|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT390 | |
| **Submission Date** | May 1, 2020 | |
| **Project Title** | Understanding the genetic architecture of COPD | |
| **Tentative Lead Investigator** *(first author)* | Victoria Martucci | |
| **Tentative Senior Author**  *(last author)* | Melinda Aldrich | |
| **All Other Authors** | Bradley Richmond, Digna Velez Edwards, Lea Davis | |
| **Sites Participating** | Open to all sites  Current participants:  Vanderbilt | |
| **Background / Significance** | Chronic obstructive pulmonary disease (COPD) is the third leading cause of death worldwide. COPD is characterized by progressive and irreversible decline in lung function. While genome-wide association studies (GWAS) have identified loci associated with COPD, the biological effect of these variants requires further investigation. We currently have access to GWAS summary statistics for several lung function traits from a study of 400,102 individuals published by Shrine et al. (2019). We propose to use these data to develop polygenic risk scores (PRS) in the eMERGE dataset and perform phenome-wide association studies (PheWAS) to examine the relationship between lung function PRS and the medical phenome. We also plan to use eMERGE genetic data to impute gene expression levels with PrediXcan to investigate the role of gene expression in COPD. Findings will be independently validated in the BioVU population at Vanderbilt University Medical Center. | |
| **Outline of Project** | We have GWAS summary statistics for 4 lung function traits that we will use to develop PRS for each phenotype. We will then use each PRS as a predictor in PheWAS analyses to look for associations across the medical phenome. We have already built the PRS and performed an initial PheWAS in the BioVU population. We will use the eMERGE dataset to validate our findings. Phecodes will be used as the outcome in all PheWAS analyses. To examine the relationship between COPD and gene expression, we have performed S-PrediXcan with the results from the lung function GWAS analyses. We will use PrediXcan with the genotyping data from eMERGE to predict gene expression levels across all tissues available in GTEx. We will use our S-PrediXcan results to prioritize genes of interest and look at associations between expression levels of these genes and COPD status. We have already performed PrediXcan in BioVU but expanding to eMERGE will allow us to validate our findings. COPD status will be defined using an algorithm we previously developed. We will stratify by race to examine potential differences across racial groups. We will also perform combined and sex-stratified analyses to investigate whether there are sex-specific effects of PRS and gene expression changes. | |
| **Desired Data - Common Variables\***  *(Available from the CC)* | Demographics  ICD9/10 codes  CPT codes  Phecodes  BMI | Common Variable Labs  Common Variable Meds  Other: Case/Control status on Phase I and Phase II phenotypes |
| **Other Desired Data *(Available from participating sites)*** |  | |
| **Desired Genetic Data** | eMERGE I-III Merged set (HRC imputed, GWAS)  eMERGE PGx/PGRNseq data set  eMERGEseq data set (Phase III)  eMERGE Whole Genome sequencing data set  eMERGE Exome chip data set  eMERGE Whole Exome sequencing data set  Other (not listed above): | |
| **Does project pertain to an existing eMERGE Phenotype?** | Yes, if so please list  No | |
| **Planned Statistical Analyses** | PRS-PheWAS, PrediXcan | |
| **Ethical Considerations** | None | |
| **Target Journal** | Depending on the results of the analyses, we will consider submitting to a genetics journal such as Human Molecular Genetics or a thoracic journal such as Annals of the American Thoracic Society | |
| **Milestones**  *(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | Gather data from coordinating center: April 2020  Conduct analyses: April-August 2020  Write manuscript: August-October 2020  Circulate and submit manuscript: November-December 2020 | |