|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT315 | |
| **Submission Date** | 11/30/2018 | |
| **Project Title** | Eliminating Genomic Medicine Research Barriers through Informatics: the eMERGE Network’s Resource Library and its Impact on Multi-Site Projects | |
| **Tentative Lead Investigator** *(first author)* | Brandy Mapes (co-lead), Luke Rasmussen (co-lead) | |
| **Tentative Senior Author**  *(last author)* | Josh Denny | |
| **All Other Authors** | Jodie Jackson, Kayla Howell, Melissa Basford, Josh Peterson, Paul Harris, CC programmers who help maintain/develop eMERGE tools, Jyoti Pathak, Marilyn Ritchie, John Connolly, David Crosslin, Ian Stanaway, George Hripcsak, Peggy Peissig, Sandy Aronson, Casey Overby, Megan Roy-Puckelwartz, Patrick Sleiman, other members welcome to participate | |
| **Sites Participating** | CC + all other interested sites | |
| **Background / Significance** | The eMERGE Network has developed a suite of web-based informatics tools aimed at sharing products and data and improving the quality and efficiency of genomic medicine projects implemented across diverse healthcare settings. From supporting organization and coordination needs of early multi-site startups to assisting with the standardization and collation of large-scale metadata, these tools assist with efforts spanning the entire implementation pipeline. This paper will focus on the development and deployment of the Network’s key resources including those related to translational research (GWAS.org, PheKB, eleMaP, the eMERGE Record Counter, PheWAS catalog, and SPHINX), and those related to clinical implementation and the return of results (CDSKB, MyResults.org, DocUBuild). Co-authors are welcome to propose additional tools that should be considered. We will also review the factors that influenced their creation and lessons learned from promoting their use across the scientific community. | |
| **Outline of Project** | 1. Background on eMERGE Network 2. Describe why development and deployment of informatics tools were/are needed 3. Describe the current tool library (descriptions, usage statistics, publication statistics, use cases) broken down by category according to their role in the translational/implementation pipeline 4. Describe lessons learned/best practices for implementing similar tools across multi-site studies 5. Describe observed barriers to adoption, and how addressed (or how to address in the future) | |
| **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)*  None | |
| **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** | None, paper will focus on best practices and lessons learned from developing/deploying the suite of eMERGE tools | |
| **Ethical Considerations** | None | |
| **Target Journal** | JAMIA | |
| **Milestones**  *(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | Initial Draft: February 2019  Revisions by Co-Authors: March 2019  Final Draft reviewed by Co-Authors: April 2019  Final Draft Submitted: May 2019 | |

**\*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