuals with or without a CAD diagnosis with a risk of misclassifications, and correlates little with known risk factors, complicating

its interpretation and clinical implications.[27] In the presence of genetic information, biomarker PRSs could guide which intervention is taken to lower the CAD risk of an individual. How much this applies to other CAD risk factors than lipids remains unknown.

Our study has several limitations. First, as FINRISK participants fasted for a minimum of 4 hours before measuring lipid profiles, our association estimates may have been attenuated particularly between the TG PRS and TG levels. The association between the TG PRS and CAD risk, however, remains unaffected by this. Second, because the Friedewald formula is invalid for individuals with TG > 4.52 mmol/l, 456 (1.7%) FINRISK samples were excluded from LDL-C analyses.[14] Third, some variants included in the lipid PRSs are not specific to their primary lipids and have residual effects on others. Excepting a negative association between the TG PRSs and HDL-C, however, the PRSs had only minor associations with other than their primary lipids (S2 Figure). Fourth, our weights for the lipid PRSs came from the UK population and were tested in the Finnish population; our results may have limited accuracy in other ethnicities. Replication and validation in other cohorts with lipid measurements and populations is warranted in the future.

In summary, the CAD risk associated with a high polygenic load for LDL-C or TG -increasing genetic variants was proportional to their impact on lipid levels. In contrast with a PRS for CAD, the lipid PRSs point to a known and directly modifiable risk factor enabling more straightforward clinical translation. As polygenic risk scores can also be measured at any point in life, they provide powerful tools for prioritising individuals for blood lipid panel screening and subsequent evidence-based intervention.

Acknowledgements

We would like to thank Sari Kivikko, Huei-Yi Shen, and Ulla Tuomainen for management assistance. The FINRISK data used for the research were obtained from THL Biobank. We thank the THL DNA laboratory for its skilful work to produce the DNA samples used in the genotyping work, which was used in this study. Part of the genotyping was performed by the Institute for Molecular Medicine Finland Technology Centre, University of Helsinki. We thank all study participants for their generous participation in the FINRISK and UKBB studies. PR and JTR acknowledge support from the Doctoral Programme in Population Health, University of Helsinki, and JTR acknowledges support from the MD-PhD Programme, University of Helsinki. This research has been conducted using the UK Biobank Resource under application no 22627. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding Sources

This work was supported by the NIH [grant number HL113315 to S.R., M.-R.T., and A.P.]; Finnish Foundation for Cardiovascular Research [to S.R., V.S., M.-R.T. and A.P.]; Sigrid Jusélius Foundation [to S.R. and M.-R.T.]; Biocentrum Helsinki [to S.R.]; EU-project RESOLVE (EU 7th Framework Program) [grant number 305707 to M.-R.T.]; Helsinki University Central Hospital Research Funds [to M.-R.T.]; Leducq Foundation [to M.-R.T.]; Ida Montin Foundation [to P.R.]; European Atherosclerosis Society [to P.R.]; MD/PhD Program of the Faculty of Medicine, University of Helsinki [to J.T.R.]; Doctoral Programme in Population Health, University of Helsinki [to J.T.R. and P.R.]; Finnish Medical Foundation [to J.T.R.]; Emil Aaltonen Foundation [to P.R. and J.T.R.]; Biomedicum Helsinki Foundation [to J.T.R.]; Academy of Finland Center of Excellence in Complex Disease Genetics [grant numbers 213506, 129680, and 312062 to S.R. and grant number 312076 to M.P.]; Horizon 2020 Research and Innovation Programme (ePerMed) [grant number 692145 to S.R.]; HiLIFE Fellow grants 2017-2020 [to S.R.]; and Academy of Finland [grant number 298149 to I.S., grant number 288509 to M.P., grant numbers 251217 and 285380 to S.R.]. The FinnGen project is funded by two grants from Business Finland (HUS 4685/31/2016 and UH 4386/31/2016) and nine industry partners (AbbVie, AstraZeneca, Biogen, Celgene, Genentech, GSK, MSD, Pfizer and Sanofi). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosures

AP is a member of the Pfizer Genetics Scientific Advisory Panel. SR holds a HiLIFE Fellowship. VS has participated in a conference trip sponsored by Novo Nordisk and received an honorarium from the same source for participating in an advisory board meeting. He also has ongoing research collaboration with Bayer Ltd.

References

1. Emerging Risk Factors C, Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray KK, et al. Major lipids, apolipoproteins, and risk of vascular disease. *Jama*. 2009;302:1993-2000.

