emerge network

eMERGE Network: External Collaborator Manuscript Concept Sheet

Reference Number *(to be assigned by CC)*	NT467
Submission Date	Monday, December 12, 2022
Project Title	Genomics of osteoarthritis, prevalence, progression and response to rehabilitation
Tentative Lead Investigator *(first author)*	Alison Rocco
Tentative Lead Investigator Email Address	arocco@uab.edu
Tentative Senior Author *(last author)*	Merry-Lynn McDonald, MSc, PhD
eMERGE Site Sponsor & Contact	UAB: Nita Limdi, PharmD, PhD, MSPH
All Other Authors	**University of Alabama at Birmingham/ Birmingham VAHC** Jasvinder Singh, MD, MPH Joshua Richman, MD, PhD Hemant Tiwari, PhD Brittney Davis, PharmD **Brigham and Women's Hospital/ VABHS (Boston)** Saiju Pyrajan, PhD **Vanderbilt University Medical Center/ VA Tennessee Valley Health Care** Ran Tao, PhD Others welcome
Sites Participating	UAB (others welcome)

Background / Significance	The societal and patient-centered impacts of end-stage osteoarthritis (OA) are profound – (1) costs for OA treatment in the US exceed $41.7 billion annually; (2) by the year 2030, more than 2 million patients with end-stage OA are predicted to undergo total hip or knee arthroplasty (THA/TKA) and subsequent rehabilitation (3) patients who undergo THA/TKA experience profound deficits in health-related quality of life (HRQL), severe limitations in activities of daily living (ADL, i.e., functional limitations); increased healthcare utilization, and higher incidence of comorbidities and hospitalization; and (4) alarmingly the prevalence of moderate-severe functional limitations 2-5 years post-surgery is 30-35% post-THA and 46-50% post-TKA. Rehabilitation approaches are typically not personalized, while escalations in annual THA/TKAs continue. Failed functional restoration after THA/TKA remains a major problem. New evidence-based, personalized approaches should be a high priority and will significantly reduce the economic burden on the healthcare system. There remain gaps in our understanding of the etiology of why some patients experience adverse outcomes post-THA/TKA despite prescribed rehabilitation. **Thus, there is a need for personalized pre and post-rehabilitative biomarkers to enhance physicians' intuition for predicting patient prognosis.**
Outline of Project	Our **overall goal** is to **identify pre and post-rehabilitative genomic biomarkers to enhance physicians' intuition for predicting patient prognosis with OA and/or PTOA.**
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	**Aim 1: To identify genetic variants associated with OA prevalence and progression to end-stage OA.** OA and progression to end-stage hip and/knee replacement will be coded using ICD codes. We will perform association testing with either SAIGE or BOLT-LMM adjusting for age, sex, BMI and principal components capturing population structure. Rare variants from available whole-genome sequence data will be analyzed using burden tests such as SMMAT. Findings from these analyses will provide replication for our research in the Million Veteran Program and All of Us and vice versa.

	Aim 2: To develop and evaluate the performance of a polygenic risk score (PRS) to identify patients at risk of progression to total hip/knee joint replacement. We will train a PRS for OA in MVP. The PRS will be tested and validated in additional subjects from All of Us, eMERGE and Genetics of Osteoarthritis (GO) consortium cohorts. Several PRS tools will be considered, and we will also develop new PRS methods which are robust to population diversity. **Aim 3: To determine whether genetic variants associated with the development of post-traumatic OA are distinct from known genetic determinants of idiopathic OA.** We hypothesize underlying genetic variation predisposes the development of post-traumatic OA. Similar to the approach in Aim 1, we will test for variants associated with PTOA in MVP and eMERGE using SAIGE or BOLT-LMM adjusting for age, sex, BMI and principal components capturing population structure. Rare variants from available whole-genome sequence data will be analyzed using burden tests such as SMMAT. We will assess overlap with OA associated loci in MVP as well as other studies and published findings.
Ethical Considerations	None. Data will be deidentified. No attempts will be made to identify or contact subjects. Individual consent and withdrawal at any stage will be honored.
Available Funding or Resources	VA merit award pending.
Target Journal	Nature Genetics
Milestones *(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)*	Milestones: 1) Phenotype file generation (3 months from approval) 2) GWAS and WGS analyses (2 years from approval) 3) Manuscript initial draft (3 year from approval) 4) Manuscript approval from eMERGE and other studies involved, manuscript submission (3.5 years from approval)

*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