|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT338 | |
| **Submission Date** | 4/4/2019 | |
| **Project Title** | Effect of genomic regulation on Clopidogrel Response in African Americans | |
| **Tentative Lead Investigator** *(first author)* | Tanima De | |
| **Tentative Senior Author**  *(last author)* | Minoli A Perera | |
| **All Other Authors** | Wei-Qi Wei, Joshua C. Denny, Jennifer Allen Pacheco | |
| **Sites Participating** | Northwestern University and other sites with African American cohorts | |
| **Background / Significance** | Clopidogrel, an antiplatelet drug widely used for thromboprophylaxis of cardiovascular diseases (CVDs), is associated with substantial inter-patient differences in response. The clinical consequence of inadequate response is recurrent cardiovascular events (CVEs) that can lead to myocardial infarction, stroke and death. The underlying mechanism is multifactorial and genetic polymorphisms is one of the main causes of variable drug response within an individual or across populations. African Americans (AAs) are disproportionately affected by death and disability from CVDs and are also at a higher risk of CVE and mortality from poor clopidogrel response. Identifying the population-specific genetic predictors for poor clopidolgrel response in AAs may explain this health disparity. In addition, exploring the pleiotropic effects of these genetic predictors across extensive set of phenotypes may identify novel associations with other drug responses and disease conditions which could be of clinical relevance.  Clopidogrel undergoes hepatic cytochrome P450 (CYP)–mediated biotransformation to an active metabolite. Regulatory single nucleotide polymorphisms (SNPs) that significantly influence CYP gene expression (known as expressed quantitative trait loci [eQTLs]), may contribute to the inadequate clopidogrel-induced platelet inhibition, accounting for differences in response. Also, SNPs that act by modifying the DNA methylation status (methylation quantitative trait loci; meQTLs), leading changes in gene expression through altered chromatin accessibility and transcription factor binding may be an additional source of variability. Evaluating the regulatory and epigenetic modulation of gene expression may uncover AA-specific factors that influence drug metabolizing enzyme (DME) expression.  Genotype, methylation and gene expression data derived from 62 AA primary hepatocytes will be used to identify SNPs that are significantly associated with variation in CYP expression. The identified hepatocyte QTLs will be tested on clopidogrel-treated AAs from Electronic Medical Records and Genomics (eMERGE) Network, to detect those SNPs that expose AAs to a higher risk of CVEs from poor response to clopidogrel. All the identified hepatocyte QTLs will also be tested for their association with various other drug and disease phenotypes, by Phenome-wide association study (PheWAS). This will help to unravel the genotype-phenotype relationships, which could be of clinical relevance beyond clopidogrel response. | |
| **Outline of Project** | 1. Genotype, methylation and gene expression data derived from 62 AAs primary hepatocytes will be used to identify CYP-specific QTLs (i.e. eQTLs and meQTLs)  2. The identified QTLs will then be tested on 500 clopidogrel-treated AAs from Electronic Medical Records and Genomics (eMERGE) Network and AA Cardiovascular Pharmacogenomics Consortium (ACCOuNT), to detect those SNPs that expose AAs to a higher risk of CVEs from poor response to clopidogrel  3. The role of the QTLs will be explored for their association with other phenotypes by Phenome-wide association study (PheWAS) | |
| **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)*** | Platelet reactivity unit (Optional covariate) | |
| **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** | **QTL Mapping:** This will be performed with FastQTL using covariates identified through the Probabilistic Estimation of Expression Residuals (PEER) method, first three principal components (PCs), age and gender. A false discovery rate (FDR) of ≤0.05 will be applied to identify genes with a significant QTL (eGenes). To detect the significant variant-gene pairs associated with eGenes, an empirical p-value threshold, pt, defined as the empirical p-value of the gene closest to the 0.05 FDR, will be used to calculate a nominal p-value threshold for each gene based on the beta distribution model (from FastQTL) of the minimum p-value distribution f(pmin) obtained from the permutations for the gene.  **Study the association of the QTLs with clopidogrel response**: Five hundred clopidogrel-treated AAs recruited through eMERGE and ACCOuNT projects will be used to test the effect of QTLs on clopidogrel response, using the available genotype and phenotype information. Subjects will be assigned case-control status depending on the occurrence of a recurrent cardiovascular event within 1 year of clopidogrel treatment. Association analysis will be done using PLINK, after adjusting for clinical covariates significantly associated with response variability.  **Explore the role of the eQTLs for their association with other phenotypes:** Case-control status will be defined for each ICD-9 code. Individual with ≥ 3 instances of an ICD-9 code will be considered as a case and control status is assigned based on the absence of an ICD-9 code. Samples with >0 but <3 ICD-9 code will be removed from analysis for that ICD-9 code. For a Phecode to be used there must be at least 20 cases. A logistic regression analysis will be carried out, to calculate the odds ratios (ORs) and 95% confidence intervals (95% CI), for each PheWAS code, after adjusting for gender, age, genotyping platform and the first three PCs. Analyses will be performed and graphed using the PheWAS package in R version 3.2.5. | |
| **Ethical Considerations** | This is a retrospective study with no direct interaction with subjects. There will be no loss of confidentiality as the data will be stored in a secured environment and password protected and encrypted computer maintained per NU policy that only the investigators will be able to access. There are no direct benefits to the patients for their participation in this study. However, our research findings may benefit other patients | |
| **Target Journal** | The Lancet | |
| **Milestones**  *(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | Project Approval – March 2019  Project duration – April 2019 to December 2019  Manuscript preparation first draft – February 20th, 2020  Final draft of manuscript– March 20th, 2020  Manuscript submission to journal- April 15th, 2020 | |

**\*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