**eMERGE Network Proposal for Analysis**

Project/Manuscript Concept Sheet

|  |  |
| --- | --- |
| **Submission Date** | 4/3/2018 |
| **Project Title** | **GWAS-PheWAS Approach to Infection-Associated Stroke** |
| **Tentative Lead Investigator (first author)** | Neal Parikh (stroke and neuro-epidemiology fellow) |
| **Tentative Senior Author (last author)** | Mitchell Elkind |
| **All other authors**  | Krzysztof Kiryluk, Amelia Boehme, George Hripscak, Chunhua Weng, Ning Shang, Atlas Khan ***additional eMERGE authors interested in participating*** |
| **Sites Involved** | We propose a network-wide study (all sites invited to participate). The analyses will be led by Columbia University. |
| **Background / Significance** | Acute infections of many types have been identified as a trigger of ischemic stroke. For example, in the Cardiovascular Health Study, hospitalization for acute infection was associated with short-term increased stroke risk (Elkind et al, *Stroke*, 2011). Similar results were recently published from ARIC (Cowan et al, *Stroke*, 2016). Among children, acute infection, particularly herpes virus infections, were recently reported to be associated with acute stroke risk (Fullerton et al, *Neurology,* 2015; Elkind et al, *Circulation*, 2016). Influenza, other respiratory infections, urinary tract infections, sepsis, and other specific infections have been associated with short-term stroke risk (Smeeth L et al, *NEJM*, 2004; Boehme A et al., *Stroke*, 2017). While these data provide evidence of associations with stroke risk, there is little known about which individuals who experience acute infection are at risk of developing stroke. The absolute risk of stroke after an infection is low, but the risk may be higher in those with genetic or environmental risk factors. Given well-reported associations of innate immunity with risk of cardiovascular disease and stroke, and evidence of other associations of immune response genes with cardiovascular disease, it is plausible that individuals who are genetically “primed” may be at highest risk of infection-triggered stroke. For example, polymorphisms in the toll like receptor 4 gene, which encodes the protein important in innate response to bacterial pathogens, is associated with increased progression of carotid atherosclerosis (Kiechl et al, *NEJM,* 2002). The potential to identify high-risk individuals prior to or at the time of an infectious event would allow precision medicine approaches to minimizing their risk of subsequent stroke and other cardiovascular events.Here, we propose to create a novel group of ‘combined’ phenotypes (“infection-triggered ischemic stroke” or ITIS) derived from multiple independent phenotypes: each combined phenotype will include an *infection phenotype* (the trigger) and an *ischemic stroke phenotype* (the outcome). These ITIS phenotypes would be created by linking the individual phenotypes in the EMR; each one would have several individual sub-phenotypes depending on the time interval between the two events (15 days, 16-30, 31-45 days, 46-60 days, etc.) We would begin with the following three ITIS phenotypes:1. Influenza-triggered ischemic stroke;
2. Sepsis-triggered ischemic stroke;
3. Bacterial respiratory infection-triggered stroke.

Once the phenotypes are created and validated, we propose a stepwise GWAS-PheWAS approach that combines GWAS for ITIS phenotypes with follow-up candidate PheWAS to define pleiotropic disease associations (i.e., other outcomes including hemorrhagic stroke, MI, sudden cardiac death, etc.). We hope that this approach will offer a powerful new method for detection of novel susceptibility loci that are associated with outcomes across a broad spectrum of infectious and cardiovascular diseases. There is innovation both in the concept of the infection-triggered stroke as well as in the methodological aspects of linking two individual phenotypes using time as the criterion; this would actually create a large family of phenotypes, based on three elements: the exposure (infection), the outcome (stroke), and the time interval between them. Our phenotypes will be written fully for OMOP to allow for efficient data extraction and implementation across the entire eMERGE network. |
| **Outline of Project** | The project will be conducted in several stages:1. Building Columbia EMERGE multi-infection and stroke phenotypes and sub-phenotypes. Please note that multi-infection phenotype will be constructed under a separate concept sheet.
2. Validate the EMR phenotypes.
3. Implementation of phenotype algorithms for all individuals with available GWAS datasets network-wide
4. Phenotype quality control analyses
5. Genome-wide association analyses
6. Targeted PheWAS to discover pleiotropic associations for significant loci
7. Manuscript preparation and submission
 |
| **Desired****Variables (essential for analysis****indicated by \*)** | Implementation of standardized Columbia EMERGE phenotypes related to infection and stroke across all sites with available GWAS datasets: * Columbia multi-infection phenotype
* Columbia Ischemic and Hemorrhagic stroke phenotype
* Dates or time intervals between these events will be required to create the composite ITIS phenotypes
* Age, sex, race/ethnicity\*
* ICD9/ICD10 codes for PheWAS of significant loci for the phenotypes defined as above

