|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT410 | |
| **Submission Date** | 10/27/2020 | |
| **Project Title** | A systematic evaluation of a polygenic risk score for ischemic stroke across diverse populations and different geographic areas | |
| **Tentative Lead Investigator** *(first author)* | Thevaa Chandereng (tc3123@columbia.edu) | |
| **Tentative Senior Author**  *(last author)* | Mitchell S. V. Elkind (mse13@cumc.columbia.edu) | |
| **All Other Authors** | Columbia: Chunhua Weng, Wendy Chung, Krzysztof Kiryluk, Cong Liu, George Hripcsak, Phyllis Thangaraj, Ken Cheung  UAB: Marguerite Ryan (Ryan) Irvin, Nita Limdi, Hemant Tiwari  Other eMERGE investigators interested in the project | |
| **Sites Participating** | We propose a network-wide study (all sites are invited to participate). The analyses will be led by Columbia University, and supported by the University of Alabama-Birmingham. | |
| **Background / Significance** | An ischemic stroke meta-Genomic Risk Score (metaGRS, hereafter referred to as a metaPRS) including 19 different GRS for stroke (n=5) and stroke risk factors and comorbidities (n=14) was validated in the UK Biobank among White participants (n= 407,388).[[1]](#endnote-1) Although this score is well-calibrated and validated for European ancestry individuals, it is unknown how this score performs in individuals of non-European ancestry and how generalizable this PRS is in North Americans across different geographic areas. In the US, stroke incidence among Black individuals, compared to that of the White population, is approximately 50% higher to 4 times as high for those under 85 years of age.[[2]](#endnote-2),[[3]](#endnote-3),[[4]](#endnote-4),[[5]](#endnote-5) Hispanic/Latino Americans are also at increased risk of ischemic stroke compared to White populations.4,5 eMERGE provides a potentially valuable data source to evaluate the transferability of PRS to other ancestries, settings, and outcomes. By testing a previously developed PRS in the eMERGE dataset, we will be able to examine (1) whether a PRS for stroke well-optimized for European ancestry can be generalized to other populations; (2) whether a PRS optimized in one geographic area can be applied directly to other geographic areas; (3) how much additional information is gained from the PRS beyond routinely assessed risk factors for stroke, such as hypertension and diabetes. | |
| **Outline of Project** | 1. Calculate and validate the previously developed metaPRS for stroke (Abraham et al, Khera et al. Nat Communication 20191) in the REGARDS cohort, including 10,000 Black participants age >18 years old. 2. Compare specificity/sensitivity among UK Biobank and REGARDS participants. 3. Create and calculate a modified metaPRS for stroke incorporating the results from these analyses to be validated in the full eMERGE dataset. 4. Compare the new metaPRS performance in European ancestry population and non-European ancestry populations based on self-reported demographics and PCAs. 5. Optimize the metaPRS for one eMERGE site and test its performance across different sites and evaluate the transferability of metaPRS. 6. Validate the metaPRS using Mendelian Randomization (MR). | |
| **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)*   * self-reported demographics * family history when available | |
| **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: ischemic stroke  No | |
| **Planned Statistical Analyses** | 1. Examine ROC AUC for the previously developed Uk Biobank metaGRS for REGARDS participants (n=10,000). 2. Examine ROC AUC for the newly developed metaPRS for eMERGE participants. 3. Compare ROC AUC and incremental change in AUC for the metaPRS between different ancestral populations. 4. Compare ROC AUC for the metaPRS between different geographic areas. 5. Validate the metaPRS using both median-based and MR-Egger regression methods. 6. Quantify the gain in predictive performance from adding metaPRS to routine clinical risk assessment based on reclassification indices. | |
| **Ethical Considerations** | NONE | |
| **Target Journal** | AJHG, PLoS GENETICS, Circulation, Stroke, or similar | |
| **Milestones**  *(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | * 11/2020: Develop initial ischemic stroke metaGRS calculation for REGARDS data; * 12/2020: Evaluate the sensitivity, specificity and other characteristics of metaGRS of ischemic stroke for REGARDS; * 02/2021: Create and validate ischemic stroke metaPRS for eMERGE; * 03/2021: Determine gain in predictive performance for metaPRS over established risk factors for eMERGE participants; * 04/2021: Manuscript draft completion; * 05/2021: manuscript draft submission. | |

**\*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

Reference

1. Abraham G, Malik R, Yonova-Doing E, et al. Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. *Nat Commun.* 12 2019;10(1):5819. [↑](#endnote-ref-1)
2. Koton S, Sang Y, Schneider ALC, Rosamond WD, Gottesman RF, Coresh J. Trends in Stroke Incidence Rates in Older US Adults: An Update From the Atherosclerosis Risk in Communities (ARIC) Cohort Study. *JAMA Neurol*. 2019;77:109-113. [↑](#endnote-ref-2)
3. Howard VJ, Kleindorfer DO, Judd SE, McClure LA, Safford MM, Rhodes JD, Cushman M, Moy CS, Soliman EZ, Kissela BM, et al. Disparities in stroke incidence contributing to disparities in stroke mortality. *Ann Neurol*. 2011;69:619-27. [↑](#endnote-ref-3)
4. White H, Boden-Albala B, Wang C, Elkind MSV, Rundek T, Wright CB, Sacco RL. Ischemic stroke subtype incidence among whites, blacks, and Hispanics: the Northern Manhattan Study. *Circulation*. 2005;111:1327. [↑](#endnote-ref-4)
5. Gardener H, Sacco RL, Rundek T, Battistella V, Cheung YK, Elkind MSV. Race and Ethnic Disparities in Stroke Incidence in the Northern Manhattan Study. *Stroke*. 2020;51:1064-1069. [↑](#endnote-ref-5)