External Scientific Panel background materials
## TABLE of CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase II Citation Analysis</td>
<td>Pg. 2</td>
</tr>
<tr>
<td>Phase II Publications</td>
<td>Pg. 3</td>
</tr>
<tr>
<td>Overview of eMERGE Tools</td>
<td>Pg. 21</td>
</tr>
<tr>
<td>Site Summaries</td>
<td>Pg. 22</td>
</tr>
<tr>
<td>Minutes from ESP Meeting, May 2015</td>
<td>Pg. 32</td>
</tr>
<tr>
<td>Minutes from ESP Meeting, May 2015, Executive Session</td>
<td>Pg. 33</td>
</tr>
</tbody>
</table>
PHASE II CITATION ANALYSIS

CITATIONS OF EMERGE PUBLICATIONS BY CATEGORY

- Genomics 4934
- CERC/ROr 1210
- Privacy 644
- PGx/Implementation 349
- Phenotyping 2005
- EHRI 459
- All Network/Foundation 1045

Cumulative Citation Count
2007 - January 2016: 10,646

Number of published projects through January 2016

- Site Projects (356)
- Network Projects (190)
- 546 Total Projects
PHASE II BIBLIOGRAPHY & REMAINING IN-PROCESS PROJECTS

[ Digital Reference Library Available Here ]

eMERGE Phase II Publications from June 2011 – January 2016

Digital Reference Library Available Here

In Progress Phase III Network Manuscripts
1. Approaches to Returning Clinically Actionable Results from Next Generation Sequencing Panel in a Healthy Population. Lead Investigator: Tracy McGregor (VU)
2. Knowledge driven rare variant PheWAS in eMERGE to identify regions associated with disease using collapsing based approach. Lead Investigator: Anna O Basile (Geisinger)
3. GWAS study on non-alcoholic fatty liver disease (NAFLD) in pediatric and adult population: comparison of size effect between adult and children using participants of the eMERGE Network. Lead Author: Bahram Namjou (CCHMC)
4. The identification and reporting of actionable incidental genetic variants from large scale clinical sequencing of drug response genes. Lead Investigator Quinn Wells (VU)
5. Feasibility of using geocoded US Census/ACS variables as proxy for socioeconomic status in genotype-phenotype interaction studies of T2DM and obesity. Lead Investigator: Kathryn Jackson (NU)

Published/Accepted and Submitted Phase II Network Manuscripts
experience. J Pathol Inform. 2015;6:50. PMCID: PMC4629307


5 of 34


82. Chute CG, Kohane IS. Genomic medicine, health information technology, and patient care. JAMA. 2013 Apr 10;309(14):1467–1468. PMCID: PMC3959893


86. Kullo IJ, Jarvik GP, Manolio TA, Williams MS, Roden DM. Leveraging the electronic health record to implement genomic medicine. Genet Med. 2012 Sep 27; PMCID: PMC4206937


