emerge network
ELECTRONIC MEDICAL RECORDS & GENOMICS

External Scientific Panel

background materials

Boston Children's Hospital
The Children's Hospital of Philadelphia
CIDR
Cincinnati Children's Hospital Medical Center
Essentia Health Here with you
GEISINGER
GroupHealth
MARSHFIELD CLINIC
MAYO CLINIC
Mount Sinai
NORTHWESTERN UNIVERSITY 1851
PENNSTATE 1855
UNIVERSITY OF WASHINGTON 1861
VANDERBILT UNIVERSITY

ESP Conference Call
May 4, 2015

Background Information Table of Contents

eMERGE Network Collaborations

Boston Children's Hospital

1. Joris A. Veltman, Nijmegen Centre for Molecular Life Sciences, Radboud University Nijmegen Medical Centre: Trio Exome Sequencing
2. Children's Hospital of Philadelphia: Asthma and ADHD (both led by CHOP); Autism (led by CCHMC/BCH)
3. Vanderbilt (Dr. Sara Van Driest) /CHOP: Early Childhood Obesity (led by CCHMC/BCH)
4. The International Consortium for the Genetics of Systemic Lupus Erythematosus (SLEGEN): Harley, CCHMC member site
5. Dr. Hutton (CCHMC) leads the consortium to federate studies based on the EMR across 30 centers that treat children with Inflammatory Bowel Disease (R01 HS020024).
6. Dr. Louis Kunkel and Dr. Christopher Walsh (BCH) Co-PIs with co-investigators Dr. Isaac Kohane and Dr. Ingrid Holm: Phenotypic and Genetic Factors in Autism Spectrum Disorders. This study led to BCH designation as a Simons Simplex Collection site and a member of the Boston-area Autism Consortium.
7. Dr. Amy Roberts (BCH): PI Cardiac Gene Project.
8. Dr. Alan Beggs (BCH): Director Manton Center for Orphan Disease Research (Dr. Ingrid Holm and Dr. Isaac Kohane - Gene Discovery Core scientists). This center frequently collaborates with the Gene Partnership program, from which BCH's eMERGE cohort originates.
9. Dr. Ingrid Holm at BCH leads Informed Cohort Oversight Board – multidisciplinary team from multiple institutions to oversee research functions and communication of results back to patients; Return of Results Consortium – funded by, and part of, the NHGRI ELSI Research Program.
10. Dr. Isaac Kohane: Led the collaborative development of i2b2 (piloted across hospitals in the Harvard Catalyst) which has been adopted by over 60 academic health centers internationally. Supports an Academic Users' Group of over 250 members from over 65 independent institutions that meets biannually for code workshops, discussion of application issues, preview of coming software, and networking; New efforts directly related to i2b2 have resulted in at least 27 new collaborative grants. Established Summer Institute in Bioinformatics and Integrative Genomics – in partnership with the Harvard-MIT Division of Health Sciences and Technology (HST).
