|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT348 | |
| **Submission Date** | 6/12/2019 | |
| **Project Title** | Mendelian randomization study of the association of lipid risk SNPs with the development of breast, prostate and other cancers | |
| **Tentative Lead Investigator** *(first author)* | Daniel Lee MD (Dept of Urology) and Shefali S. Verma | |
| **Tentative Senior Author**  *(last author)* | Kara Maxwell MD PhD | |
| **All Other Authors** | Benjamin Voight and colleagues  Marlyn Ritchie and Shefali S. Verma  Bioinformatician: Heena Desai in the Maxwell Lab  Other participating eMERGE sites and co-authors | |
| **Sites Participating** | TBD | |
| **Background / Significance** | A number of epidemiological studies have attempted to determine whether levels of circulating lipids are associated with risks of various cancers, including breast cancer. If alteration in lipid levels also reduced risk of breast cancer, this could present a target for disease prevention. We have previously demonstrated significant associations for HDL SNPs with breast cancer in the UK Biobank. | |
| **Outline of Project** | Genotypes of known lipid SNPs will be tested by a Mendelian randomization model for association with breast, prostate and other cancers. | |
| **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)*** | *None* | |
| **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 Common variables as listed above are already available for sites.  No | |
| **Planned Statistical Analyses** | Mendelian randomization mode | |
| **Ethical Considerations** | None | |
| **Target Journal** | Nature Genetics | |
| **Milestones**  *(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | Analyses completion: August 2019  Manuscript Draft: October 2019  Manuscript Submission: November 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