|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT466 | |
| **Submission Date** | 9 Nov2022 | |
| **Project Title** | **Evaluation and return of polygenic risk scores in a diverse population: Design of the eMERGE-4 study** | |
| **Tentative Lead Investigator** *(first author)* | Nita Limdi……………… Noura Abul-Husn, David Veenstra | |
| **Tentative Lead Investigator Email Address** | nlimdi@uabmc.edu | |
| **Tentative Senior Author**  *(last author)* | Noura Abul-Husn, David Veenstra | |
| **All Other Authors** | PI’s, Provider uptake and outcomes group, CC leadership, NHGRI –eMERGE leadership, Invitae, others | |
| **Sites Participating** | All emerge IV sites, NIH, Invitae (others) | |
| **Background / Significance** | Genome-wide association studies (***GWAS***) have enabled the development of polygenic risk scores **(*PRS*)** that can identify patients at higher risk of common diseases, and in concert with clinical, demographic, and family history information (genomic-informed risk assessment, **GIRA**) provide a novel approach for risk-based disease screening and prevention of future disease.  Although development and validation of PRSs is a highly active area of research, implementation and evaluation of their use in clinical care, particularly in diverse patient populations, has been limited.  NHGRI’s eMERGE IV network was created to take this vital first step to leverage the power of genomics to prevent disease. To this end, the network aimed to:   1. Develop and validate polygenic risk scores for multiple complex diseases across racial groups 2. Recruit 25,000 diverse adults and children 3. Conduct genomic-informed risk assessment for 10 common diseases and develop treatment recommendations 4. Communicate GIRA and treatment recommendations to patient and physician 5. Assess incremental uptake of risk-reduction recommendations, incidence of new disease diagnosis, and explore impact on related clinical outcomes | |
| **Outline of Project** | The purpose of this manuscript is to describe….   1. Introduction to e4 aims and design 2. Health recommendations (CARE group + Phenotype leads) 3. Data collection (Surveys, EMR) 4. Challenges and iterative refinement of study design 5. Outcomes overall and by phenotype 6. Analysis framework – Regression discontinuity design 7. Detectable differences at 80% power 25k pts | |
| **Desired Data - Common Variables\***  *(Available from the CC)* | N/A |
| **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): **N/A** | |
| **Does project pertain to an existing eMERGE Phenotype?** | Yes, if so please list:  No | |
| **Planned Statistical Analyses** |  | |
| **Ethical Considerations** | N/A | |
| **Target Journal** | AJHG, Genetics in Medicine….please add | |
| **Milestones**  *(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | * Nov 2022 manuscript concept sheet approval * Jan 2023 finalize sections/ outline * Jan 2023 - writing assignment/ groups * March 2023- draft manuscript * May 2023- complete manuscript * May 2023 – submit manuscript | |

**\*Common Variables available across all datasets:**

* Demographics: sex, year of birth, decade of birth, race, ethnicity
* Codes: (repeated values & age at event): ICD9/10, 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, Blood pressure medications
* Other: Case/Control status on Phase I and Phase II phenotype: only on GWAS dataset participants