11. Dr. Ingrid Holm at BCH is a co-investigator on a newly-awarded U19 cooperative agreement – "Genome Sequence-Based Screening for Newborn Illness and Childhood Risk". The Principal Investigators for this study are Dr. Robert Green of Brigham and Women's Hospital and Dr. Alan Beggs of BCH.
12. Dr. Isaac Kohane is the PI of the Undiagnosed Diseases Network (UDN) Coordinating Center and Dr. Ingrid Holm is a co-investigator and overseeing the enrollment of patients for the UDN.
13. In addition to acting as the coordinating center, BCH is also a joint site with Brigham and Women's and Massachusetts General Hospital for the UDN.
14. Dr. Guergana Savova (BCH) is involved in various multisite collaborative studies, including *Cancer Deep Phenotype Extraction from EMR*; *Annotation, development, and evaluation for clinical information extraction; Temporal Relation Discovery for Clinical Text; Genetic predictors of response to anti-TNF therapy in rheumatoid arthritis.*
15. Dr. John Harley (CCHMC) is the Director of the Cincinnati Biobank Core Facility and the PI of the Better Outcomes for Children Project at Cincinnati Children's to consent patients to use their leftover samples for research.
16. Marc Rothenberg, MD, PhD (CCHMC) shared GWAS data and samples from his Eosinophilic Esophagitis cohort.
17. Dr. Sue Thompson (CCHMC) shared GWAS data and samples from her Juvenile Idiopathic Arthritis (JIA) cohort.
18. Dr. John Harley (CCHMC) has overseen the Cincinnati Control Cohort at Cincinnati Children's; ~1000 local controls, ~850 of which are of European ancestry; all have been extensively phenotyped and have been genotyped on the Affymetrix 6.0 platform
19. Dr. Teresa Smolarek (CCHMC); Cytogenetics Lab Director
20. Dr. Senthil Sadhasivam (CCHMC); PI of the post-surgical pain management protocol;
21. Dr. Patty Manning (CCHMC), Co-Director of Developmental & Behavioral Pediatrics and the Susan Wiley Center, champion for the Autism phenotype
22. Dr. Stephanie Kennebeck (CCHMC); champion for Appendicitis
23. Dr. Nancy Crimmins (CCHMC); champion for Early Childhood Obesity
24. Dr. Keith Marsolo (CCHMC); Co-Investigator on grants supporting the following quality improvement and learning networks - ImproveCareNow (62 centers), Ohio Perinatal Quality Collaborative (85 centers), and Solutions for Patient Safety (79 centers)