2. Cholesterol Treatment Trialists C, Fulcher J, O'Connell R, Voysey M, Emberson J, Blackwell L, et al. Efficacy and safety of LDL-lowering therapy among men and women: meta-analysis of individual data from 174,000 participants in 27 randomised trials. *Lancet*. 2015;385:1397-405.

3. Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet*. 2013;45:1345-1352.

4. Tg, Hdl Working Group of the Exome Sequencing Project NHL, Blood I, Crosby J, Peloso GM, Auer PL, et al. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *N Engl J Med*. 2014;371:22-31.

5. Nordestgaard BG and Varbo A. Triglycerides and cardiovascular disease. *Lancet*. 2014;384:626-35.

6. Hegele RA, Ginsberg HN, Chapman MJ, Nordestgaard BG, Kuivenhoven JA, Averna M, et al. The polygenic nature of hypertriglyceridaemia: implications for definition, diagnosis, and management. *Lancet Diabetes Endocrinol*. 2014;2:655-66.

7. Bhatnagar D, Soran H and Durrington PN. Hypercholesterolaemia and its management. *Bmj*. 2008;337:a993.

8. Talmud PJ, Shah S, Whittall R, Futema M, Howard P, Cooper JA, et al. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolaemia: a case-control study. *Lancet*. 2013;381:1293-301.

9. Ripatti P, Ramo JT, Soderlund S, Surakka I, Matikainen N, Pirinen M, et al. The Contribution of GWAS Loci in Familial Dyslipidemias. *PLoS Genet*. 2016;12:e1006078.

10. Khera AV, Won HH, Peloso GM, Lawson KS, Bartz TM, Deng X, et al. Diagnostic Yield and Clinical Utility of Sequencing Familial Hypercholesterolemia Genes in Patients With Severe Hypercholesterolemia. *Journal of the American College of Cardiology*. 2016;67:2578-89.

11. Bhatt DL, Steg PG, Miller M, Brinton EA, Jacobson TA, Ketchum SB, et al. Cardiovascular Risk Reduction with Icosapent Ethyl for Hypertriglyceridemia. *N Engl J Med*. 2019;380:11-22.

12. Catapano AL, Graham I, De Backer G, Wiklund O, Chapman MJ, Drexel H, et al. 2016 ESC/EAS Guidelines for the Management of Dyslipidaemias. *European heart journal*. 2016;37:2999-3058.

13. Borodulin K, Tolonen H, Jousilahti P, Jula A, Juolevi A, Koskinen S, et al. Cohort Profile: The National FINRISK Study. *International journal of epidemiology*. 2017.

14. Friedewald WT, Levy RI and Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem*. 1972;18:499-502.

15. Finland IfMM. FinnGen. 2019;2019.

16. Pajunen P, Koukkunen H, Ketonen M, Jerkkola T, Immonen-Raiha P, Karja-Koskenkari P, et al. The validity of the Finnish Hospital Discharge Register and Causes of Death Register data on coronary heart disease. *Eur J Cardiovasc Prev Rehabil*. 2005;12:132-7.

17. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562:203-209.

18. Vilhjalmsson BJ, Yang J, Finucane HK, Gusev A, Lindstrom S, Ripke S, et al. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *Am J Hum Genet*. 2015;97:576-92.

19. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature*. 2015;526:68-74.

20. Loh PR, Tucker G, Bulik-Sullivan BK, Vilhjalmsson BJ, Finucane HK, Salem RM, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet*. 2015;47:284-90.

21. Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet*. 2018;50:1335-1341.

22. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM and Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7.

23. R: A Language and Environment for Statistical Computing [computer program]. R Foundation for Statistical Computing; 2018.

24. Abul-Husn NS, Manickam K, Jones LK, Wright EA, Hartzel DN, Gonzaga-Jauregui C, et al. Genetic identification of familial hypercholesterolemia within a single U.S. health care system. *Science*. 2016;354.

25. Jorgensen AB, Frikke-Schmidt R, Nordestgaard BG and Tybjaerg-Hansen A. Loss-of-function mutations in APOC3 and risk of ischemic vascular disease. *N Engl J Med*. 2014;371:32-41.

26. Myocardial Infarction G, Investigators CAEC, Stitziel NO, Stirrups KE, Masca NG, Erdmann J, et al. Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *N Engl J Med*. 2016;374:1134-44.

27. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*. 2018;50:1219-1224.

Tables

Table 1. Clinical and Metabolic Characteristics of Individuals.