In Process Phase II Network Manuscripts

1. Association of rare and common variants with insulin resistance: results from the eMERGE Network. Lead Investigator: Daniel Kim
   (UW)
2. Investigation of CETP SNPs with LDL-C, BMI and risk of T2D. Lead Investigator: Brendan Keating (UPenn)
3. Quantitative and discreet trait analysis across eMERGE-I phenotypes. Lead Investigator: Joseph Glessner (CHOP)
4. Association of Y chromosomal variants with leukocyte count and cardiovascular disease: results from the eMERGE Network. Lead
   Investigator: Yatong Li (UW)
5. Multi-site IRB review – experience of the eMERGE Network. Lead Investigator: Jen McCormick (Mayo)
6. A phenome-wide association study to discover pleiotropic effects of PCSK9. Lead Investigator: Maya Saforova (Mayo)
7. An investigation of somatic mutations in PGx-eMERGE dataset. Lead Investigator Kenneth Kaufman (CCHMC)
8. The eMERGE Network: The practice of patient education in the return of genomic medicine results. Lead Investigator: Cassandra
   Perry (BCH)
9. Assessing opinions about healthcare provider-targeted MyResults.org content. Lead Investigator: Casey Overby (NU)
10. Genome wide association studies and gene based interaction study for Ocular Hypertension. Lead Investigator: Shefali Verma
    (Penn State)
11. Agreement between research-grade sequencing and CLIA validation genotyping in eMERGE-PGx. Lead Investigator: Laura R Torvik
    (NU)
12. Replication of another sites’ GWAS findings for Squamous Cell Carcinoma (SCC) and Actinic Keratosis (AK) findings using existing
    PheWAS codes from the “Systematic comparison of phenome-wide association study of electronic medical record data and
    genome-wide association study data” paper. Lead Investigator: Josh Denny (VU)
13. The identification of adverse events in the eMERGE PGx cohort using the electronic health record, and assessing association with
    genetic variation in the 84 pharmacogenes. Lead Investigator: David Crosslin (UW)
14. An investigation into the genetics of Intractable Epilepsy in the pediatric population. Lead Investigator: Berta Castillo (CHOP)
15. Patients’ views on consent and data sharing in biobank research: A large multisite experimental survey in the US. Lead Investigator:
    Saskia Sanderson (MSSM)
16. The eMERGE Network: Healthcare provider education to support genomic medicine in practice. Lead Investigator: Carolyn R.
    Vitek (Mayo)
17. A targeted sampling scheme utilizing both EHR and census information. Lead Investigator: Nate Mercaldo (VU)
18. Feasibility of using geocoded US Census/ACS variables as proxy for socioeconomic status in genotype-phenotype interaction
    studies of T2DM and obesity. Lead Investigator: Kathryn Jackson (NU)
19. Cognitive Interviews associated with developing a national survey on consent across a national network of genomic medicine
    sites. Lead Investigator: Melanie Myers (CCHMC)
20. Developing a national survey on consent across a national network of genomic medicine sites. Lead Investigator: Ingrid Holm &
    Maureen Smith
    Investigator: Naniba’ Garrison (VU)
23. A Highly Accurate Electronic Algorithm for the Classification of Asthma Severity in Children. Lead Investigator: Erik Hysinger
    (CHOP)
24. Epistatic gene-based interaction analyses for glaucoma in eMERGE network and NEIGHBOR consortium. Lead Investigator: Shefali
    Verma (PSU)
25. Exploring the genetic architecture of Age-Related Macular Degeneration (AMD) in the eMERGE network. Lead Investigator: Molly
    Hall (PSU)
26. Genome-wide Association Study of Gastroesophageal Reflux Disease (GERD) in Adult and Pediatric Populations. Lead Investigator:
    Patrick Sleiman (CHOP)
27. Genome-wide Association Study of Atopic Dermatitis in Adult and Pediatric Populations. Lead Investigator: Berta Almoguera
    (CHOP)
28. Genome-wide Association Study of Asthma in Adult and Pediatric Populations. Lead Investigator: Berta Almoguera (CHOP)
29. Genome-wide Association Study of Attention Deficit Hyperactivity Disorder (ADHD). Lead Investigator: John Connolly (CHOP)
30. Investigation of PCSK9 SNPs with LDL-C, BMI and risk of T2D. Lead Investigator: Brendan Keating (CHOP)
31. Phenotype transportability across Electronic Health Records. Lead Investigator: Joshua Denny (VU)
32. The COGENT consortium meta-analysis of blood pressure African ancestry cohorts. Lead Investigator: Dinga Velez Edwards (VU)
33. Genetic variation among 84 pharmacogenes from the PGRNSeq in the eMERGE Network. Lead Investigators: Will Bush (Case) & David Crosslin (GH/UW)
34. PheWAS analysis of a functional variant in CDHR3. Lead Investigator: Michael March (CHOP)
35. PheWAS analysis of homozygous deletions in GWAS data. Lead Investigator: Michael March (CHOP)
36. Using PheWAS to assess disease comorbidity and potential pleiotropy of genetic risk scores for rheumatoid arthritis. Lead Investigator: Robert Carroll (VU)
37. Association of rare and common variants in LDLR, HMGCR, NAT2, ABCA1, and APOA1 with plasma lipid levels: results from 9000 participants of the eMERGE Network. Lead Investigators: Daniel Kim (Michigan), Erin Austin (Mayo)
38. Exploring genetics and outcomes associated with acute kidney injury (AKI) using electronic health records and genomics. Lead Investigator: Girish N Nadkarni (Mt. Sinai)
39. PGRNseq and GWAS predictors of Methylphenidate (MPH) response. Lead Investigator: Tanya Froelich (CCHMC)
40. Discovery, Replication and Clinical Associations of Pathway-Based Trans-eQTL. Lead Investigator: Laura Wiley (VU)
41. Multiscale Analysis Of Influenza Host-Pathogen Interactions: Fluomics. Lead Investigator: Ellie Sang Sukerman (Northwestern)
42. Rare RYR1, CACNA1S variant annotation, exposure history, observed phenotypes in cases and controls. Lead Investigator: Senthilkumar Sadhasivam (CCHMC)
43. Variant Calling and Annotation for 82 known pharmacogenes (tentative). Lead Investigator: Will Bush (CC)
44. Characterizing the individual and shared genetic components of pheWAS phenotypes. Lead Investigator: Jonathan Mosley (VU)
45. Examining gene variants in eMERGE samples for association with uterine fibroids. Lead Investigator: Todd Edwards (VU)
46. A Phenome-wide Survey of the Phenotypic Effects of Neanderthal Admixture. Lead Investigator: Tony Capra (VU)
47. MVtest: a method to flexibly model the genetic determinants of trait variability. Lead Investigator: Todd Edwards (VU)
49. Pediatric Providers are Poor at Identifying Severe Obesity in Young Children at Two Tertiary Pediatric Medical Centers. Lead Investigator: Cassandra Brady (CCHMC)
50. Developing an Algorithm to Detect Early Childhood Obesity Two Tertiary Pediatric Medical Centers. Lead Investigator: Todd Lingren, Vidhu Thaker (BCH)
51. Association of APOL1 G1/G2 risk alleles with metabolic and cardiovascular traits. Lead Investigator: Girish Nadkarni (Mt. Sinai) & Miriam Udler (BCH)
52. Development of a dynamic XML event-driven ophthalmologic data capture framework. Lead Investigator: Peggy Peissig (Marshfield)
53. Burden of structural variation and PheWAS. Lead Investigator: David Crosslin (GroupHealth)
54. Knowledge driven search for gene-gene interactions associated with hypothyroidism in the eMERGE network. Lead Investigator: Molly Hall (MC/EIRH/PSU)
55. Discovery and replication of genetic interactions for quantitative lipid traits. Lead Investigator: Marylyn Ritchie (MC/EIRH/PSU & CC)
56. Genetic Risk Scores for Complex Diseases in the eMERGE Network: Characterization and Predicative Abilities in Clinical Settings. Lead Investigator: Logan Dumitrescu (VU)
57. Phenotypes Seen in Cohorts with Rare Variants in Six PGRN-Seq (VIP) Genes also Identified by the ACMG as Priority Genes for Reporting Incidental Findings. Lead Investigator: Josh Denny (VU)
58. Chromosomal Anomalies that Affect Levels of White Blood Count (WBC) and its Differential. Lead Investigator: David Crosslin (UW)
59. Extracting the Quality of Prostate Cancer Care from Electronic Healthcare Records. Lead Investigator: Tina Hernandez Boussard
60. Practical Approaches to the Omic Chasm. Lead Investigator: Justin Starren (NU)
61. Developing Consents for Returning Pharmacogenomics Results: The eMERGE Experience. Lead Investigator: Maureen Smith (NU)
62. Portable Applications for Implementing Multi-Site Clinical NLP Algorithms. Lead Investigator: David Carrell (GroupHealth)
63. Evaluation of a Differentially Private Top-k SNP Publication Strategy. Lead Investigator: Mehmet Kuzu (External Collaborator); Brad Malin (VU)
64. Effective Use of Electronic Health Records to Identify Venous Thromboembolism: Results from the eMERGE Network. Lead Investigator: Jyoti Pathak (Mayo)
65. GWAS of Infection or Colonization with Community Associated Methicillin-Resistant Staphylococcus Aureus (CA-MRSA). Lead Investigator: Abel Kho (NU)
67. Diverticulosis. Lead Investigator: Will Thompson (NU)
68. The Geographic Distribution of Colon Polyps. Lead Investigator: Will Thompson (NU)
69. Colon Polyps. Lead Investigator: Abel Kho (NU)
70. Genetic Variants Associated with Response to Heart Failure Treatment: The Electronic Medical Records and Genomics (eMERGE) Network. Lead Investigator: Sue Bielinski (Mayo)
71. Genome Wide Association of Risk of Heart Failure: The Electronic Medical Records and Genomics (eMERGE) Network. Lead Investigator: Sue Bielinski (Mayo)
72. Genetic Variation that Predicts Susceptibility to Clostridium Difficile. Lead Investigator: Josh Denny (VU)
73. Genome-wide study of resistant hypertension using existing genomic data and electronic medical records. Lead Investigator: Logan Dumitrescu (VU).
74. GWAS of Venous Thromboembolism (VTE) among White Americans. Lead Investigator: John Heit (Mayo)
75. GWAS of Venous Thromboembolism (VTE) among African-Americans. Lead Investigator: John Heit (Mayo)
76. Genome-Wide Association Study of Abdominal Aortic Aneurysms with Electronic Medical Record Phenotyping. Lead Investigators: Helena Kuivaniemi and Gerard Tromp (Geisinger)