25. Dr. Keith Marsolo (CCHMC); Co-Investigator on a grant with Bill Nichols to develop a tool that will allow investigators to query and request both case report form and sample data for a patient cohort with Pulmonary Arterial Hypertension.
26. Dr. Keith Marsolo (CCHMC); Co-Investigator on 3 unique contracts to build the PCORI (Patient Centered Outcomes Research Institute) National Patient-Centered Clinical Research Network (one with CCHMC as prime, another with BCH as prime, another with CHOP as prime).
27. Cindy Prows, MSN, APN, FAAN (CCHMC); leads the PGx supplement
28. Dr. Melanie Myers (CCHMC); leads the CERC supplement
29. Dr. Kenneth Kaufman (CCHMC) – collaboration in establishing exome sequencing
30. Dr. Sander Vinks (CCHMC) – Director of Clinical Pharmacology; ongoing pharmacogenomics consultation
31. Dr. Bill Nichols (CCHMC) ; PI of an NLHBI R24 to collect 3,500 Pulmonary Arterial Hypertension patients

Cincinnati Children's Hospital Medical Center

1. **Nijmegen Centre for Molecular Life Sciences**, Radboud University Nijmegen Medical Centre: Trio Exome Sequencing (Joris A. Veltman, PhD).
2. **Children's Hospital of Philadelphia (CHOP)**: Asthma, ADHD and Atopic Dermatitis (both led by CHOP); algorithm validation by CCHMC/BCH.
3. **Vanderbilt** (Sara Van Driest, MD) and **CHOP** (John Connolly, PhD): Autism Algorithm Validation
4. **Geisinger** (Marc Williams, MD and Lisa Bailey-Davis, DEd.; site specific implementation ideas and contribution to publication) & **CHOP** (John Connolly, PhD): Early Childhood Obesity validation.
5. **The International Consortium for the Genetics of Systemic Lupus Erythematosus (SLEGEN)**: Harley, CCHMC member site.
6. **"Building Modular Pediatric Chronic Disease Registries for QI and CE Research"**; PI, John Hutton, MD (CCHMC) led the consortium to link EHRs to disease-specific registries across 30 centers that treat children with Inflammatory Bowel Disease (R01 HS020024).
7. **Phenotypic and Genetic Factors in Autism Spectrum Disorders**: Louis Kunkel, PhD and Christopher Walsh, MD, PhD (BCH) Co-PIs with co-investigators Isaac Kohane and Ingrid Holm on this study led to BCH designation as a Simons Simplex Collection site and a member of the Boston-area Autism Consortium.
8. **PI Cardiac Gene Project:** Amy Roberts, MD (BCH)
9. **Manton Center for Orphan Disease Research:** Alan Beggs, PhD (BCH) is the Director and Drs. Ingrid Holm and Isaac Kohane are Gene Discovery Core scientists. This center frequently collaborates with the Gene Partnership program, from which BCH's eMERGE cohort originates.
10. **Informed Cohort Oversight Board**: Ingrid Holm, MD, MPH at BCH leads this multidisciplinary team from multiple institutions to oversee research functions and communication of results back to patients; Return of Results Consortium – funded by, and part of, the NHGRI ELSI Research Program.
11. **i2b2 (Informatics for Integrating Biology and the Bedside)**: an NIH-funded National Center for Biomedical Computing based at Partners HealthCare System led by Isaac Kohane, MD, PhD, has been adopted by over 60 academic health centers internationally. The i2b2 framework has been extended by the **CCHMC i2b2 team**, adding several new capabilities to the platform. These include the ability to view clinical data in a web-based form (similar to a chart review), the ability to enter data directly into i2b2, and the ability to run reports and perform other visualizations. These new features allow i2b2 to serve as a platform for research patient registries (either identified or de-identified), and when coupled with the SHRINE federated query platform, provide a mechanism for creating distributed, multi-center registries.
12. **"Genome Sequence-Based Screening for Newborn Illness and Childhood Risk"**: PI Ingrid Holm, MD, MPH at BCH is a co-investigator on this newly-awarded U19 cooperative agreement. The Principal Investigators for this study are Dr. Robert Green of Brigham and Women's Hospital and Dr. Alan Beggs of BCH.
13. **Undiagnosed Diseases Network (UDN)**: Isaac Kohane, MD, PhD is the PI of the Undiagnosed Diseases Network (UDN) Coordinating Center and Ingrid Holm, MD, MPH at BCH is a co-investigator and overseeing the enrollment of patients for the UDN.
14. **Guergana Savova, PhD** (BCH) is involved in various multisite collaborative studies, including *Cancer Deep Phenotype Extraction from EMR*; *Annotation, development, and evaluation for clinical information extraction; Temporal Relation Discovery for Clinical Text; Genetic predictors of response to anti-TNF therapy in rheumatoid arthritis.*