	FINRISK		FinnGen	
Characteristic	***n***	**Mean±SD**	***n***	**Mean±SD**
n (male/female)	27039 (12884/14155)		135300 (59074/76226)	
CAD, *n* (%)	2750 (10.2%)		13695 (10.1%)	
Lipid-lowering medication usage, *n* (%)	1658 (6.1%)			
Smoking, *n* (%)	6739 (25%)		19634 (22.0%)	
Age* (year)	27039	48.9±13.5	135300	59.2±16.6
BMI (kg/m^2)	26941	26.8±4.69	95251	27.2±5.6
Total cholesterol (mmol/l)	27024	5.49±1.08		
LDL-C (mmol/l)	26568	3.47±1.01		
Triglyceride (mmol/l)	27024	1.47±1.00		
HDL-C (mmol/l)	27024	1.44±0.381		
Apolipoprotein B (g/l)	22464	0.965±0.248		
Non-HDL-C (mmol/l)	27024	4.05±1.10		
Remnant-C (mmol/l)	26568	0.630±0.340		

LDL-C was calculated using the Friedewald formula; the effect of lipid-lowering therapy in those using medication at the time of lipid measurement was adjusted for by dividing LDL-C by 0.7 as utilized previously.[10] FinnGen lacks lipid measurements and lipid-lowering medication usage information. *Age at recruitment for FINRISK and age at end of follow-up for FinnGen. SD, standard deviation. CAD, coronary artery disease. BMI, body mass index. LDL-C, LDL-cholesterol. HDL-C, HDL-cholesterol. Non-HDL-C, non-HDL-cholesterol. Remnant-C, remnant cholesterol.

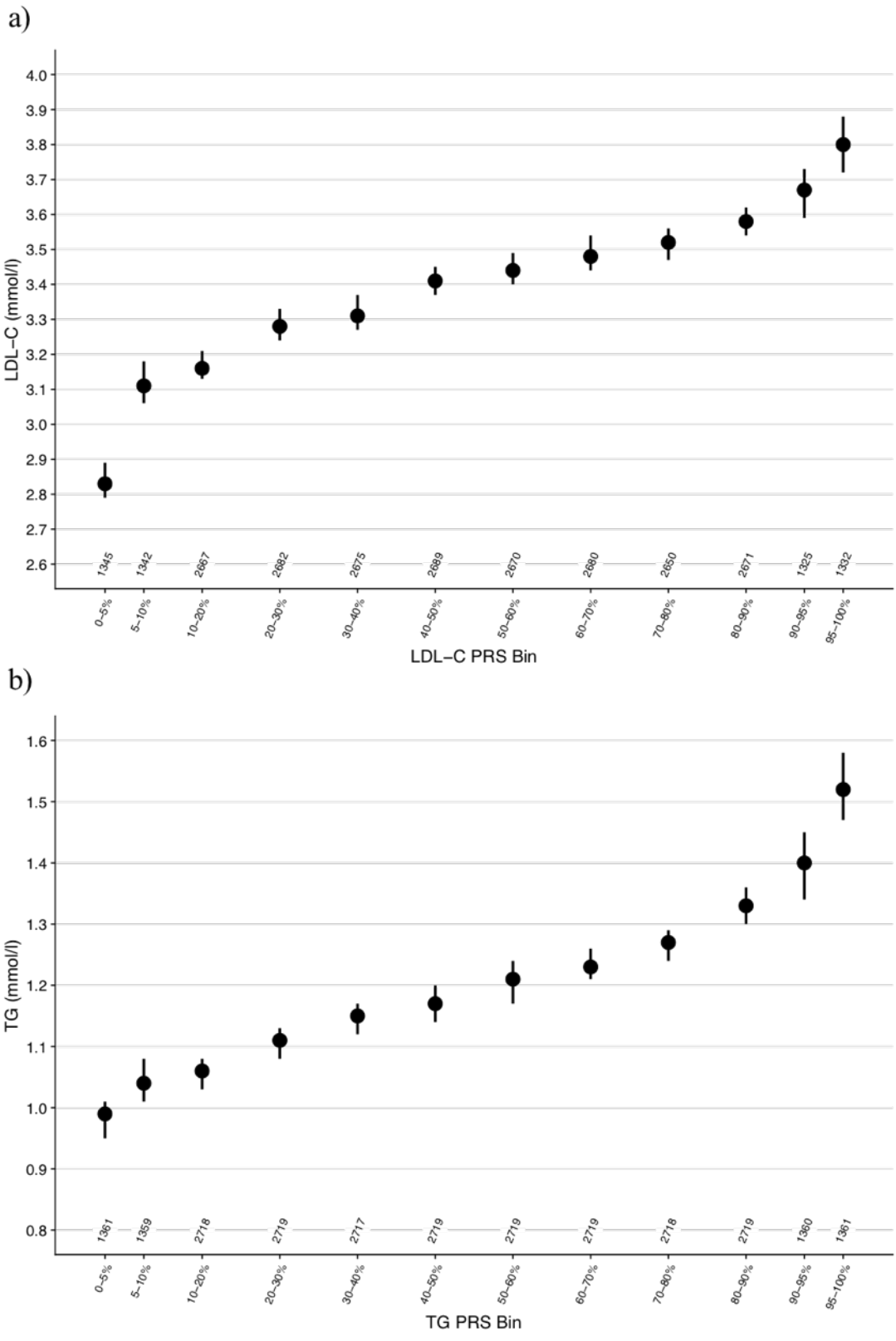
Table 2. CAD Prediction with Lipid and CAD PRSs.