**Site-Specific Manuscripts Published During Phase II**

**BCH/CCHMC**

10. Kohane IS. *An Autism Case History to Review the Systematic Analysis of Large-Scale Data to Refine the Diagnosis and Treatment of Neuropsychiatric Disorders.* Biol Psychiatry. 2014 Jun 12; PMCID: PMC4260993


CHOP


60. Mitchell JA, Hakonarson H, Rebbeck TR, Grant SF. *Obesity-susceptibility loci and the tails of the pediatric BMI distribution*. Obesity (Silver Spring) [Internet]. 2013 Feb 14. PMID: 23661695


**Geisinger**


2. Williams MS. *Perspectives on what is needed to implement genomic medicine*. Mol Genet Genomic Med. 2015 May;3(3):155–159. PMID: 244444156


10. Shirts BH, Jacobson A, Jarvik GP, Browning BL. Large numbers of individuals are required to classify and define risk for rare variants in known cancer risk genes. Genet Med. 2013 Dec 19. PMCID: PMC4063879
17. Larson EB. Building trust in the power of “big data” research to serve the public good. JAMA. 2013 Jun 19;309(23):2443–2444. PMID: 23780455


30. Haasl RJ, McCarty CA, Payseur BA. Genetic ancestry inference using support vector machines, and the active emergence of a unique American population. Eur J Hum Genet [Internet]. 2012 Dec 5; PMID: PMC3641388

Mayo


Mount Sinai


Northwestern


Vanderbilt


OVERVIEW of eMERGE TOOLS

Main resources and tools produced and supported:

- **CDS KB**: A repository of clinical decision support knowledge designed to support clinical processes from diagnosis and investigation through treatment and long term care.

- **eleMAP**: A tool that allows researchers to harmonize their local phenotype data dictionaries to existing metadata and terminology standards such as the caDSR (Cancer Data Standards Registry and Repository), NCIT (NCI Thesaurus) and SNOMED-CT (Systematized Nomenclature of Medicine-Clinical Terms).

- **eMERGE Infobutton Project**: A template containing important content topics for genomic medicine.

- **eMERGE Record Counter**: A web-based research tool that provides exploratory data figures for research planning purposes and feasibility assessment.

- **Model consent language**: A publication representing the compiled work of eMERGE I investigators and consultants on consent language for the collection and storage of human biospecimens and data for future research, particularly those collections that have an electronic medical records component.

- **MyResults.org**: A website to educate providers and patients about genetic test results and the related medications.

- **PheKB**: An online collaborative environment for building and validating electronic algorithms to identify characteristics of patients within health data.

- **PheWAS Catalog**: A catalog that contains the PheWAS results for 3,144 single-nucleotide polymorphisms (SNPs) present in the NHGRI GWAS Catalog.

- **SPHINX**: A data exploring tool for genetics related drug response hypothesis generation.

- **Synthesis-View, PheWAS-View** and Phenogram: Visualization tools for genome & phenome-wide data.
SITE SUMMARIES

The Center for Applied Genomics (CAG) at Children’s Hospital of Philadelphia

- **Aim 1: Phenotype Characterization in Four Pediatric Cohorts**
  - We have been active in the Phase II Phenotyping workgroup, leading the asthma (4,498 CHOP samples, 3,680 additional eMERGE samples), asthma severity (4,498; >1000 eMERGE), ADHD (1,013 CHOP samples, +335 eMERGE), atopic dermatitis (1,695 CHOP, +383 eMERGE), and GERD (1,573 CHOP, eMERGE in progress) algorithms. For Phase III, we propose to develop algorithms for the following primary phenotypes: intellectual disability, epilepsy, autism and obesity, all of which have multiple covariate phenotypes to pursue (i.e. obesity patients are enriched for asthma, obstructive sleep apnea, diabetes and metabolic syndrome). We will also mine EHRs to determine pharmacogenomic (PGx) response profiles (efficacy and adverse events), and search for polymorphisms impacting variation in response to commonly-used drugs in the existing pediatric dataset, leveraging CHOP and eMERGE data. We hypothesize that data-mining of information from GWAS and EHRs from >60,000 children of different ethnic backgrounds will result in discovery of biomarkers and disease/PGx profiles that can inform clinical care. (Note: Here, we collapse Pediatric/PGx workgroups under Phenotyping)

- **Aim 2: Characterization of Common and Rare Variants in >3,000 Pediatric Patients**
  - CAG has built a comprehensive infrastructure for detecting novel rare variants, yielding more than 375 publications since our founding in 2006. Under eMERGE-II, we (re-)developed a suite of genomics tools including ANNOVAR (950 citations), PennCNV (751), SNVER (57), ParseCNV (11), and PennCNV-Seq (10). While continuing to drive discovery in the complex phenotypes (as per eMERGE-II), these tools will be further optimized for detecting rare variants from existing genotype and sequencing data, as well as prospective data generated under eMERGE Phase III. We aim to build upon extensive preliminary data and public resources to catalog and validate rare variants in those phenotypes proposed above, as well as those proposed by other eMERGE sites. We hypothesize that continued optimization of CAG’s analysis pipeline will uncover a range of actionable variants of diagnostic/clinical significance in a range of pediatric and adult diseases.

- **Aim 3: Return of Results (RoR) to 160+ Patients with Actionable Genomic Variants**
  - Under eMERGE Phase II, we extended our Clinical Laboratory Improvement Act (CLIA)/College of American Pathologists (CAP) certified workflow status. We are CLIA-approved for genotyping, clinical pathology- , whole exome- and whole genome- sequencing (ThermoFisher and Illumina). We have developed a highly efficient genomics pipeline for sample-tracking, quality-control, base calling, and annotating variants all of which are being conducted in-house in a CLIA-approved environment. Our analysis team and state-of-the-art genotyping/sequencing resources, affords flexibility to efficiently validate actionable variants in a wide range of phenotypes. We hypothesize that state-of-the-art informatics analyses will accelerate characterization of actionable variants, maximize validation success and clinical utility, and ultimately improve patient care.

