|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT442 | |
| **Submission Date** | February 10, 2022 | |
| **Project Title** | Returning integrated genomic risk assessments to providers and participants: the eMERGE IV study. | |
| **Tentative Lead Investigator** *(first author)* | Jodell E. Linder (Jackson) | |
| **Tentative Lead Investigator Email Address** | Jodell.jackson@vumc.org | |
| **Tentative Senior Author**  *(last author)* | Josh F. Peterson | |
| **All Other Authors** | Leadership, PIs, co-chairs; Niall Lennon, Maegan Harden, Lori Orlando, Tejinder Rakhra-Burris, Sophie Forman, Sofia Labrecque, Lynn Seabolt, Alanna DiVietro, etc. | |
| **Sites Participating** | eIV sites | |
| **Background / Significance** | eMERGE has launched a fourth phase of the network, focused on estimating risk of common, complex disease using a report of integrated genomic risk. This paper will describe the study goals, the Network structure, the implementation framework, and early challenges encountered prior to launching prospective recruitment. This paper can be used as a reference to downstream network wide papers on the cohort. | |
| **Outline of Project** | Introduction:   * Overarching emerge goal is to study the implementation of novel genomic interventions through use of scalable technologies * Components of genomic risk for common complex disease * Phase IV aims and rationale for implementation study * Focus on underrepresented, large age range, inclusiveness in common complex disease. Importance to advancing field of genomics (and risk prediction).   Methods:   * Logistics for organizing network (workgroups/contracts/duas/trackers/dashboards) * sIRB & central redcap data repositories (not in depth about R4, will be separate paper). * Prospective genotyping workflow * Partnership with Invitae (Tier 1 + 5), Broad (GDA array), Duke & complexities with CLIA labs & industry partners   Results:   * Elements leading to launch of study - brief descriptions of goals of each workgroup, early ELSI input (set up for ELSI paper). * Brief section on condition selection (logistics) touching on PRS validation (not in depth; set up for PRS paper) and final list of conditions. * Brief estimations of high risk/overall sample sizes of prospective cohort (impact the project will have) * General study/protocol design (but leaving room for downstream recruitment outcomes paper) * Main outcomes to be studied (mirror clin trial.gov submission; set up outcomes studies)   Discussion:   * Improvements in PRS performance and overall risk assessment in the field of genomics * ELSI components of risk assessment and underrepresented participants in genomic medicine * Implications on clinical implementation and next steps for the field | |
| **Desired Data - Common Variables\***  *(Available from the CC)* | Demographics  ICD9/10 codes  CPT codes  Phecodes  BMI | Common Variable Labs  Common Variable Meds  Geocoding 2015 ACS variables  Other: Case/Control status |
| **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  No | |
| **Planned Statistical Analyses** | None | |
| **Ethical Considerations** | None | |
| **Target Journal** | GIM? ACMG? AJHG? | |
| **Milestones**  *(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | February 2022 MCS approved  March 2022 Outline to authors  April 2022 Draft to Authors  May 2022 Submission to journal | |

**\*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