National Advisory Council for Human Genome Research (NACHGR), February 11, 2019

Concept Clearance for RFAs

Electronic **Me**dical **R**ecords and **Ge**nomics (eMERGE): Genomic Risk Assessment and Management

Purpose

NHGRI proposes three RFAs to continue to use biorepositories linked with electronic medical records (EMRs) for genomic research by developing, evaluating, and disseminating genomic risk assessment and management tools for clinical use. Specifically, this effort will use genomic variant and EMR data to: 1) calculate validated polygenic risk scores (PRS) for several complex diseases retrospectively using available datasets, and share the distributions and associations of these PRS; 2) develop EMR-based methods to communicate genomic risk profiles and relevant clinical recommendations based on PRS, family history, and other clinical data; 3) recruit and genotype 20,000 individuals of diverse ancestry, prospectively calculate their genomic risk for selected e-phenotypes, and provide risk estimates and management recommendations to them and their providers; and 4) use EMR-based methods to assess uptake of risk-reduction recommendations and impact on related clinical outcomes.

Background

eMERGE is a national network of biorepositories linked to EMRs funded by NHGRI to conduct genomic medicine discovery and implementation research. Since its inception in 2007, eMERGE has demonstrated the utility of EMR-derived e-phenotypes for genomic discovery, conducted pilot clinical implementation studies, expanded to include pediatrics and pharmacogenomics (PGx), and assayed genomic variation in 109 clinically relevant genes to assess penetrance of identified variants and use them to improve the clinical outcomes of patients harboring them.

Parallel to the advances and resources produced by eMERGE, substantial progress has also been made in developing clinically certified tests of genomic variation, reaching consensus on interpretation of actionable variants, and providing professional guidelines for their clinical use. Genomic variants predictive of a variety of complex diseases have been combined into genomic or "polygenic" risk scores (PRS) for identifying persons at several-fold increased risk of diseases such as atrial fibrillation, type 2 diabetes, coronary disease, Alzheimer's disease, and breast cancer. Automated tools for systematic, patient-driven collection of family history have also been developed and integrated into EMRs, and PGx variants predictive of altered drug response have been identified and shown to be carried by nearly everyone.

These developments are setting the stage for rapid clinical adoption of genomically-defined assessment and management of complex disease risk, but several gaps remain to be filled before widespread adoption and implementation can be expected. If genomic risk estimates are to be utilized effectively by clinicians, these estimates should be seamlessly integrated into EMRs and delivered with recommended clinical actions in a user-friendly manner, but such EMR estimation tools have yet to be developed and validated. If EMR-derived phenotypes are to be

used as outcomes for assessing the impact of genomic risk reporting, their predictability by these methods should be defined. Estimation and validation of PRS have been conducted almost exclusively in European ancestry (EA) populations; their predictive value should be assessed in non-EA populations. PRS based on e-phenotypes are needed for non-EA populations as well. Patient and provider uptake of risk-reduction recommendations across a range of conditions should be assessed. The added value of genotypic information over standard clinical data in a variety of complex diseases and demographic subgroups should also be assessed.

Proposed Scope and Objectives

This initiative will support 6-10 Clinical Sites and a Coordinating Center to: 1) validate existing PRS and related risk algorithms against EMR-defined phenotypes, using existing genotype and phenotype data; 2) develop and evaluate EMR-based methods to communicate genomic risk profiles and relevant clinical recommendations; 3) recruit and consent 20,000 persons of diverse ancestry who will undergo genomic risk assessment using dense SNP genotyping arrays, online family history tools, and other clinical data, and will receive risk estimates and risk-reduction recommendations using EMR tools; and 4) quantify the uptake of recommendations and their impact on related clinical outcomes. Awardees would likely spend the first year validating genomic risk scores against e-phenotypes using publicly available data on 83,000 eMERGE participants or other available data from studies such as *All of Us* or the Million Veteran Program. Addition of high-quality genotype and phenotype data from ancestrally diverse populations will be strongly encouraged. The first year would also be devoted to validating risk algorithms against e-phenotypes and refining them as needed, and to developing risk communication and electronic clinical decision support (eCDS) to be deployed at each site.