Predictors	AUC	OR (95% CI)	*p*
$PRS_{LDL\text{-}C}$	0.877	1.16 (1.13-1.18)	$< 2\times10^{-16}$
PRS_{TG}	0.877	1.13 (1.10-1.15)	$< 2\times10^{-16}$
PRS_{CAD}	0.881	1.33 (1.30-1.36)	$< 2\times10^{-16}$
$PRS_{CAD} + PRS_{LDL\text{-}C}$	0.881		
PRS_{CAD}		1.32 (1.29-1.34)	$< 2\times10^{-16}$
$PRS_{LDL\text{-}C}$		1.12 (1.10-1.15)	$< 2\times10^{-16}$
$PRS_{CAD} + PRS_{TG}$	0.881		
PRS_{CAD}		1.32 (1.29-1.35)	$< 2\times10^{-16}$
PRS_{TG}		1.08 (1.06-1.11)	5.86×10^{-13}
$PRS_{CAD} + PRS_{LDL\text{-}C} + PRS_{TG}$	0.881		
PRS_{CAD}		1.31 (1.28-1.34)	$< 2\times10^{-16}$
$PRS_{LDL\text{-}C}$		1.10 (1.08-1.13)	$< 2\times10^{-16}$
PRS_{TG}		1.05 (1.03-1.08)	1.28×10^{-5}

ORs and AUCs for CAD with continuous LDL-C, TG, and CAD PRSs as predictors estimated using logistic regression. All models were additionally adjusted for age and sex. AUC, area under the ROC curve. ROC, receiving operating characteristic. OR, odds ratio. CI, confidence interval. CAD, coronary artery disease. PRS, polygenic risk score. LDL-C, LDL-cholesterol. TG, triglycerides.

