

## **NHGRI Workshop: *Future Directions for the eMERGE Network*, January 22, 2014 Executive Summary**

The NHGRI Workshop on Future Directions for the Electronic Medical Records and Genomics (eMERGE) Network gathered eMERGE principal investigators and workgroup co-chairs, the eMERGE External Scientific Panel, and external experts in key areas of genomic medicine and discovery to review the current goals and accomplishments of eMERGE and suggest future directions for NHGRI to consider as the possibility of an eMERGE Phase III approaches. NHGRI Director Eric Green presented NHGRI's current portfolio of genomic medicine research and eMERGE Steering Committee Chair Rex Chisholm reviewed eMERGE progress to date, followed by eight panel discussions.

Goals and accomplishments: In eMERGE Phase I (2007-2011), the Network of 5 study sites demonstrated the utility of linking electronic medical records (EMRs) to biorepositories for genomic research, allowing eMERGE to identify novel associations through genome-wide association studies (GWAS) and to re-use those GWAS data for a variety of EMR-defined phenotypes. A closely integrated ELSI component explored consent, privacy, and other ELSI issues surrounding this kind of research. In eMERGE Phase II (2011-2015), two additional adult sites and three pediatric institutions joined the Network, and the scope broadened to include pilot clinical implementation studies. eMERGE II has expanded the electronic phenotyping library to continue GWAS discovery studies and has initiated integration of actionable variants into the EMR for use in clinical care. Currently, the network has over 325,000 participants with EMRs and ~76,000 samples genotyped. Imputation of the genotyped data created a merged dataset including ~52,000 participants. Roughly 9,000 participants are undergoing targeted sequencing and reporting of pharmacogenetics genes in the eMERGE-PGx project. Other accomplishments include establishing a public library of electronic phenotypes; developing new methods for genomic discovery such as PheWAS; sequencing pharmacogenetic genes for both discovery and use in clinical care; and developing software/tools to integrate genetic information into EMRs for clinical decision support.

Panel discussions: Panelists were asked to focus on the following questions:

- How should eMERGE continue to contribute to the scientific community?
- How can eMERGE balance genomic discovery and implementation?
- What approaches should eMERGE apply for genomic discovery?
- What implementation research should eMERGE focus on?
- What issues should eMERGE address in terms of ethical, legal and social implications?

Workshop participants agreed that eMERGE is well-positioned for genomic research in several ways, including its focus on EMRs for phenotyping, phenome-wide studies, data re-use for genomic discovery, and integration of genomic findings and implementation in the EMR; its large size and diverse population; its inclusion of pediatric centers and potential for bringing genomic translation to pediatrics; its closely integrated infrastructure for empirical ELSI research on implementation; and its site-specific and network-wide efforts at implementing genomic findings in diverse clinical settings. Recommendations for future directions included:

### Phenotyping and EMRs

- Develop new methods/tools to make phenotyping fast, accurate, and reproducible and to account for confounders for genomic discovery and implementation;
- Leverage the richness of EMR data by applying phenomic approaches and assessing longitudinal, rare, pharmacogenomics, and disease subtype phenotypes;
- Identify undiagnosed phenotypes in carriers of rare genotypes;
- Create “modular” phenotypes to facilitate transportability and restructuring into new phenotypes;

- Use structured data dictionaries across the Network to increase poolability of data;
- Address EMR transportability and standards for sharing genomic data between institutions;
- Investigate how the ‘phenotyping workforce’ can be expanded to scale beyond eMERGE institutions, particularly with a view to collaborating with CTSA awardee institutions;
- Convene a forum with health systems’ clinical leaders and EMR vendors/other stakeholders on improving integration of genomic data and enhancing use of EMRs for genomic research.

