**eMERGE Network Proposal for Analysis**

Project/Manuscript Concept Sheet

|  |  |
| --- | --- |
| **Reference Number** | NT260 |
| **Submission Date** | 10/20/2017 |
| **Project Title** | Association of *CYP2D6* genotypes in the eMERGE PGx Cohort with EHR-derived phenotypes. |
| **Tentative Lead Investigator (first author)** | John Logan Black III MD |
| **Tentative Senior Author (last author)** | Iftikhar Kullo MD |
| **All other authors** | Hugues Sicotte PhD, Sandra Peterson, Laura Rasmussen-Torvik, Cindy Prows, Adam Gordon, Teri Manolio, Josh Denny, Dan Roden and other interested eMERGE investigators |
| **Sites Involved** | All eMERGE PGx sites |
| **Background and Significance** | * CYP2D6 is difficult to genotype from short read NGS data due to a large number of CNV-related conditions and star alleles that are present with variable haplotypes * We have generated a novel approach to determining CNV of CYP2D6 using NGS short read technology based upon an understanding of the CYP2D locus * This novel approach has been encoded into a software solution called CNVAR. The software not only identified CNVs (deletions, duplications, multiplications, hybrid genes) but also provides star allele genotypes * CNVAR is in clinical production at Mayo Clinic |
| **Outline of Project** | Aim I. Determine CYP2D6 genotype of samples in the eMERGE PGx dataset that were sequenced using PGRN\_SEQv1 and estimate allele frequencies, noting novel alleles where possible.  Aim II. Investigate the association of CYP2D metabolizer status with adverse drug response phenotypes and conduct PheWAS of metabolizer status. |
| **Desired**  **Variables (essential for analysis**  **indicated by \*)** | Age, sex, ADR codes, phecodes for PheWAS. |
| **Desired data** | We require and have access to the required sequencing files. |
| **Planned Statistical Analyses** | Simple descriptive analysis of allele frequency and number of samples that cannot be genotyped using CNVAR. Results will have unique identifier, genotype (\*allele), phenotype data (metabolizer status). A genotype to phenotype look up table will be provided. |
| **Ethical considerations** | None |
| **Target Journal** | * TBD |
| **Milestones** | October 2017: Proposal submission  November 2017: Initiate analysis of sequencing files  November 15, 2017: Possible poster submission to Pacific Symposium on Biocomputing (stretch goal)  December 2017: Complete analysis of all samples  January 2018: Return results to eMERGE  March 2018: Manuscript preparation  May 2018: Manuscript submission |