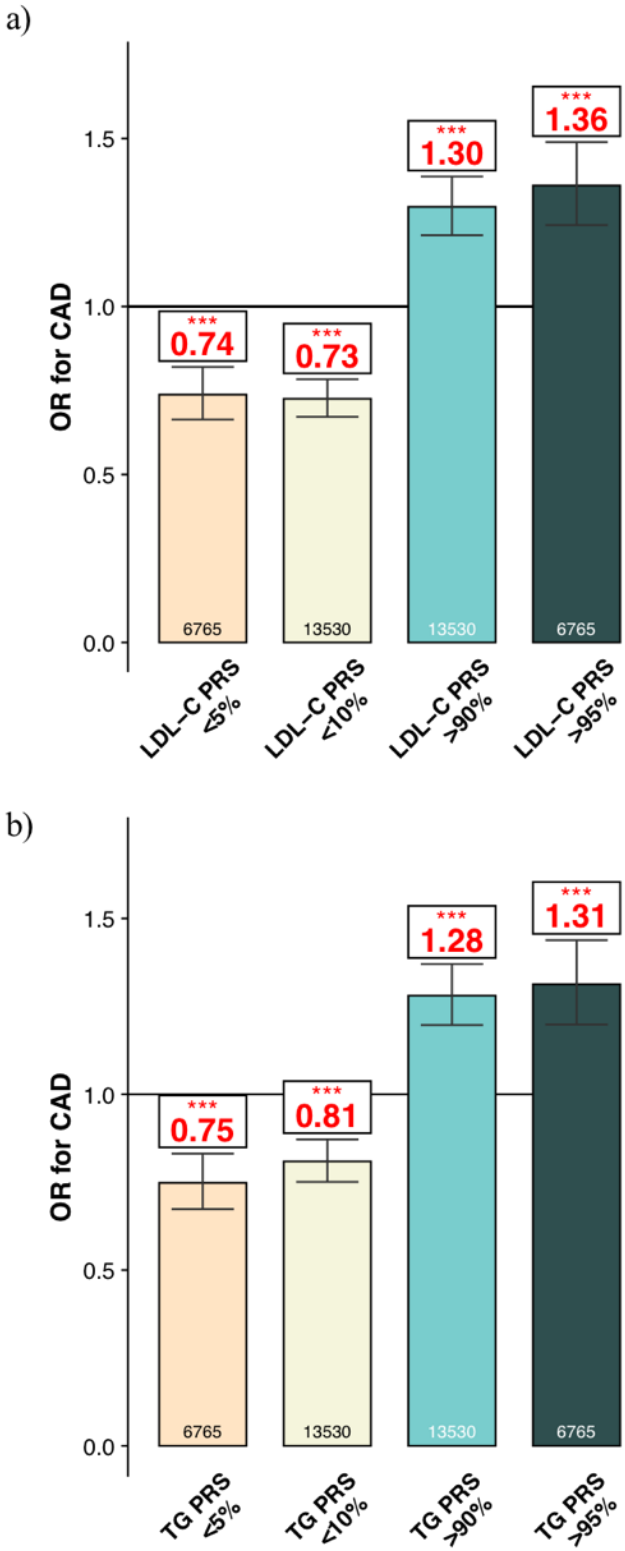
Figures

Figure 1.



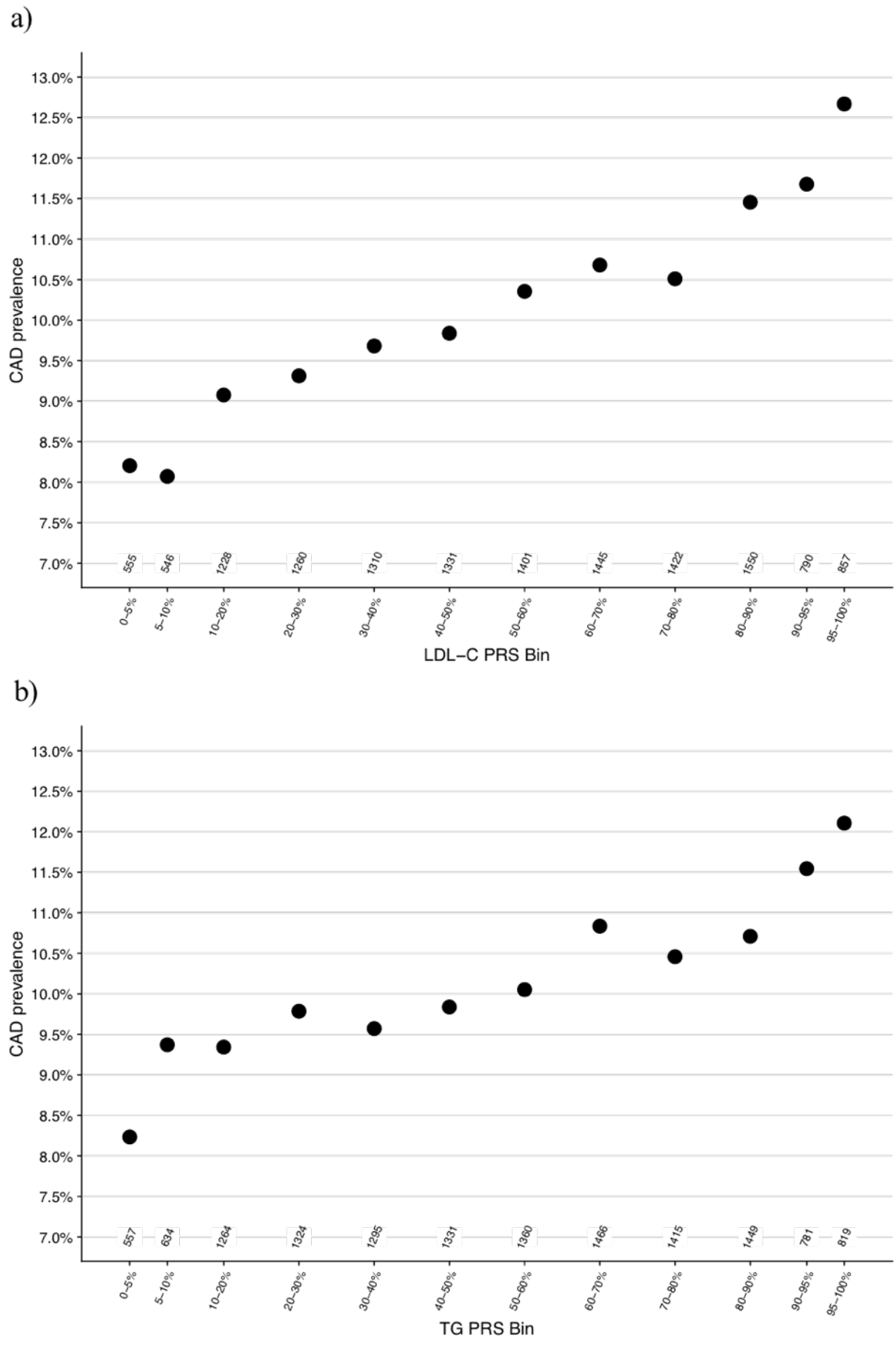
Median LDL-C (a) and TG (b) levels across the distributions of the respective PRSs in the FINRISK cohort. Numbers of individuals in the PRS bins are reported. Vertical lines represent 95% CIs. PRS, polygenic risk score. LDL-C, LDL-cholesterol. TG, triglycerides. CI, confidence interval.

