**eMERGE Network Proposal for Analysis**

Project/Manuscript Concept Sheet

|  |  |
| --- | --- |
| **Submission Date** | February 2, 2017 |
| **Reference Number** | NT212 |
| **Project Title** | A phenome-wide association study to discover pleiotropic effects of lipid metabolism genes (*LPA*). |
| **Tentative Lead Investigator (first author)** | Ben Satterfield |
| **Tentative Senior Author (last author)** | Iftikhar Kullo |
| **All other authors** | M de Andrade, D Schaid, JC Denny, SJ Hebbring, Wei Qi, L Bastarache, TA Manolio  *Representation from other eMERGE sites and Million Veteran Program* |
| **Sites Involved** | All sites |
| **Background and Significance** | * Variants in several genes implicated in the lipid and lipoprotein metabolism influence susceptibility to atherosclerotic cardiovascular disease. Increasingly, however, it is being recognized that these genes may also have pleiotropic effects. Recent studies demonstrate that *LPA* influence glucose homeostasis, blood pressure, triglyceride-rich lipoprotein metabolism, aortic stenosis, heart failure, and chronic kidney disease suggesting that genetic variants in this gene are likely to have pleiotropic effects. An emerging drug class of medications targeting Lp(a) is undergoing clinical trials, motivating an investigation of associations of *LPA* variations with diverse phenotypes in the electronic health record (EHR). * We propose an agnostic comprehensive scan of the phenome, to test the associations between *LPA* and the entire array of phenotypes in the EHR. |
| **Outline of Project** | Aim I. Perform an agnostic scan of the phenome to identify variant level pleiotropic effects of *LPA.* (eI-III merged imputed genotype data).  Aim II. Replicate phenome-wide significant associations from Aim 1 (additional genotyped datasets from Mayo, Harvard, and Million Veteran Program as well as publically available data from UK Biobank and Finngen).  Aim III. Performing a previously published *LPA* genetic score to evaluate the relationship between Lp(a) levels and ASCVD phenotypes and determine differences based on racial ancestry. |
| **Desired**  **Variables (essential for analysis**  **indicated by \*)** | * Genotyped/imputed variants in *LPA*; Genotyping platform, quality control metrics * Total cholesterol, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, lipoprotein(a) / data on lipid-lowering treatment * Baseline clinical characteristics (age, gender, race, ethnicity, current smoking, hypertension, diabetes, BMI, ASCVD subtypes) * Phenocodes |
| **Desired data** | * All genotyped and imputed data for *LPA* variants * Case-control status via application of PheWAS algorithm |
| **Planned Statistical Analyses** | Common variants will be analyzed for associations using both single point and agglomerative tests.  Each phenotype meeting inclusion/exclusion criteria will be tested for association at the gene and variant levels using weighted adaptive sum of powered score test and logistic regression assuming an additive genetic model adjusted for age, sex, study site and principal components, respectively.  The quality control and data analyses will be conducted using a combination of PLINK and the R statistical package, including PheWAS R package. |
| **Ethical considerations** | None |
| **Target Journal** | * JAMA Cardiology |
| **Milestones** | |  |  |  |  |  | | --- | --- | --- | --- | --- | | **2017** | **2018-2020** | **03-04/2020** |  | **04/2020** | | Obtain information from all sites | Data review and analyses | Draft of manuscript circulated |  | Manuscript submission | |