

Cohort information

UK Biobank (UKB) is a cohort of 500,000 individuals from the United Kingdom. All participants gave electronic consent for genotype, phenotype, and electronic health record access, and their information was collected at multiple timepoints¹. Analysis was approved by UCLA under UK Biobank application 33127.

Cardiometabolic Phenotype Ascertainment

Direct LDL, HDL, and triglycerides (respectively, fields 30780, 30760, 30870; mmol/L) for each participant were obtained and converted to mg/dL from mmol/L (LDL & HDL: multiplied by 38.67; triglycerides multiplied by 88.57).² The mean of these measurements taken across multiple visits were used to represent each individual. Patients that were taking statins by the field then had their LDL and triglycerides adjusted to account for their medication use. For LDL, patient's measurement was divided by 0.7 and triglyceride was divided by 0.85.^{3,4}

Body mass index (BMI, kg/m²) was also pulled down for each individual. The maximum BMI recorded across all instances of measurement was used to represent each individual. A BMI greater than 25 kg/m² but less than 30 kg/m² was considered overweight; a BMI greater than or equal to 30 kg/m² was used to identify obese participants.

Glycated hemoglobin or HbA1c (field 30750; mmol/mol) measurements and the "Diabetes diagnosed by doctor" field were used to identify participants with a diagnosis of Type 2 diabetes (T2D). HbA1c was converted from mmol/mol to percentage, and the maximum HbA1c measured across all instances was used to represent each individual. Participants were identified as having T2D if they fulfilled at least one of the following criteria: 1) HbA1c greater than or equal to 6.5%, 2) at least one instance of "Diabetes diagnosed by doctor" marked TRUE. Participants were identified as pre-diabetic if their HbA1c was greater than or equal to 5.7% but less than 6.5%.

Testing for epistasis via PRS

Testing for genetic epistasis occurring between common background genetic variation and monogenic variant carrier status was completed using the model, $y = \sum G \cdot \beta_G + C \cdot \beta_C + \sum G \cdot C \cdot \beta_{G \cdot C} + \epsilon$, where y is the phenotype of interest, G represents common genetic variation and its associated effect β_G on the phenotype of interest, C is an indicator if an individual is carrying a pathogenic variant and β_C is the effect size of the monogenic variant on the phenotype of interest, and $G \cdot C$ is the interaction between common background variation and carrier status (i.e., genetic epistasis) with its associated effect size on the phenotype of interest, $\beta_{G \cdot C}$. Covariates, such as age, sex, and the individual's first 10 genetic PCs are also adjusted for in this model. One method this project employed to test for genetic epistasis was to use PRS as a proxy for $\sum G \cdot \beta_G$, leading to the model $y = \beta_{PRS} \cdot PRS + C \cdot \beta_C + C \cdot PRS \cdot \beta_{C \cdot PRS} + \epsilon$. Age, sex, and the first 10 genetic PCs were adjusted for in this model.

1. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
2. Haney, E. M. *et al.* *Screening for Lipid Disorders in Children and Adolescents.* (2007).
3. Zhao, Z. *et al.* Comparative efficacy and safety of lipid-lowering agents in patients with hypercholesterolemia: A frequentist network meta-analysis. *Medicine* **98**, e14400 (2019).

4. Cholesterol Treatment Trialists' (CTT) Collaboration *et al.* Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet* **376**, 1670–1681 (2010).