- **Aim 4: Exploration of Challenges Facing At-Risk Families**
  - We propose a series of formative and summative analyses to assess the health impact, cost-effectiveness, and ELSI implications of the site for patients. We will follow patients from our RoR program, leveraging CAG’s re-contact protocol to assess the health consequences of patients for whom results have been returned under SA3. Ultimately, this will feed-forward to improving guidelines and governance rules for CAGs biorepository, databases, and RoR procedures. In addition, we propose a series of developments of the CAG-developed MyResults.org website, established
under eMERGE-II. We hypothesize that these programs will allow refinement to RoR protocols and yield successful and improved healthcare-related outcomes for CHOP patients.

- **Aim 5: Integration of Genomic Data into CHOP’s EHR**
  - Under eMERGE II, we worked extensively with the IRB and EpicCare teams to develop a protocol for returning results to CAG patients, and our relevant infrastructure is live and operational. We propose to dramatically expand our EHRI platform in eMERGE-III, and to integrate actionable results identified under SA3. In addition MyResults will be redeveloped to improve transparency and include a physician education sub-section, with tie-in to clinical decision support (CDS) interfaces at CHOP, and other eMERGE sites. We hypothesize that these efforts will leave CHOP poised for large-scale integration of genomic results into its Epic EHR platform, which will leverage the cumulative efforts of SA1-4.

**Cincinnati Children’s Hospital Medical Center (CCHMC)**

- **Aim 1: Genomics**
  - Rank at least 100 genes for DNA sequence analysis and develop an eMERGE III Targeted Gene Panel (eTGP) with the eMERGE III Steering Committee and other eMERGE III network members.
  - Provide >2000 DNA samples for DNA sequencing on the eMERGE III Targeted Gene Panel and an aggregate of 6,544 data records to dbGAP or the Short Read Archive (SRA).
  - Identify candidate causal somatic variants in 4,000 next generation sequencing (NGS) targeted gene panels obtained in the course of clinical care for patients at CCHMC.
  - Further develop and provide user access to advanced genomic sequence analysis informatics (CASSI).

- **Aim 2: Phenotypes**
  - Develop, validate, and distribute new electronic health record (EHR)-based phenotype algorithms informed by natural language processing using heuristic and machine learning methods.
    - Phenotypes underway include: Appendicitis, Methylphenidate response in Attention Deficit Hyperactivity Disorder (ADHD) and Malignant Hyperthermia.
    - Phenotypes for the theme of pain and its management include: Hypermobility (including Ehlers-Danlos Syndrome), Migraine, Fibromyalgia, Neonatal Abstinence Syndrome (NAS) and Narcotic Dependence.
    - Other phenotypes of present interest include Familial Hypercholesterolemia, Primary Pulmonary Hypertension, and Pyloric Stenosis.
  - Explore the variants in the eMERGE III Gene Panel (eTGP) for candidate phenotypic expression.
  - Pilot demonstration project: Translate pertinent EHR-based algorithms into “executable code” that will run against the PCORnet Common Data Model.
  - Pilot demonstration project: Apply an eMERGE algorithm (e.g., Abdominal Aortic Aneurysm (AAA)) to the Million Veterans Project data and develop an algorithm in the Veterans Affairs EMR (e.g., Chronic Obstructive Lung Disease (COPD)) and bring to eMERGE III network sites.

- **Aim 3: Implementation and Evaluation**
  - Determine the perspectives of patients, guardians, and physicians.
    - Pilot demonstration project: Development of tools to assess adolescent references.
    - Pilot demonstration project: Assessment of adolescent preferences for return of negative incidental genomic results and of electronic portal effectiveness (MyChart).
  - Pilot demonstration project: Evaluate the genetics of outpatient pain after tonsillectomy.
o Pilot demonstration project: Assess the economic impact of CYP3A5 genotype guided tacrolimus dosing.
o Integrate clinical decision support processes into the EHR for identified cohorts, for follow-up recommendations, and for using specific discrete genetic information to trigger consequences.
  ▪ Pilot demonstration project: *PTEN* analysis in patients who are evaluated for possible autism.
  ▪ Pilot demonstration project: Assess outlier yield of testing for Primary Ciliary Dyskinesia (PCD).
  ▪ Demonstration project: Evaluate >2000 eMERGE III Targeted Gene Panel (eTgP) data records for clinically actionable and returnable results; develop and apply workflows for returning results to physicians, and patients or guardians; and provide specific recommendations and clinical decision support tools to physicians and patients.
o Explore and integrate ethical, legal, and social considerations into EHR and clinical care decision making processes as genomic medicine advances, particularly re-evaluation of genomic variation.

Columbia University

- **Aim 1: Advance next-generation phenotyping**
o Designing, validating, and sharing high-throughput, data quality-aware, standards-based phenotyping methods.

- **Aim 2: Perform genetic association studies of rare variants with diverse clinical phenotypes**
o Broad collaboration with the eMERGE network and other phenotyping research communities.

- **Aim 3: Develop practical, scalable learning mechanisms for returning results**
o Leveraging a genomic patient portal and genetic providers to dynamically elicit and incorporate patient preferences for return of genomic results, returning results, and studying patient understanding of returned results.

- **Aim 4: Provide genomic decision support**
o Enhancing and validating our clinical and informatics infrastructure for genomic decision support with learning mechanisms for tailored shared decision-making between patients and health care providers.

Geisinger

- **Aim 1: Use existing biospecimens, genotype and sequence data and EMR-generated phenotypes for discovery in the proposed disorders: familial hypercholesterolemia (FH) and chronic rhinosinusitis (CRS).**
o For FH discovery activities will include: characterization of variants in FH genes chosen for sequencing and correlation with lipid levels across a large, diverse US population; exploration of variants in these genes for extreme lipid phenotypes and Phenome Wide Association Studies (PheWAS) to identify other associated disorders; use of machine learning techniques to develop the most effective population screening algorithm to identify patients likely to have FH; Develop and test economic models of population based screening informed by data accruing from the project. For CRS discovery activities will include: Perform large scale replication studies of the small genetic epidemiology studies of CRS; study the impact of variants in the set of CRS candidate genes selected for sequencing and use for PheWAS; gene environment interactions including Environment Wide Association Studies (EWAS); study similarities and differences between CRS in children and adults.

