|  |
| --- |
| **eMERGE Network: External Collaborator Manuscript Concept Sheet** |
| **Reference Number** *(to be assigned by CC)* | NT324 |
| **Submission Date** | January 9, 2019 |
| **Project Title** | Replication of Million Veterans Program GWAS on Abdominal Aortic Aneurysm  |
| **Tentative Lead Investigator** *(first author)* | Derek Klarin |
| **Tentative Senior Author** *(last author)* | Phil Tsao, Scott Damrauer |
| **eMERGE Site Sponsor & Contact** | Marylyn Ritchie and Shefali Verma |
| **All Other Authors**  | Shefali Setia Verma, Marylyn Ritchie, Yuki Bradford, And other eMERGE AuthorsRenae Judy, Cuiping Pan, Themistocles L. Assimes , Bill Boden, Kyong-Mi Chang, Kelly Cho, Scott Duvall , Jie Huang, Sekar Kathiresan, Jennifer Lee, Julie Lynch, Donald Miller, Christopher J. O’Donnell, Daniel J. Rader, Yan V. Sun, Peter W.F. Wilson |
| **Sites Participating** | All |
| **Background / Significance** | Aside from occlusive disease, aneurysmal dilatation of the peripheral vasculature is the leading form of peripheral artery pathology. This form of arterial disease most commonly affects the infrarenal abdominal aorta (AAA). There are many shared cardiovascular risk factors with AAA and also risk for conditions like diabetes mellitus are negatively correlated. Understanding genetic etiology of AAA could help in elucidating differences among correlated conditions. Capitalizing on the polygenic nature of AAA association studies could also help in developing risk scores for samples. |
| **Outline of Project** | MVP has performed a GWAS study where cases are defined by 2 or more occurrences of AAA ICD codes. In eMERGE, we will be replicating their top associations (60 SNPs only) by performing logistic regression analyses where cases and controls will be defined in same way. Additional analyses for may include:1. PheWAS of top novel AAA associated SNPs with other ICD codes to identify potentially pleiotropic associations and comorbidities.
2. Risk prediction in eMERGE population.
3. Global GWAS meta-analysis of AAA

MVP is trying to assemble a larger group to do a large meta-analysis. Depending on the results of replication in eMERGE and the success of finding additional partners, that will determine if this is one replication analysis only or a full GWAS meta-analysis. |
| **Desired Data - Common Variables\*** *(Available from the CC)* |  **Demographics**  **ICD9/10 codes** |  |
| **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 |
| **Desired Genetic Data** |  **eMERGE I-III Merged set (HRC imputed, GWAS)** |
| **Does project pertain to an existing eMERGE Phenotype?** | **Yes, if so please list (AAA) However, we will be defining AAA case status with ICD 9/10 codes only as described below.**  |
| **Planned Statistical Analyses** | 1. Defining case/control based on following criteria:
* The controls are no occurrences of the above ICD codes AND no occurrences of any codes in the 440-448 or 557 range in ICD9 or I71-I75, I77-I79, K55.
* The case definition is ICD 9/10 (441.3, 441.4, I71.3, I71.4) on at least two distinct dates
1. Extracting chr:pos for SNPs to replicate from MVP study (60 SNPs)
2. GWAS using logistic regression in PLATO, adjusting for year of birth, sex, 6 PCs and eMERGE Site
3. For novel replicating associations, we will conduct a PheWAS with other ICD codes.
4. Polygenic risk score (PRS) to predict risk on samples. This step would include applying genomic risk scores calculated on MVP datasets to validate the PRS in eMERGE population.
5. There are efforts for generating a large AAA dataset which would initiate a global meta-analysis for AAA. In conducting such analysis, we would run a GWAS on all imputed SNPs where the case/control definition will remain the same (based on ICD codes as explained in #1).
 |
| **Ethical Considerations** | None |
| **Available Funding or Resources** |  Dr. Ritchie start-up funding. |
| **Target Journal** |  |
| **Milestones***(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | 1. Perform analyses in February 2019
2. Submit results to MVP in March 2019
 |

**\*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