15. **Cincinnati Biobank**: John Harley, MD, PhD (CCHMC) is the Director of the Cincinnati Biobank Core Facility and the PI of the **'Better Outcomes for Children' Project** at Cincinnati Children's to consent patients to use their leftover samples for research. To date, a collection of >100,000 DNA samples from >35,000 unique MRNs and >25,000 urines from >4,300 unique MRNs has been built.
16. **The Cincinnati Center for Eosinophilic Disorders (CCED)**: Director, Marc Rothenberg, MD, PhD (CCHMC), shared GWAS data and samples from his Eosinophilic Esophagitis cohort.
17. **Pediatric Rheumatology Tissue Repository (PRTR)**, a funded component of a P30 rheumatic disease core center, assists local and national investigators in the collection, processing, storage and distribution of biospecimens for rheumatic disease-related research. Director, Sue Thompson, PhD (CCHMC), shared GWAS data and samples from her Juvenile Idiopathic Arthritis (JIA) cohort.
18. **Cincinnati Control Cohort**: John Harley, MD, PhD (CCHMC) oversees the Cincinnati Control Cohort at Cincinnati Children's; ~1000 local controls, ~850 of which are of European ancestry; all have been extensively phenotyped and have been genotyped on the Affymetrix SNP Array 6.0 and Illumina HumanOmni5.0 platforms.
19. **CCHMC Cytogenetics Lab**: Teresa Smolarek, PhD (CCHMC), is Cytogenetics Lab Director. This lab combines state-of-the-art techniques with comprehensive interpretation of test results by experienced, board-certified cytogenetics experts. The Cytogenetics Lab has provided data for the eMERGE project work.
20. **Senthil Sadhasivam, MD** (CCHMC): 1. PI of the post-surgical pain management protocol with preemptive and preoperative genotyping and CYP2D6 guided opioid managment; 2. PI of a large morphine pharmacokinetic and pharmacogenetic clinical study in children and 3. PI of a multicenter opioid pharmacogenetic study with 1000 children from 15 participating sites in USA and China.
21. **Autism Clinician Experts**: Patty Manning, MD (CCHMC), Co-Director of Developmental & Behavioral Pediatrics and the Susan Wiley Center, and Julie Bickel, MD (BCH); champions for the Autism phenotype.
22. **Appendicitis Clinician Experts**: Stephanie Kennebeck, MD (CCHMC) and Amir Kimia, MD (BCH); champions for Appendicitis.
23. **Early Childhood Obesity Clinician Experts**: Nancy Crimmins, MD (CCHMC), Cassandra Brady, MD (BCH) and Vidhu Thaker, MD (BCH); champions for Early Childhood Obesity.
24. **Keith Marsolo, PhD** (CCHMC eMERGE Key Personnel); Co-Investigator on grants supporting the following quality improvement and learning networks - ImproveCareNow (62 centers), Ohio Perinatal Quality Collaborative (85 centers), and Solutions for Patient Safety (79 centers); Co-Investigator on a grant with Bill Nichols to develop a tool that will allow investigators to query and request both case report form and sample data for a patient cohort with Pulmonary Arterial Hypertension; Co-Investigator on 3 unique contracts to build the PCORI (Patient Centered Outcomes Research Institute) National Patient-Centered Clinical Research Network (one with CCHMC as prime, another with BCH as prime, another with CHOP as prime).
25. **Genetics Expert Panel**: Cindy Prows, MSN, APN, FAAN (CCHMC) is Chair of the Genetics Expert Panel at CCHMC.
26. **Cincinnati Analytical Suite for Sequencing Informatics**: Developers Kenneth Kaufman, PhD and John Harley, MD, PhD (CCHMC) – active collaboration in establishing exome sequencing
27. **Sander Vinks, PharmD, PhD, FCP** (CCHMC) – Director, Division of Clinical Pharmacology; ongoing pharmacogenomics consultation.
28. **National Biological Sample and Data Repository for Pulmonary Arterial Hypertension (PAH)**: Bill Nichols, PhD (CCHMC) is the PI of an NLHBI R24 to collect 3,500 Pulmonary Arterial Hypertension patients; Dr. Nichols is also Interim Director of Human Genetics at CCHMC.
29. **'Methylphenidate in ADHD' project**: Tanya Froehlich, MD, MS (CCHMC) is the Clinician Expert of the 'Methylphenidate in ADHD' project.
30. **Center for Pediatric Genomics (CpG)**: John Harley, MD, PhD and Peter White, PhD, Professor and Chair, Biomedical Informatics, UC & CCHMC, are Co-Chairs of this Center at CCHMC.

Children's Hospital of Philadelphia

1. **Neurodevelopmental Genomics: Trajectories of Complex Phenotypes**: The goal of this study is to perform comprehensive neurocognitive phenotyping on 10,000 children who are already genotyped and perform methylation

profiling, imaging and establish EBV cell lines on a subset of them. This will be made available as a public resource. (PIs: H. Hakonarson, CHOP and R. Gur,UPenn)