Figure 2.



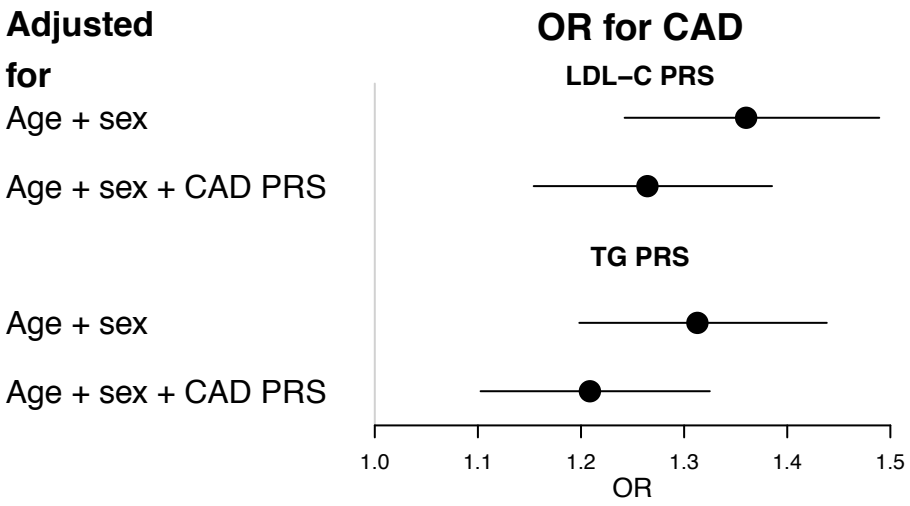
ORs for CAD across the LDL-C (a) and TG (b) PRS distributions in FinnGen. Total numbers of individuals in PRS bins are reported. ORs were estimated using logistic regression. PRS bins were compared with the remainder of the population. Error bars represent 95% CIs. PRS, polygenic risk score. LDL-C, LDL-cholesterol. CAD, coronary artery disease. OR, odds ratio. TG, triglycerides. CI, confidence interval. '$p <$ 0.1. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Figure 3.



CAD prevalence across the LDL-C (a) and TG (b) PRS distributions in FinnGen. Numbers of CAD cases in PRS bins are reported. PRS, polygenic risk score. CAD, coronary artery disease. LDL-C, LDL-cholesterol. TG, triglycerides.

Figure 4.



ORs for CAD for those in the highest 5% of the PRSs compared to the remainder of the population with and without adjusting for the CAD PRS in FinnGen. ORs were estimated using logistic regression. All models were additionally adjusted for age and sex. Horizontal lines represent 95% CIs. OR, odds ratio. CI, confidence interval. CAD, coronary artery disease. PRS, polygenic risk score. LDL-C, LDL-cholesterol. TG, triglycerides.

Appendices

S1 Text. Supplemental Methods.

S1 Figure. LDL-C and TG Levels in FINRISK Surveys from 1992 to 2012.

S2 Figure. HDL-C, TG, and LDL-C levels in FINRISK Across the PRS Distributions.

Supplemental Files

Supplemental Material.

Supplemental_Material.pdf, Ripatti, Polygenic Hyperlipidemias and Coronary Artery Disease Risk.doc