|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT321 | |
| **Submission Date** | 1/24/2019 | |
| **Project Title** | Pleiotropic Associations Between a Uterine Leiomyoma Polygenic Risk Score (PRS) and the Clinical Phenome | |
| **Tentative Lead Investigator** *(first author)* | Jacklyn N. Hellwege | |
| **Tentative Senior Author**  *(last author)* | Digna R. Velez Edwards | |
| **All Other Authors** | Eric S. Torstenson  Todd L. Edwards  Brian Mautz  Sarah Jones  Josh C. Denny  Dan M. Roden  Sarah Pendergrass  Yanfei Zhang | |
| **Sites Participating** | Open to all sites  Current participants:  Vanderbilt  Geisinger | |
| **Background / Significance** | Uterine leiomyoma, or fibroids, are the most common female pelvic tumor, with prevalence up to 77% by menopause. Fibroids are benign smooth muscle growths which share fibroproliferative features with many physiological conditions, including keloids and nephrosclerosis. The genetic architecture of fibroids currently includes ~30 susceptibility loci across populations. We combined fibroid polygenic risk scores (PRS) with a phenome-wide association study (PheWAS) approach to gain understanding about the shared genetic contribution across many clinical phenotypes. | |
| **Outline of Project** | We constructed PRS with PRSice software using effect sizes derived from imaging-confirmed fibroids genome-wide association (GWAS) data from the Electronic Medical Records and Genomics (eMERGE) network, with optimization in additional imaging-confirmed cases and controls from the BioVU Repository. PRS were built in white women from BioVU using racially-consistent GWAS results.  We will perform PheWAS of the PRS stratified by race and sex in the eMERGE data omitting the samples used to develop the PRS, adjusted for age, body mass index, and principal components of ancestry. We will also use meta-analysis to calculate estimates of effects and significance across races and sexes. We are also utilizing the MEGA array data at VUMC (currently ~85K), and the array data for non-eMERGE participants at Geisinger (~80k) for this analysis.  We will summarize the results of the association tests, with secondary analysis of Phecode groupings (hypergeometric tests for overrepresentation and sign tests for trends in effect sizes), as well as network analysis of results. | |
| **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 Uterine  No | |
| **Planned Statistical Analyses** | PheWAS | |
| **Ethical Considerations** | None | |
| **Target Journal** | Depends on results, likely a genetics journal such as Human Molecular Genetics or a general journal such as elife, Nature Communications, or Scientific Reports | |
| **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: 3/1/2019  Conduct statistical analyses: 3-6/2019  Write manuscript: 6-9/2019  Circulate and submit manuscript: 9-10/2019 | |