|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT320 | |
| **Submission Date** | 1/25/2019 | |
| **Project Title** | Identify potential risk genes for common calcium kidney stones | |
| **Tentative Lead Investigator** *(first author)* | Krzysztof Kiryluk MD, MS | |
| **Tentative Senior Author**  *(last author)* | John Lieske, MD | |
| **All Other Authors** | Peter C. Harris, PhD, Ali Gharavi, MD, Ozan Dikilitas, MD, Daniel Schaid, PhD, Iftikhar Kullo, MD | |
| **Sites Participating** | Any other interested eMERGE sites | |
| **Background / Significance** | Kidney stones are common affecting approximately 10% of the population. The majority (~80%) are majority calcium oxalate stones. There are clear familial patterns of inheritance and risk. We have developed an algorithm that uses a combination of ICD 9/10 and CPT codes to identify common calcium stones and exclude secondary causes. After 3 iterations the most recent version was 94% sensitive and 90% specific in the Mayo Clinic Electronic Heath Record System when confirmed by manual chart review.  Thus this algorithm can be used to identify kidney stone cases and controls using data universally available from the electronic health record. eMERGE data will help to provide a cohort of kidney stone cases and controls to identify kidney stone risk genes using a GWAS approach. | |
| **Outline of Project** | We will apply this definition to identify cases and controls for GWAS studies.   1. Validate the electronic definition of calcium kidney stones and controls in collaboration with the Columbia University team*.* 2. Identify calcium kidney stone cases in the Columbia University and Mayo Clinic eMERGE cohorts, together with other interested eMERGE collaborators. 3. In a case control design we will identify genetic regions associated with kidney stone risk using the eMERGE Phase III v2 GWAS data sets. In these analyses we will adjust for sex, age and weight (BMI), all key demographics that associate with kidney stone risk. 4. In future studies, we will follow up on any implicated genetic regions, and identify candidate genes. | |
| **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)*** | *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)* | |
| **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 breast cancer, cancer  No | |
| **Planned Statistical Analyses** | 1. GWAS to identify genetic regions associated with kidney stone risk | |
| **Ethical Considerations** | Genomics data and phenotypic data will be de-identified to protect confidentiality. | |
| **Target Journal** | Kidney International | |
| **Milestones**  *(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | 1. Complete validation of electronic definition of a kidney stone with Columbia University by March Dec 2019 2. Obtain kidney stone phenotype data of eMERGE subjects by June 2019 3. Statistical analysis by September 2019 4. Manuscript by December 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 for breast cancer
* 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