N.B.: No extra work for creation of stroke phenotype since we are committed to creating it already. The Columbia multi-infection phenotype will be submitted as a separate concept sheet, but implementation of this phenotype will be required for this proposal. |
| **Desired data** | * Standardized multi-infection/ and stroke phenotypes as above
* All genotype data from EMERGE sites
* Imputed genotype data for all sites.
 |
| **Planned Statistical Analyses** | Primary and confirmatory case-control GWAS for ITIS phenotypesCases: ITIS phenotypesControls:  Primary: specific infection phenotype without any other infection or any stroke phenotype Confirmatory: neither infection OR stroke phenotypeSecondary analyses will include conditional analyses, haplotype analyses and rare-variant association scans. Data from individual centers will be meta-analyzed genome-wide using standard approaches. These analyses will be performed in the Elkind Lab at Columbia University.Targeted PheWAS analysis will be performed for cardiovascular outcomes on a small number of genome-wide significant loci to better define their disease associations and potential pleiotropic effects. The summary statistics will be combined by meta-analysis across all EMERGE sites. |
| **Ethical considerations** | There are no additional risks involved. The EMR and genomic data will be stored at a secured location in the data storage system at Columbia. No data will be shared with unauthorized third parties. Patient identity will not be compromised by the proposed analysis. We will also abide by the EMERGE guidelines in this regard. |
| **Target Journal** | NEJM, Circulation |
| **Milestones\*\*** | Total Duration of the study: 2 years Completion of study design/approvals: May 2018Implementation of phenotyping algorithms: July 2018Implementation of GWAS analyses: September 2018Draft of manuscript to authors: July 2019First submission: September 2019 |

**\*\*** This section should include: Timeline for completion of project, including approval, project duration, first and second draft of the paper and submission.

**References:**

Boehme AK, Ranawat P, Luna J, Kamel H, Elkind MSV. Risk of acute stroke after hospitalization for sepsis: A case-crossover study. Stroke 2017;48(3):574-580.

Cowan LT, Alonso A, Pankow JS, Folsom AR, Rosamond WD, Gottesman RF, Lakshminarayan K. [Hospitalized Infection as a Trigger for Acute Ischemic Stroke: The Atherosclerosis Risk in Communities Study.](http://www.ncbi.nlm.nih.gov/pubmed/27165961) Stroke. 2016;47(6):1612-7.

Elkind MSV, Carty CL, O’Meara ES, Lumley T, Lefkowitz D, Kronmal RA, Longstreth WT. Hospitalization for infection and risk of acute ischemic stroke: The Cardiovascular Health Study. Stroke 2011;42:1851-1856. [PMCID: PMC3125478]

Elkind MSV, Hills NK, Glaser CA, Lo WD, Amlie-Lefond C, Dlamini N, Kneen R, Hod EA, Wintermark M, deVeber GA, Fullerton HJ, IPS Investigators. Herpesvirus Infections and Childhood Arterial Ischemic Stroke: Results of the VIPS Study. Circulation 2016;133(8):732-41.[PMCID: PMC4766042]

Fullerton HJ, Hills NK, Elkind MSV, Dowling MM, Wintermark M, Glaser CA, Tan M, Rivkin MJ, Titomanlio L, Barkovich JA, deVeber GA, VIPS Investigators. Infection, Vaccination, and Childhood Arterial Ischemic Stroke: The Vascular effects of Infection in Pediatric Stroke Study Results. Neurology 2015;85(17):1459-66.

Kiechl S, Lorenz E, Reindl M, Wiedermann CJ, Oberhollenzer F, Bonora E, Willeit J, Schwartz DA. [Toll-like receptor 4 polymorphisms and atherogenesis.](http://www.ncbi.nlm.nih.gov/pubmed/12124407) N Engl J Med. 2002;347(3):185-92.

[Smeeth L](http://www.ncbi.nlm.nih.gov/pubmed/?term=Smeeth%20L%5BAuthor%5D&cauthor=true&cauthor_uid=15602021), [Thomas SL](http://www.ncbi.nlm.nih.gov/pubmed/?term=Thomas%20SL%5BAuthor%5D&cauthor=true&cauthor_uid=15602021), [Hall AJ](http://www.ncbi.nlm.nih.gov/pubmed/?term=Hall%20AJ%5BAuthor%5D&cauthor=true&cauthor_uid=15602021), [Hubbard R](http://www.ncbi.nlm.nih.gov/pubmed/?term=Hubbard%20R%5BAuthor%5D&cauthor=true&cauthor_uid=15602021), [Farrington P](http://www.ncbi.nlm.nih.gov/pubmed/?term=Farrington%20P%5BAuthor%5D&cauthor=true&cauthor_uid=15602021), [Vallance P](http://www.ncbi.nlm.nih.gov/pubmed/?term=Vallance%20P%5BAuthor%5D&cauthor=true&cauthor_uid=15602021). Risk of myocardial infarction and stroke after acute infection or vaccination. [N Engl J Med.](http://www.ncbi.nlm.nih.gov/pubmed/?term=smeeth+l+nejm) 2004;351(25):2611-8.