- **Aim 2: Develop and test approaches for implementation of genomic information in clinical practice.**
• **Aim 3:** Explore, develop and implement novel approaches for family-centered communication around clinically relevant genomic results.
  o Utilizing mixed-methods, explore attitudes of biobank participants toward the familial implications of returned results, including factors that may impede or facilitate communication with other family members. We will subsequently develop and test alternative approaches for notifying family members of genomic results. Deliberative engagement will also be used to inform the envisioned two-phase study and achieve this specific aim. The best practices that result will be implemented for all GHS return of results.

**Group Health Cooperative/University of Washington**

• **Aim 1:** Genomic medicine discovery and implementation focused on CRC/P, TG, and NPC.
  o **Aim 1a:** Collaboratively develop phenotype algorithms for the network, including our own CRC/P and TG phenotypes. This work builds on our E1/E2-derived expertise, including natural language processing (NLP). Our other site-specific phenotype, NPC, is easily derived from WBC/differential phenotypes established in E2.
  o **Aim 1b:** We will have 1000 participants with CRC/P and 1000 Asian ancestry participants sequenced. We will then use eMERGE-wide phenotype and sequence data, as well as collaborators’ CRC/P data, to classify the pathogenicity of variants and to determine the age-dependent penetrance of pathogenic variants for CRC/P genes. Further, using a phenotype-wide association (PheWas) approach, we will determine whether pathogenic variants in these genes are associated with previously unrecognized cancer phenotypes.
  o **Aim 1c:** We propose to investigate rare variants in 2 novel genes for association with TG and NPC and explore new imputation methods to assess rare variant association with these quantitative traits.

• **Aim 2:** Integrate genomic information into GH-wide clinical care and the EMR.
  o **Aim 2a:** We will develop intuitive, comprehensive reports to return genes deemed actionable by the American College of Medical Genetics and Genomics (ACMG). If sequencing should be Clinical Laboratory Improvement Amendment (CLIA) compliant, negative results can be incorporated into the EMR. Medical Genetics will care for participants with actionable results.
  o **Aim 2b:** We will incorporate formal stakeholder input to implement integrated processes and tools to ensure that patients and providers are informed of appropriate actions based on genomic results, developing CDS and online information links for all providers including support for pathogenic variants for CRC/P and long QT syndrome (LQTS). These models can generalize to other genetic disorders.
  o **Aim 2c:** We will also implement processes and tools to support education for primary care and specialty providers, including development and evaluation of educational outreach and online resources.

• **Aim 3:** Evaluate the effectiveness and economic impact of result return to patients and their families.
Aim 3a: We will evaluate the economic impact and cost effectiveness of returning IFs from the ACMG6 list of actionable genes.

Aim 3b: We will test, in a nested randomized controlled trial (RCT), the effectiveness and social and economic impact of an innovative online tool to increase family communication about CRC/P risk and screening aimed at maximizing benefits and reducing harm.

Harvard

- **Aim 1 (Discovery): Hypothesis:** Common and rare variants from a custom chip including 50,000 loss of function (LoF) alleles will be associated with cardiovascular, neuropsychiatric and immune-mediated phenotypes derived from the EMR.
  
  We have analyzed 63,000 whole exome sequences, collected all LoF alleles observed among these individuals, bioinformatically filtered them, and designed a custom genotyping array that includes 50,000 novel LoF alleles as well as an enhanced common marker panel. In 25,000 Partners HealthCare Biobank subjects that will have been genotyped on this custom chip, we will test the association of variants with EMR phenotypes for cardiovascular, neuropsychiatric and immune-mediated diseases. Our goal is to discover clinically actionable genes and variants in Partners HealthCare Biobank subjects.

- **Aim 2 (Penetrance and Pleiotropy): Hypothesis:** Sequencing a set of established genes or loci will allow us to discover additional variation, and define penetrance and pleiotropy using EMR phenotypes.
  
  Genes for sequencing will be chosen from five sources: associations found in specific aim 1, the ACMG list of actionable genes, the PGRN list of actionable pharmacogenetic genes, high impact common CVAS variants, and other eMERGE groups’ nominations. Examples of these genes in our three disease areas might be: a) **LDLR, PCSK9, and APOB** for association with hypercholesterolemia, coronary heart disease, and stroke; b) **DRD2, CACNA1C, TCF4, CHD8** for association with ADHD, BD, depression, and schizophrenia; and c) **HLA-DRB1, IL23RA, TYK2, TNFRSF1A, and CTLA4** for association with RA, IBD, and MS, and asthma. We will study the penetrance and pleiotropic effects of rare variants in these genes by PheWAS, chart review and recall of subjects for additional phenotyping. Finally, genes will be ranked in terms of effect size and population-attributable risk for prioritization of return of research results.

- **Aim 3 (Implementation): Hypothesis:** Physicians will alter their surveillance and treatment of patients based upon voluntary return of actionable variants to provide safe and cost-effective benefits to patients.
  
  To better understand the implications of returning genetic information from a Biobank into the clinical care stream for patients, we will screen our entire Biobank population of 25,000 individuals for pathogenic variants in the gene **LDLR**, the leading genetic cause of premature coronary artery disease, and conduct an exploratory trial in disclosing this information. Biobank participants with pathogenic variants in **LDLR** will be offered enrollment into a randomized trial in which their finding will be CLIA confirmed, and in one arm, this result will be communicated to their physicians through the EMR. Over one year, we will collect the following outcomes through participant surveys and EMR queries: physician visits, laboratory testing, changes in medication prescriptions, low-density lipoprotein cholesterol levels, medical costs and the number of family members screened and treated as a result of the intervention. We will collaborate with the entire eMERGE III Network to incorporate what we learn from this pilot trial into large-scale implementation protocols for the genes selected by the Network.
Mayo Clinic

- **Aim 1: Classify variant pathogenicity and develop clinical decision support for actionable variants discovered by sequencing of a panel of 100 genes including genes related to FH/CRC.**
  - We will: a) use state-of-the-art bioinformatics pipelines for variant annotation and assign pathogenicity and actionability to variants through an expert multidisciplinary variant curation group; b) evaluate variants of uncertain significance (VUS) by investigating associations with EHR-ascertained phenotypes; and c) develop EHR-based clinical decision support for actionable variants to facilitate clinical decision-making at point of care.

- **Aim 2: Return clinically actionable results including those relevant to FH/CRC and assess the subsequent patient outcomes, costs and utilization of health care resources, and behavioral and psychosocial effects.**
  - Building on methods we developed in eMERGE II, we will assess: a) impact of genomic results on patient outcomes, including diagnostic and therapeutic interventions and new case detection in families; b) costs and utilization of health care resources following return of results; and c) understanding of genomic results, concerns about including genomic data in the EHR, worries about future employability or insurance coverage, and views regarding sharing genomic results with family members and others.