### Discovery

- Balance discovery and implementation research, emphasize research *on* implementation;
- Sequence ~100 disease-related genes including some of the 56 ACMG-reportable genes and return results to determine how best to address incidental findings in such genes;
- Select maximally informative phenotypes for sequencing such as extreme discordant phenotypes, tails of distributions, etc.;
- Determine the most appropriate approaches to study rare but collectively common variants that could inform improved treatment, especially of underrepresented minority patients;
- Generate dense data through sequencing (genome, exome, targeted genes) and/or a genotyping platform that includes Loss of Function (LoF) variants and CNVs;
- Expand data collection processes to include additional sources of RNA/DNA, environmental data, family history, EMR-defined co-morbidities, etc.;
- Develop analytic tools to incorporate other sources of data (e.g., ENCODE, GTEx) and environmental data into analyses to explore potential pathogenic or causal variants;
- Address penetrance (particularly of “Mendelian” variants), heritability, and pathogenicity;
- Collaborate with similar genomics research efforts such as UK Biobank, Kaiser, and VA.

### Implementation

- Develop best methods for conducting pragmatic clinical trials across the network to assess implementation;
- Establish learning health care systems and engage experts in clinical care and workflow;
- Include diverse populations (race/ethnicity and age) in clinical implementation;
- Expand genotyping focus to gene-phenotype pairs instead of genotype-phenotype pairs allowing for closer collaboration with gene-based groups like ClinGen;
- Study the impact of re-contact and re-annotation of actionability;
- Develop best approaches to re-annotation and who is best responsible for it, including methods for versioning actionability information to view past clinical decisions in context;
- Generate data on efficiency, cost-effectiveness, and ease of implementation when engaging chief medical informatics officers (CMIOs) and institutional IT groups;
- Develop cross-network models for conducting economic/value-based research on genomic implementation, such as economic analyses assessing the impact of RoR;
- Engage and educate stakeholders such as IRBs, clinicians, EMR vendors, administrators.

### Pediatrics

- Lead translation of genomics into pediatrics;
- Target conditions for genomic analysis that may have early clinical utility;
- Expand the network over time to increase the diversity of the pediatric population;
- Increase overlap between targeted adult and pediatric phenotypes, as with cross-cutting phenotypes such as appendicitis or opiate use and fetal addiction syndromes;
- Take opportunities to use the national birth defects registry to expand its population base;
- Explore clinical utility of intervening on potential adult onset disease risk in children;
- Store samples for future studies of methylation changes from childhood to adulthood.

## ELSI

- Establish broad biobank consent to return CLIA research results to EMRs for clinical use;
- Explore legal barriers including CLIA, HIPAA, re-contacting participants, and RoR;
- Examine potential thresholds for clinical tests requiring specific consent and counseling across a range of sensitivity of clinical tests, such as TPMT to Huntington's disease;
- Develop cross-network models for conducting economic/value-based research;
- Ensure RoR studies include diverse populations;
- Study local differences across IRBs in genomics expertise and promote IRB education;
- Evaluate the impact of increasing access to EMRs and poll what risks/variants information patients want in their records and how they would like this information to be displayed;
- Explore the risks of re-identification of de-identified patients for use in re-consent and RoR;
- Examine the impact of RoR on proband's relatives;
- Assess what happens long-term after RoR, such as lifestyle/behavior changes.

Summary and Conclusions: eMERGE can best contribute to the scientific community by continuing its emphasis on EMRs, developing new methods and tools for phenotyping, and leveraging the EMR for genomic discovery and clinical implementation. eMERGE should continue to include both discovery and implementation, and is uniquely positioned to conduct research *on* implementation. Discovery in eMERGE should leverage the rich EMR phenotyping data and utilize state-of-the-art techniques for measuring genomic variation, including CNVs. Discovery studies in eMERGE should include additional sources of RNA/DNA, EMR-defined comorbidities and environmental data to account for confounding factors, EMR-defined phenotypes of rare variant carriers; and functional data for identifying possible causative variants. Implementation research in eMERGE should include rare but collectively common disease-related variants that could inform improved treatment, explore differences in implementation among diverse subgroups, develop best approaches to variant re-annotation, and generate data on efficiency, cost-effectiveness, and ease of implementation. One convergence of discovery and implementation research utilizing eMERGE's unique strengths could involve sequencing highly clinically relevant genes, developing best methods for returning these results, and assessing their impact. The integrated ELSI infrastructure in eMERGE should continue to explore legal and economic implications of genomic medicine implementation, examine and minimize local differences across IRBs in genomics expertise and decision-making; and assess impact of RoR.