2. **1,000 Rare Diseases Project to Advance Gene Discovery:** The project employs integrative genomic approaches and analysis pipelines focusing on sequencing 1,000 rare diseases, including ones that affect both children and adults. It uses next-generation sequencing (NGS) technologies to analyze well characterized DNA samples from patients and families with single-gene inheritance patterns. (Co-Directors: H. Hakonarson, CHOP and Jian Wang, BGI)
3. **CHOP/PENN Center for Excellence for Autism Research:** This is a comprehensive study of neuroanatomical features of children with ASD. It aims to characterize brain structure, connectivity (using diffusion tensor imaging), and brain function to investigate possible causal mechanisms for the heterogeneity in ASD. (Co-PI H. Hakonarson, CHOP)
4. **Cholestatic Liver Disease Consortium:** An ancillary study, Genetic Modifiers of Liver Disease Severity in Alagille Syndrome is designed to identify genetic factors associated with liver disease severity in Alagille syndrome. (PI, N. Spinner, CHOP)
5. **Brain, Behavior and Genetic Studies of 22q11.2 Deletion Studies**: The objective of this collaboration between CHOP and Penn is to combine genetic and neurobiologic paradigms for understanding pathogenesis and for detection of genes that modulate susceptibility to psychosis with phenotypic features of schizophrenia and related disorders (SCZ). (PI, B. Emanuel, CHOP)
6. **Alzheimer's Disease Genetics Consortium:** The goal of this study is to perform a whole genome scan to test for association of AD to a high-throughput tag-SNP arrays. (PI, G. Schellenberg, UPenn; PI-CHOP, H. Hakonarson)
7. **Pediatric Cardiac Genomics Consortium:** We are studying the genetic basis of conotruncal defects. Genetic risk factors for disease are identified using genome wide association studies and copy number variant analyses. An inception cohort with tetralogy of Fallot would be ascertained. (PI, E. Goldmuntz, CHOP; PI on genotyping network facility: H. Hakonarson)
8. **NICHD multi-center, multi-ethnic longitudinal Bone Mineral Density in Childhood Study:** We are conducting a genome-wide association study of bone mineral accretion during childhood. Our goal is to identify the genetic determinants of bone health early in life with the ultimate goal of identifying new pathways for osteoporosis prevention. (Co-PI, S. Grant, CHOP and B. Zemel, CHOP)
9. **Clinical Sequencing Exploratory Research**: The Applying Genomic Sequencing in Pediatrics group is bringing genomic sequencing into pediatric clinical settings. Researchers will work with families, scientists, and ethicists to determine how patients should be counseled and educated before testing, what data should be provided back to individual families, and what educational tools will help in understanding the implications of the testing. (Co-PIs, I. Krantz, CHOP and N. Spinner, CHOP)
10. **3/3 Networks from Multidimensional Data for Schizophrenia and Related Disorders:** Overall goal is to identify schizophrenia genes (PI H. Hakonarson, CHOP).....Others: Stephen Friend; Eric Schadt, Pamela Sklar; Raquel Gur……..
11. **Deep Sequencing of 296 Genes Implicated in Inflammatory Bowel Disease by Genome Wide Association Studies in Pediatric Cohort**: Overall goal is to identify variants associated with IBD (PI, H. Hakonarson); ……..others: Robert Baldassano; Mark Daly MGH/Broad…
12. **NGS:** Whole genome scan to test for association to autism utilizing high-throughput tag-SNP arrays. The project used genotyping and sequencing data generated by CAG and CAG-collaborators for association studies of both common and rare variants in autism. (PI H. Hakonarson)
13. **Genetic Modifiers of Liver Disease Severity in Alagille Syndrome:** Overall goal is to resolve AGS (Co-I, H. Hakonarson, CHOP)
14. **Fine Mapping and Functional Evaluation of Selected Type 1 Diabetes Loci:** Overall goal is to identify genetic variants associated with T1D (PI. H. Hakonarson)
15. **The Genetic Basis of Conotruncal Defects:** Overall goal is to uncover CHD genes. (Co-I, H. Hakonarson, CHOP)
16. **NGS:** Type 1 diabetes, targeted sequencing. (PI, H. Hakonarson)
17. **NGS:** Crohn's disease, targeted sequencing. (PI, H. Hakonarson, CHOP)
18. **NGS:** Exome sequencing, early forms of schizophrenia. (PI, R. Gur, UPenn; PI-CHOP, H. Hakonarson)
19. **GWAS:** Bone Mineral Accretion during Childhood (Co-I, H. Hakonarson, CHOP)
20. **GWAS:** Childhood obesity. (PI, S. Grant, CHOP)
21. **GWAS**: Latent autoimmune diabetes in adults (PI, S. Grant, CHOP)
22. **GWAS:** Gene-Environment Interactions in Asthma. (Co-PI, H. Hakonarson, CHOP)
23. **GWAS:** Neuroblastoma. (PI, J. Maris, CHOP)
24. **GWAS:** High density lipoprotein cholesterol. (PI, D. Rader, Upenn; PI-CHOP, H. Hakonarson)
25. **GWAS:** Schizophrenia. (PI, H. Hakonarson, CHOP)
26. **PGx:** Ventilation in pediatric patients with respiratory failure. (PI: A. Zuppa, CHOP)

