|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT408 | |
| **Submission Date** | October 15, 2020 | |
| **Project Title** | Genomic architecture of pharmacogenetic traits. | |
| **Tentative Lead Investigator** *(first author)* | Ayesha Muhammad (VUMC) | |
| **Tentative Senior Author**  *(last author)* | Sara Van Driest (VUMC) | |
| **All Other Authors** | Jonathan Mosley, Dan Roden, Ida Aka, Brendan Armstreet, Christian Shaffer, WeiQi Wei | |
| **Sites Participating** | VUMC (performing analyses and writing manuscript)  This analysis uses data from previously published eMERGE II ACE-inhibitor associated cough analysis, which included data from VUMC, Mt Sinai, Marshfield Clinic, Northwestern, University of Washington, Mayo and Geisinger. | |
| **Background / Significance** | The vast majority of clinical pharmacogenetic testing is done using a small number of moderate to large effect size variants in a few genes to predict drug response. The complex nature of most drug metabolism and response pathways conceptually support the use of polygenic predictors, which are largely unexplored. Applying mixed model approaches to existing drug response datasets available on dbGaP and extractable from electronic health records provide the opportunity to assess the genomic architecture of drug response phenotypes and the potential for improved prediction using polygenic approaches. | |
| **Outline of Project** | We have applied Bayesian linear mixed models to drug response datasets obtained from dbGaP, BioVU, the ICPC, and one dataset from eMERGE II (ACE inhibitor associated cough). In all, the drugs studied include clopidogrel, statins, ACE inhibitors, methotrexate, cyclosporine, tacrolimus, vancomycin and gentamicin. Our analyses quantified the SNP-based heritability for these traits and the contributions of small, medium, and large effect size variants. The data will be presented at the 2020 ASHG meeting, and a manuscript has been drafted for submission. | |
| **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  **None required from the CC** |
| **Other Desired Data *(Available from participating sites)*** | *Please specifically list out any data elements that participating sites would collect or extract from clinical or other sources for this project (i.e. not common variables above)*  **None** – all data are in hand | |
| **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):  **None** – all data are in hand | |
| **Does project pertain to an existing eMERGE Phenotype?** | Yes, ACE-inhibitor associated cough (from eMERGE II)  No | |
| **Planned Statistical Analyses** | Bayesian linear mixed models | |
| **Ethical Considerations** | De-identified Individual level data were obtained to perform the analyses, which were stored on secure servers. No individual level data will be published. | |
| **Target Journal** | To be discussed by authors. | |
| **Milestones** | October and November 2020 – Finalization of manuscript  November 2020 –Manuscript Submission to Journal | |

**\*Common Variables available across all datasets:**

* Demographics: sex, year of birth, decade of birth, race, ethnicity
* Codes: (repeated values & age at event): ICD, CPT, Phecodes
* BMI: (repeated value & age at event) height, weight, BMI
* Labs: (lab name, repeated lab value & age at event) Serum total cholesterol, LDL, HDL, Triglycerides, Glucose fasting/non-fasting/unknown, & White Blood Cell count
* Medications: (medication name, repeated, & age at event) Cerivastatin sodium, Rosuvastatin, Simvastatin, Fluvastatin, Pravastatin, Lovastatin, Atorvastatin, & Pitavastatin
* Other: Case/Control status on Phase I and Phase II phenotype: only on GWAS dataset participants