- **Aim 3: Perform EHR-based phenotyping and genomic discovery**
  - We will use validated electronic algorithms including natural language processing, to extract phenotypes from the Mayo and MPHC EHRs. For the sequence data set (n= 25,000 across the network), our phenotypes of interest are plasma lipid levels. We will work with the FH Foundation and other eMERGE sites to develop an EHR-based algorithm for identifying FH cases. For the high-density genotyping data set (n= ~50,000 across the network) our phenotypes of interest include adverse drug reactions and treatment response. We will: a) perform common and rare variant association analyses for phenotypes of interest; b) explore penetrance using family data from cascade screening for FH/CRC; and c) assess the pleiotropic effects of genetic variants using PheWAS and additional novel approaches.

Northwestern University

- **Aim 1: Perform electronic phenotyping and genomic analysis on at least four novel phenotypes of medical, scientific and public health importance.**
  - See Section 3 for contact information. We will use both the eMERGE GWAS dataset (common variants) and deep DNA sequence of at least 100 genes likely to contain medically actionable variants (rare variants) to identify novel phenotype-genotype associations. New phenotyping algorithms for cardiovascular, gastrointestinal and neurological phenotypes, plus chronic rhinosinusitis will be developed at NU. Those developed at other sites will be implemented in the NU environment. We propose deep sequence analysis of candidate genes associated with our phenotypes and those recommended for reporting by the ACMG and some CPIC guidelines. We will deploy modular phenotyping algorithms and a PheWAS (2) approach that may allow sub-classification of conditions associated with the sequenced genes in addition to the broader collection of imputed variants available in the combined eMERGE I and II dataset.
• **Aim 2: Consent at least 2000 ethnically diverse participants who agree to receive results about clinically actionable variants found in their genotyped genes and return appropriate information to them.**
  o Participants will be enriched for the possibility of carrying an actionable genetic variant in one or more of the sequenced genes. We will implement strategies to identify actionable results using resources such as ClinGen and ClinVar, and ACMG and CPIC guidelines and return them in a clinical setting. A learning healthcare systems approach will be used in returning results that engages participants through capturing family history and other valuable outcome information that may aid in clinical care.

• **Aim 3: Expand our ancillary genomics system (AGS) to incorporate sequence data, including whole genome and exome sequence, and develop strategies for transferring actionable variants into EHR to enable decision support.**
  o We will assess the implications of this model on institutional policy, clinical care and research.

• **Aim 4: Develop and share best practices, expertise and experience for:**
  o Creating tools and methods for rapid and accurate phenotyping; (b) assisting patients and families in making choices about genomic results; (c) sharing of clinical and family history data with families; (d) storing and representing genomic data in the EHR; and (e) understanding downstream effects of genomic sequencing results on the healthcare system, patient and family.

**Vanderbilt University**

• **Aim 1: To leverage EMR-based phenotyping to define genetics of disease and disease subsets.**
  o We will expand the network’s phenotyping library by creating increasingly granular phenotype definitions that identify specific subsets of disease with predictable clinical courses or response to therapies. These will be analyzed by GWAS using existing eMERGE genotypes. The increased phenotype granularity will also enable PheWAS for replication here and for discovery in Aim 2.

• **Aim 2: To identify rare variants with strong associations with human traits by sequencing 100 genes in 25,000 subjects across eMERGE-III.**
  o We propose that the 100 genes be drawn from (1) the ACMG list of 56 genes with variants known to confer large effect sizes for important human diseases; (2) eMERGE PGx genes; and (3) genes with very strong associations in PheWAS we have conducted using exome chip data in nearly 30,000 BioVU records to analyze associations with disease codes and lab values. We will use both genotype- and phenotype-driven methods to search for new associations between sequence variants and disease. Actionable sequence variants will be returned to subjects and their EMRs in Aim 3.

• **Aim 3: To expand PREDICT, our pre-emptive pharmacogenomic implementation program, to deliver actionable variants to patients and providers and to assess their response.**
  Specific Aims 1 and 2 will expand the range of genomic or phenotypic information that could be considered relevant to an individual patient’s care (“actionable”). We have developed an initial framework for delivering such information in the flow of healthcare within the PREDICT program at Vanderbilt. Our initial focus has been on pharmacogenomic variation, and we will collaborate across eMERGE to develop, implement, and assess tools to deliver new information, measuring impact to ensure optimal benefit to patients.
Partners – Broad Institute

• **Aim 1: Assay Design, Sequence Generation and Analysis.**
  o We will design and clinically validate a targeted eMERGE III gene panel using Broad’s well-established clinical next generation sequencing (NGS) platform as well as additional guidance from LMM’s many years of experience in rapid development and validation of CLIA-certified NGS gene panel tests. The LMM will also leverage its expertise in evaluating gene-disease relationships as well as its linkage with relevant ClinGen expert committees to advise on the contents of this gene panel. CLIA grade sequencing, alignment and variant calling will be performed at the Broad using software that is most widely used and at the leading edge of NGS analysis. Raw data as well as DNA aliquots will be immediately available to LMM for downstream processing.

• **Aim 2: Orthogonal Confirmation, Interpretation and Report Generation.**
  o Variant calls and raw data will flow from the Broad Institute to the eMERGE Coordinating Center (CC) as well as into LMM’s clinical grade NGS tertiary analysis pipeline. Variant filtering will be done using criteria established with the eMERGE Steering Committee (SC). Variants will be validated to reach diagnostic confidence using orthogonal technologies as necessary (Sanger sequencing, ddPCR) and clinical reports will be created using LMM’s CLIA processes, including its gene/disease/variant knowledge base, as well as automated report drafting infrastructure. We propose to create a default clinical report for all genes in the eMERGE panel. We propose to report all variants with established clinical significance but are prepared to customize reports according to site specific needs and adapt them over time as gene content and return of results approaches evolve.