Geisinger

1. **HMO Research Network:** member site; collaborating with investigators at Group Health on the pharmacogenomics of extreme weight gain due to 2nd/3rd generation anti-psychotic use funded by the International Severe Adverse Events Consortium (iSAEC).
    a. **HMORN Genomics Special Interest Group:** This group focuses on use of the HMORN for genomic medicine research. It was recently reactivated, and Geisinger personnel are leading the group.
2. **Pharmacogenomics Research Network (PGRN):** associate member site; clopidogrel genetics collaborations with investigators from the University of Maryland funded by the NIH as part of the Pharmacogenomics of Antiplatelet Intervention 2 (PAPI-2) Study, a multi-center prospective double-blind randomized comparative effective trial comparing personalized anti-platelet therapy to standard of care in patients undergoing percutaneous coronary interventions.
3. **Clinical Decision Support Consortium:** The goal of the CDSC is to assess, define, demonstrate, and evaluate best practices for knowledge management and clinical decision support in healthcare information technology (IT) at scale – across multiple ambulatory care settings and EHR technology platforms.
4. **Clinical Pharmacogenomics Implementation Consortium:** The CPIC develops evidence-based guidelines for the use of pharmacogenomics data to inform medication use. Geisinger actively participates in the review of all guidelines and co-authored the recently published guideline regarding the use of *IL28B* genotyping in treatement of Chronic Hepatitis C.
    a. **CPIC Informatics Working Group:** This group is working to develop generalizable informatic solutions to lower the implementation barriers for CPIC guidelines. Geisinger is actively participating in the work of this group.
5. **American Medical Informatics Association Genomic Working Group:** This group's purpose is to focus on opportunities in biomedical informatics, that arise from the storage, retrieval, analysis, and dissemination of molecular information in a clinical setting. Geisinger joined this group in Spring 2014.
6. **AAA Meta-GWAS Consortium:** Collaborating with WTCCC, University of Utrecht, deCODE Genetics, and New Zealand AAA Study.
7. **Aneurysm Global Epidemiology Study (AGES):** Collaborating with Edward Choke from University of Leicester, UK.
8. **Mid-Atlantic Nutrition Obesity Research Center (NORC):** The Mid-Atlantic NORC brings together resources at the University of Maryland, Johns Hopkins University, The US Department of Agriculture and Geisinger Health System; and cross-cutting expertise in various fields that can address the basic mechanisms that determine individual responses to nutrient intake and energy imbalance.
9. **PA CURE:** A grant awarded by the PA-DOH and lead by Geisinger in collaboration with The University of Pittsburgh and Temple University to create a risk scoring tool and test its utility for population screening of AAA.
10. **PCORnet project EMPOWERING PATIENTS AND FAMILIES FOR COMMUNITY-DRIVEN RESEARCH**: **THE DUCHENNECONNECT PATIENT-REPORT REGISTRY INFRASTRUCTURE PROJECT**: The project will explore improvements to DuchenneConnect to increase participant engagement, increase collection and transfer of patient reported data, and enhance the accuracy of patient-reported data. The project will explore how to move information from MyGeisinger (EPIC) into DuchenneConnect. Geisinger will explore IRB issues related to the sharing of data and will serve as the IRB of record. Research Group: Parent Project Muscular Dystrophy (PPMD), UCLA, PatientCrossroads, Geisinger
11. **GIANT** (Genetic Investigation of Anthropometric Traits) – ExomeChip data contributed for discovery analysis for anthropometric traits from approximately 7,600 MyCode participants and workgroup participation.
12. **CHARGE (Cohorts for Heart and Aging Research In Genomic Epidemiology):** Exome chip data contributed for analysis of kidney disease from approximately 7,600 MyCode participants. Gerard Tromp from Geisinger is participating in all data analysis calls weekly. Group leader is Caroline Fox from NIH.
13. **AAA ExomeChip Consortium:** Exome chip data contributed for analysis of abdominal aortic aneurysms from approximately 7,600 MyCode participants. Gerard Tromp from Geisinger is participating in the data analysis. Group leader is Matthew Bown from University of Leicester, UK.
14. **Regeneron Genetics Center Collaboration:** The goal is to perform whole exome sequencing of up to 100,000 Geisinger MyCode participants and to identify novel gene-phenotype associations by linking to electronic medical record data.

