|  |  |
| --- | --- |
| **External Collaborator Proposal** *for* **eMERGE Network Analysis**  Project/Manuscript Concept Sheet | |
| **Reference Number** | NT453 |
| **Submission Date** | 7/3/22 |
| **Tentative Lead Investigator** *(first author with contact information and affiliation)* | Yogasudha Veturi, post-doctoral fellow, Department of Genetics, University of Pennsylvania [yveturi@pennmedicine.upenn.edu](mailto:yveturi@pennmedicine.upenn.edu) |
| **Tentative Senior Author**  *(last author)* | Marylyn Ritchie, [marylyn@pennmedicine.upenn.edu](mailto:marylyn@pennmedicine.upenn.edu) |
| **eMERGE Site Sponsor & Contact** | University of Pennsylvania, Marylyn Ritchie |
| **Project Title** | Identifying shared genetics between neuroimaging data and diseases in the EHR |
| **All Other Authors** | Christos Davatzikos |
| **Other eMERGE Sites Involved** | All emerge network |
| **Background / Significance** | There is evidence to suggest that genes associated with neurodegeneration are also associated with diseases of multiple organ systems, especially for cardiometabolic diseases such as coronary artery disease and type-II diabetes. From our own preliminary analyses, we have observed genetic overlaps between neuroimaging and diseases across multiple organ systems such as asthma and some neoplasms. |
| **Outline of Project** | In order to obtain a better idea of the extent of genetic overlap between neuroimaging phenotypes and diseases in the EHR, our goal is to conduct neuroimaging-guided PheWAS and gene-expression based PheWAS analyses. This will entail conducting these analyses only on the SNP subset that is significantly associated with brain imaging data based on an imaging GWAS conducted in an independent cohort like the UK Biobank. |
| **Desired Variables**  *(essential for analysis*  *indicated by* ***\*****)* | * Primary phenotypes: ICD-9 and ICD-10 disease codes * Confounding variables: age, sex, and race/ethnicity, adult/child status * Related phenotypes: systolic/diastolic blood pressure, body-mass-index, smoking status, drinking status, medications taken |
| **Desired Data** | eMERGE-III HRC imputed data |
| **Planned Statistical Analyses** | Conduct a phenome-wide association study (PheWAS) and transcriptome-wide association study (based on gene-expression weights from Genotype Tissue Expression project v8) across all ICD codes on a neuroimaging-associated SNP subset. |
| **Ethical Considerations** | None |
| **Available Funding or Resources** | eMERGE |
| **Milestones\*\*** | 1. Submit manuscript by Jan 2022. |

***\*\**** *This section should include the timeline for completion of project, including: approval, project duration, first and second draft of the paper and submission.*