• **Aim 3: Return of results, novel discovery approaches and interaction with the scientific and clinical genomics community.**
  o Standard clinical reports and subsequent knowledge updates will be transmitted electronically to the sites via site specific instances of GIC, which will link to LMM’s variant knowledgebase. An eMERGE-specific instance of GIL will **a)** provide sites with direct real-time access to all genomic knowledge within LMM’s genomic knowledge base, **b)** provide linkage to an existing variant and knowledge sharing network of multiple US and Canadian labs, and **c)** enable reciprocal exchange of variant data between eMERGE and ClinVar. To enable novel discovery analyses, primary data and variant annotations will be disseminated to the CC and/or to a portal located at the Broad Institute housing genomic analysis tools. We will liaise with the greater scientific and clinical genomics community via our established roles in the following NIH consortia: Clinical Genome Resource (ClinGen), Clinical Sequencing Exploratory Research (CSER), Newborn Sequencing In Genomic medicine and public HealTh (NSIGHT), Undiagnosed Diseases Network (UDN) and through our roles in organizations developing standards in genomics or coordinating the rollout of clinical genetic IT support including the American College of Medical Genetics and Genomics (ACMG), Association for Molecular Pathology (AMP), College of American Pathology (CAP), Global Alliance for Genomics and Health (GA4GH), Health Level Seven International (HL7), Human Gene Variation Society (HGVS), International Rare Diseases Research Consortium (IRDiRC) and Institute of Medicine (IOM) to establish knowledge feeds and share genomic data and best practices between eMERGE and these bodies.
Baylor College of Medicine Human Sequencing Center

• **Aim 1:** Provide DNA sequence from approximately 100 selected gene targets in a CAP and CLIA certified environment to advance the aims of the eMERGE network.
  - DNA samples will be received, archived and genotyped to enable sample tracking and QA/QC. Complete sequences of the defined loci will be generated by a custom DNA capture panel and high quality DNA sequence variants, including structural variants, will be identified by the BCM clinical analytical pipeline. Actionable variants, including but not limited to currently known pathological disease variants and the ACMG recommended incidental findings within the targeted areas, will be validated by Sanger sequencing. A second round of validation will be offered for variants that are nominated by eMERGE Network members. Designs can be iterated, at least annually.

• **Aim 2:** Communicate data to the eMERGE coordination center, with standardized clinical annotations and advanced data security measures.
  - Raw data and derived files will be archived for 2 years for the benefit of the consortia and submissions to the public databases such as dbGAP and CLinGen will be offered.

• **Aim 3:** Progressively provide alternative strategies for sequencing so that the eMERGE Steering Committee can determine the most efficient and cost-effective pathway to molecular diagnoses.
  - The sequencing core will apprise the eMERGE network of technical developments that may influence the Networks’ choice of technologies and directions and so economics and other factors that favor alternate approaches can be considered. As a part of this collaborative effort, we will contribute expertise and leadership in the areas of efficient high throughput sequencing, data analysis/interpretation, and clinical reporting, while participating actively with the Steering Committee and network community to continually improve data generation.

Vanderbilt University Serving as the Coordinating Center

• **Aim 1:** Accelerate EMR phenotyping and enable end-to-end genomic medicine research.
  - In Phase III, we propose advancing current infrastructure by expanding upon current tools and exploring new resources for electronic phenotype algorithm development and data sharing. We propose enabling broader PheWAS use with additional services and new methods. In addition, we propose establishing the Genomic Actionability and Implementation Repository (GenAIR), a collaborative development space and central repository for collection and dissemination of best practices and standards related to genomic medicine implementation.

• **Aim 2:** Integrate high quality genomic information across eMERGE sites.
  - The University of Washington (UW) team (led by Drs. Crosslin and Jarvik) will manage genomic data for eMERGE III. The group has extensive expertise providing sequencing data production and analysis support for many sequencing and genotyping studies. We have interacted and propelled many different studies through our involvement in the NIGMS Pharmacogenomics Research Network, the NHGRI Clinical Sequencing Exploratory Network, NHGRI/NHLBI Centers for Mendelian Genomics, the NHLBI Resequencing and Genotyping Project and Exome Sequencing Project (ESP), and many others. For Phase III, we will manage the following activities: 1) provide continued access to all existing genotype data for eMERGE data by download and improve on these data by providing new imputations to the expanding whole genome sequencing datasets; 2) transition the existing SPHINX database to a format similar to the exome variant server that is easy to use and familiar; 3) aggregate PGRNseq data to perform new network wide studies; 4) aggregate the sequence data produced in phase III by recalling the data and delivering multiple sample variant calls and
annotation for single nucleotide variant, indel, and structural variant analysis; and 5) assist with rare variation analysis for network wide projects using this new sequencing data.

- **Aim 3: Secure EMR and genomic data sharing risk mitigation.**
  - In Phase III, we propose to share tools and expertise that enable secure genomic medicine implementation while maintaining robust data access for research. We will continue to pursue research methods, tools and novel approaches that inform privacy and security policies around genomic sequencing and EMR data sharing. To bring privacy and data security issues and opportunities to the diverse investigators in the eMERGE Network, we propose a workshop for eMERGE participants on EMR and Genomic Data Privacy and Security and Clinical Data Management to be led by Brad Malin, PhD and Paul Harris, PhD (see Aim 5). By focusing efforts on privacy and anonymization policy guidance, novel methods research and real-world site implementation guidance, we will ensure eMERGE retains its reputational strength and leadership position within the field of EMR and genomic data sharing risk mitigation.

- **Aim 4: Provide excellent logistical support to the entire Network; including committees, work groups, NHGRI, and the ESP.**
  - We will draw heavily upon existing infrastructure, personnel and processes to assist the Network in accomplishing Phase III program goals. Specifically, we will continue to refine Network communications, organizational and functional support models to optimize engagement of a broad range of investigators, facilitate work, enhance learning and catalyze new ideas.

- **Specific Aim 5: Synergize with other related networks; to eliminate redundancy, to promote crosspollination of best practices, and to share eMERGE tools more broadly.**
  - The productive interactions with PGRN and CSER in Phase II resulted in large part from fortuitous collegiality among like-minded investigators involved in related research networks. We will use existing information-sharing tools to encourage, foster, and possibly seed more inter-network interaction. Project managers will follow up with project planning assistance for nascent collaborations or additional tool implementation advice to foster cross-network interactions. All network projects will continue to be governed according to each network’s process, so for eMERGE, the Steering Committee will still need to approve any work deriving from these activities. As such, the success of this work will depend on cooperation between the operations groups and investigators from the different networks.
# ESP MEETING MINUTES: May 2015

**eMERGE Network**  
**ESP Conference Call Minutes**  
May 4, 2015 - 4:00 p.m. ET (3:00 p.m. CT/1:00 p.m. PT)

**Attendees:**
- CCHMC/BCH: John Harley, Ingrid Holm; CHOP: Hakon Hakonarson; Geisinger: Marc Williams, David Carey; Group Health/UW: Gail Jarvik, Eric Larson; Marshfield/Essentia/PSU: Cathy McCarty, Peggy Peissig; Mayo: Iftikhar Kullo; Mt. Sinai: Erwin Bottinger; Northwestern: Rex Chisholm, Maureen Smith, Laura Rasmussen-Torvik; VU: Dan Roden, Sarah Stallings; NHGRI: Rongling Li, Teri Manolio, Ken Wiley, Jackie Ogdis; CC: Paul Harris, Melissa Basford, Adam Hardebeck, Mollie Bodin-Claar; ESP: Howard McLeod (Moffitt Cancer Center), Eta Berner (University of Alabama-Birmingham), Gerardo Heiss (UNC), Lisa Parker (University of Pittsburgh)

## Opening Remarks
- Rongling and Rex welcomed the ESP members, and Howard McLeod complimented the group for their work on the eMERGE Network ESP preparation packet.