The genomic risk assessment would involve using a clinically certified (preferably commercially available) SNP array assaying variants predictive of modifiable disease risk. The risk assessment would also include online tools for patient-entered family history information and any other state-of-the-art genomic risk assessments. These risk assessments would be combined to identify persons with a 2- to 3-fold increased risk of complex disease compared to the general population. Individualized risk estimates would be reported to providers and patients along with guideline-based risk-management information through eCDS incorporated in the EMR. Actions taken on these recommendations, such as increased disease screening and surveillance or changes in therapy, would be assessed through the EMR, and changes in modifiable risk factors (such as decreased blood pressure in persons at risk for stroke) would be identified.

Other clinical risk factors (such as hypertension, obesity, and health behaviors) as well as medication use and e-phenotypes for pre-existing disease, would be extracted from EMRs. Outcomes of interest would include proportions of patients in whom at least one risk-reduction intervention (such as a screening test or a drug prescription change) was applied, estimating that roughly 50% uptake would be needed for the genomic risk assessment and communication effort to be deemed worthwhile. Uptake rates lower than that during the study's course could prompt stakeholder consultation and revisions of the tools, as needed. Outcomes could also include change in a quantifiable disease measure (such as bone mineral density or cholesterol level) as a pilot measure of the effectiveness of risk reduction recommendations.

The design and components of the genomic risk assessment, including the guideline-based risk reduction recommendations to be delivered, will be proposed by each applicant; a single network-wide protocol for risk assessment, risk modification, and outcome ascertainment will be finalized by consensus among the awardees. A study of 20,000 patients would identify 5,000 patients at the highest 2% of risk for 15 independent diseases (based on binomial probabilities, as patients are at risk for more than one) and would provide 95% confidence to estimate the percent uptake of risk reduction recommendations for each condition with a 5% margin of error. Assuming a key demographic group comprises 20% of the high-risk subset (1,000 patients), the uptake could be estimated with an 11% margin of error at 95% confidence.

Through these efforts to develop and apply genomic risk assessment and management tools, eMERGE would continue its foundational efforts in developing and distributing e-phenotypes, using them to identify genomic associations, improving and implementing eCDS tools, and providing best practices for conducting genomic research in biorepositories linked to EMRs.

Criteria for selecting clinical sites will include ability to: 1) provide $\geq$ 2,500 new participants who have not previously received genomic risk information, of whom $\geq$ 35% (and preferably 75%) are from ancestrally diverse or medically underserved (https://data.hrsa.gov/tools/shortage-area/mua-find) populations; 2) implement e-phenotyping, eCDS, and outcome assessment in a comprehensive EMR; and 3) provide valid e-phenotypes for multiple clinical outcomes.

Relationship to Ongoing Activities

eMERGE will continue to build resources and tools focused on EMRs for genomic medicine and conduct implementation research. This differs from the *All of Us* (AoU) Research Program that focuses on resource building with minimal implementation. AoU sites will be encouraged to apply and bring their genome-sequencing data and additional infrastructure. eMERGE differs from the Clinical Sequencing Evidence Generating Research (CSER) project in that CSER's sites are conducting discrete studies of clinical sequencing in varied settings, while eMERGE uses network-wide genomic and phenotype datasets and develops network-wide e-phenotypes and eCDS, all of which is used for genomic discovery and implementation. eMERGE differs from the Implementing Genomics in Practice (IGNITE) network in that IGNITE is conducting discrete trials of genomic medicine implementation, while eMERGE will implement genomic risk assessment and reduction efforts and assess the uptake of this approach by providers and patients.

Mechanism of Support

These three RFAs—one each for Clinical Sites recruiting $\geq$35% or $\geq$75% of participants from diverse or underserved populations, and one for the CC—will use the cooperative agreement award mechanism (U01). Anticipated duration of the program is 5 years, starting in FY2020.

Funds Anticipated

NHGRI will commit approximately $65M over 5 years to support 6-10 Clinical Sites and the Coordinating Center. Costs will peak in years 2-4 during recruitment and genomic testing.