Group Health/University of Washington

1. Alzheimer's Disease Genetics Consortium (**ADGC**)
2. (GENE enVironment Association) (**GENEVA):** to test and validate an automated algorithmic method to identify mosaic regions

3. Northwest Institute of Genetic Medicine (**NWIGM**): to use the ongoing NWIGM biorepository to expand subject pool for eMERGE Phase 2. There are approximately 1700 subjects in NWIGM that have not been GWAS'd for eMERGE to date and which remain as an expansion cohort. All 2024 subjects in NWIGM will have exome chips available to eMERGE.
4. HMO Research Network: member site; GHRI recently completed a CommonFund Collaboratory project to help develop resources and capacity to support large-scale epidemiological studies(megaepidemiology) with potential to link to HMO based biobanks to address critical questions about the burden, causes, course, treatment, and prevention of common and rare human diseases.
5. **Scalable and Robust Clinical Text De-Identification Tools:** New R01 funded by National Library of Medicine that is a collaboration between Group Health and Vanderbilt to develop novel text de-identification methods. This work may facilitate improved access within an institution to its clinical text (e.g. for phenotype development purposes) as well as sharing of clinical text across institutions.
6. **NWIGM:** to study markers of Erectile Dysfunction (*potential collaboration*) in collaboration with Dr. Hunter Wessels, University of Washington Medical Center. Several eMERGE sites have expressed interest in participating in this collaboration.
7. **AD Sequencing Project (ADSP):** is doing whole exome sequencing (WES) on n=1,342 ACT participants, with data due June 2014. We as ACT and as eMERGE will benefit from having WES data. The phenotypes considered by ADSP are AD case/control phenotype (which we have from ACT) and the WES sequencing.
8. **CARE**: WBC (*potential collaboration*)
9. We are exploring a possible collaboration with **Kaiser Permanente Hawaii** to link EHR data from enrolled members with biobank samples. At this time, only prep research work has begun in coordination with KPHI.

Marshfield/Essentia/Penn State

1. **AMD** consortium Retinal specialists from Vanderbilt confirmed case status. Data will be shared in the near future with the network which is attempting to identify novel genetic variants associated with AMD.
2. **Glaucoma** consortium: additional samples were GWAS'd. Several papers have been published and other are under review. (1R01EY022305-01)
3. **PhenX** received admin supplement for year 1 of eMERGE II. Seven sites were funded in the PhenX RISING network. The network modeled their data sharing agreement on the eMERGE DUA. Supplement funding is complete. PhenX looked at Ecologic Stressors, Post-Traumatic Stress Disorder, and Drug Use in Detroit, University of Michigan, Ann Arbor, Allison Aiello, PhD and National Institute on Drug Abuse. PhenX has 1 paper published and 3 more papers under review.
    a. Cathy McCarty co-chairs with the Steering Committee for the evaluation and revisions of PhenX Toolkit Measures.
4. **NDPBRN National Dental Practice Based Research Network** to support dentists in individual practices or CHC to conduct research.
5. **"Oral-Systemic Health Consortium"** with Mt. Sinai, Univ. of Pittsburgh, UMDNJ, and UNC. The Oral-Systemic Health Consortium is in the process of enrolling study subjects along with microbiome samples
6. **HMORN** member site for Marshfield and Essentia; The Collaboratory will seek to enable large-scale epidemiological studies to address critical questions about the burden, causes, course, treatment, and prevention of common and rare human diseases. Development/enhancement of web-based patient portals and creation of megacohorts seem to be achievable goals of an eMERGE/HMO RN Collaboratory partnership.
7. **International Health Terminology Standards Development Organization** (IHTSDO): standards for safe, precise and effective exchange of clinical and health information.
8. **Clinical and Translational Science Awards** (CTSA): The goal of the University of Wisconsin Institute for Clinical and Translational Research (UW ICTR) is to create an environment that transforms research into a continuum from investigation through discovery and to translation into real-life community practice, thereby linking even the most basic research to practical improvements in human health.
9. **Wisconsin Genomics Initiative** (WGI): The Wisconsin Genomics Initiative advances personalized health care research. Scientists will be able to predict an individual's risk of developing a disease, precisely target a personalized treatment and ultimately prevent disease before it occurs.
10. **PGRN-PGPOP** - PharmacoGenomic discovery and replication in very large patient POPulations. PGPop was conceived as a network resource to provide to PGRN an opportunity to identify large groups of real world patients with known drug exposures and outcomes for pharmacogenomic study in a clinical setting.
11. **Iron Metabolism** (Christopher Vulpe, PI) "An integrated molecular approach to understand variation in iron metabolism" (1R01GM083198-01A1).

