**eMERGE Network Proposal for Analysis**

Project/Manuscript Concept Sheet

|  |  |
| --- | --- |
| **Submission Date** | 11/8/16 |
| **Project Title** | Incidental and secondary Findings (IFs) in 10,000 eMERGE participants |
| **Tentative Lead Investigator (first author)** | A Gordon |
| **Tentative Senior Author (last author)** | H Rehm, G Jarvik |
| **All other authors** | B Funke, C Eng, M Leduc, D Crosslin, L Amendola, Clinical Annotation and ROR WG, Seq centers, CC genomics group, others as interested |
| **Sites Involved** | All |
| **Background / Significance** | The eMERGESeq panel captures coding regions from all 56 current ACMG genes as well as 12 additional genes and 14 SNVs deemed actionable by eMERGE members. Current literature on IF rate is limited by both the lack of CNV data and poor representation of minority populations. Initial sequencing results from eMERGE can begin to address both of these issues: initial sample sets across the network will be ancestrally diverse, and eMERGEseq can detect clinically relevant CNVs. |
| **Outline of Project** | We hope to include IF data from approximately the first 1000 participants per site. This would represent an adequate sample size that includes all sites while enabling a more rapid and meaningful publication than waiting for sequencing of all samples to complete |
| **Desired**  **Variables (essential for analysis**  **indicated by \*)** | Primary phenotype indication for sequencing\*, age, self-reported ancestry\*, and site\* and result\* for each pathogenic or likely pathogenic IF in our actionable list.  These analyses will *not* include follow-up rephenotyping of those found to have IFs. |
| **Desired data** | Principle components and IF. |
| **Planned Statistical Analyses** | Analyses of IF rate will be performed by gene to adjust for test indication (ie, if the primary indication was colorectal cancer, the case will not be used in determining the IF rate for that gene. Analyses will be stratified by ancestry. An overall IF rate will be calculated for patients unselected for relevant phenotypes.  Geisinger data may be considered, if ascertainment bias can be accommodated, but may be excluded or analyzed separately. |
| **Ethical considerations** | What data are returned to whom; not all sites will return the same sets IFs to patients. There is already a plan for a separate manuscript on which sites will return which genes/SNVs. |
| **Target Journal** | AJHG |
| **Milestones\*\*** | Data template 12/2016  Introduction, methods 2/17  Pilot analysis with ~4 sites 4/17  Full data completed by 7/2017, per sequencing schedule.  Draft 8/17 |

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