## Network Overview & Response to ESP Recommendations
- Recommendations from the ESP Packet were reviewed.
- The Network’s publications have been cited 7,786 times since Phase I (2,704) since Phase II. The network’s focus on network wide projects during 2015 was reaffirmed, as over 100 network manuscripts are being developed.
- The ESP noted that local IT integration can be a major challenge.
- ESP members inquired about any feedback received from clinicians related to the Infobutton project, and Marc explained that the project is in the CDS data collection process, but early analysis points to minimal pushback from clinicians.
- The ESP also inquired about eMERGE’s experience with the correlation of genomic CDS to non-genomic CDS. They noted that the alert override audience doesn’t feel as knowledgeable about CDS. Marc noted that besides taking longer to explain genomic matters to clinicians/admin, the group hasn’t noticed significant difference between the two. The group will have more detailed answers after the CDS data analysis.

## Network Workgroup Update: CERC Survey
- Maureen Smith and Ingrid Holm provided an overview of the CERC Survey project. On April 8th, 90,000 participants were mailed a survey. The survey currently has a response rate of 7%. A second wave of reminder cards and surveys will be circulated to those that did not initially respond. The survey analysis will extend beyond the end of Phase II.
- The pilot survey was also reviewed. The pilot had an overall response rate of 12% out of 1,500 targeted patients. It was also noted that oversampling of minority and low education groups helped provide a more demographically accurate representation of the study.
- The ESP inquired about the CERC Survey target cohort. The workgroup co-chairs confirmed that the survey recipients are patients from eMERGE institutions, but are not necessarily participants in eMERGE biobanks.
- It is not possible to know the overlap between survey recipients and eMERGE participants at all sites, but this
information is collected where possible. This info was not collected in the pilot survey.

Network Workgroup Update: PGx

- Sequencing is on track for completion by the end of Phase II.
- The ESP expressed interest in evidence of the comfort level with precision prescribing for providers. Provider Education metrics are largely qualitative, and implementation has taken place at various sites, making it difficult to gather consistent quantitative data.

**ACTION ITEM:** The Provider Education team was advised to investigate methods to collect more qualitative data about providers' precision prescribing comfort level.

Conclusion

- The ESP complimented the group for their productivity throughout Phase II, and for continuing to make progress throughout the grant.

Action Items

1. The Provider Education team was advised to investigate methods to collect more qualitative data about providers' precision prescribing comfort level.

### NHGRI-ESP Executive Session Minutes

5:00-5:30 p.m. ET (4:00-4:30 p.m. CT/2:00-2:30 p.m. PT)

**Attendees:**

ESP: Eta Berner (UAB) Gerardo Heiss (UNC), Howard McLeod (Moffitt, Chair), Lisa Parker (Pittsburgh); **NHGRI:** Kongling Li, Teri Manolio, Jacqueline Odgis

<table>
<thead>
<tr>
<th>Executive Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>- The External Scientific Panel (ESP) met with members of NHGRI staff in Executive Session after the ESP teleconference held on May 4, 2015.</td>
</tr>
<tr>
<td>- The ESP was pleased with how responsive the Network has been to the ESP’s recommendations from the last joint ESP-Steering Committee meeting. The ESP was also impressed with how well Network investigators were able to anticipate the ESP’s questions in their presentations.</td>
</tr>
<tr>
<td>- The ESP appreciated the quality of the packet for this teleconference, as well as the increased amount of information on the progress of network-wide projects in the Steering Committee’s report.</td>
</tr>
<tr>
<td>- The ESP noted that the Network seems to have many activities to complete before Phase II concludes.</td>
</tr>
<tr>
<td>- Regarding the CERC Survey project, the ESP was concerned about the low (~7%) response rate and the complexity of the concepts assessed, given the anticipated knowledge and reading level of the participants. NHGRI staff noted that the CERC Survey Workgroup has addressed the projected low response rate by piloting and modifying the survey multiple times throughout its development to refine the survey language.</td>
</tr>
</tbody>
</table>
and increase the number of responders, particularly participants with low education levels. Recognizing that the survey has already been implemented, it was suggested that, for future surveys, the CERC Survey Workgroup consider simplifying the biobank scenario vignettes presented in the survey to maximize participation. If information on participants' biobank status is available, the ESP wanted the CERC Survey Workgroup to compare responses from survey participants enrolled in biobanks to responses from participants who have never been enrolled in biobank research. As a potential Phase III project, the ESP also recommended integrating a study of consent approaches into actual enrollment into biobanks, i.e., by randomizing participants to different currently used approaches for consent.

- The eMERGE-FQx Workgroup seems to be moving forward in conducting this project; however, overall success for the project will depend on whether the participants in this study are using the selected drugs or have the desired variants. The workgroup has effectively demonstrated its ability to obtain and integrate results into EHR for clinical use. The ESP agreed that it will be worthwhile for the eMERGE-FQx Workgroup to continue harmonizing multiple platforms across multiple sites as this reflects real-world platform diversity among hospital systems. It was suggested that data comparing EHR integration of results across different platforms should be made publically available.

### ESP Recommendations

1. The CERC Survey Workgroup should consider, for future surveys, simplifying the biobank scenario vignettes to maximize participation.
2. If information on participants' biobank status is available, the CERC Survey Workgroup should compare responses from survey participants enrolled in biobanks to responses from participants who have never been enrolled in biobank research.
3. For Phase III, the CERC Survey Workgroup should consider integrating a study of consent approaches into actual enrollment into biobanks, e.g., by randomizing participants to different currently used approaches for consent. The primary recommendation is to study actual attempts to recruit and enroll biobank participants using different approaches, rather than engaging in study of hypothetical scenarios.
4. The eMERGE-FQx Workgroup should continue to harmonize multiple platforms across multiple sites as this reflects real-world platform diversity among hospital systems.
5. The eMERGE-FQx Workgroup should make the data comparing EHR integration of results across different platforms publically available.