12. **A Traumatic Brain Injury (TBI) Patient Registry to Examine Variations in Diagnosis, Treatment and Outcomes** (Allen Heinemann and Abel Kho) As one specific Aim we propose to: Develop and validate a traumatic brain injury phenotype, linked to existing genetic data, using EHR data within the eMERGE consortium. In addition to defining the phenotype, we would also try and document certain objective measures of TBI measures and outcomes.
13. **GEMS Consortium** A number of pharmacogenomics research centers (the "consortium") are interested in discovering new genetic variants that are important for predicting myopathy in the context of statin therapy using candidate gene association studies (CGAS), genome wide association studies (GWAS), exome scanning, whole genome sequencing, and other methods. Many members of the consortium have collected cohorts of patients who have been treated with lipid modifying drugs and have obtained a variety of response phenotypes, including both efficacy and toxicity endpoints, and have genotype data and/or DNA available from these participants. Members of the consortium are willing to share these data for the purposes of replicating the findings of others, and for combined (meta) analysis. Based on the success of similar consortia, other disease genetics consortia, and other pharmacogenetic consortia GEMS seeks to utilize this same approach to advance pharmacogenomics for statin therapy.
14. **PCROI** (Univ Kansas Cooperative subcontract) Clinical Data Research Network (CDRN) Marshfield Clinic is one of 10 sites from the Greater Plains Collaborative as a Clinical Data Research Network within the Patient Centered Outcome Research Institute's new National Patient-Centered Clinical Research Network.

Mayo Clinic

1. **AT&T** Foundation
2. Multi-source Integrated Platform for Answering Clinical Questions (**MiPAQ**)
3. The Environmental Determinants of Diabetes in the Young Study (**TEDDY**)
4. **World Health Organization (WHO):** International Classification of Disease (ICD)-11 Revision
5. cancer Biomedical Informatics Grid (**caBIG®**)
6. Consensus Measures for Phenotypes and Exposures (**PhenX**): Network collaboration on eleMAP and other data standards projects
7. Pharmacogenomics Research Network (PGRN): member site.
8. Pharmacogenomics Ontology Network Resource (PHONT within PGRN): PGRN network resource, an effort to support meta-analyses, achieve translational goals, and facilitate the messaging of pharmacogenomics-related data from and into clinical environments such as EMRs
9. National Center for Biomedical Ontology (NCBO): collaborating site, an effort to leverage ontologies to increase access to and understanding of defined terms and relationships in the biomedical domain.
10. Office of National Coordinator for Health Information Technology (ONC).
11. PheMA (Phenotype Modeling Architecture): Vanderbilt and Northwestern

12. **Peripheral arterial disease (PAD) Genetics Consortium:** Collaborating with University of Dundee Scotland, Wellcome Trust Centre for Human Genetics, deCODE Genetics, and others.
13. **Cardiogram+4CD consortium**. Coronary heart disease genetics consortium.
14. **AAA Meta-GWAS Consortium:** Collaborating with Geisenger Medical Center, WTCCC, University of Utrecht, deCODE Genetics, and New Zealand AAA Study.
15. **Fibromuscular dysplasia (FMD) genetics consortium.**
16. **GEMS Consortium:** Discovery of new genetic variants related to statin myopathy.
17. **Amiodarone pulmonary toxicity:** (with Vanderbit University)
18. **GWAS of vancomycin induced changes in serum creatinine** (with Vanderbilt University)
19. **Susceptibility Genes for Erectile Dysfunction** (with Dr. Hunter Wessels, UW and additional eMERGE sites)

Icahn School of Medicine at Mount Sinai

1. **GIANT** (Genetic Investigation of Anthropometric Traits) – GWAS data contributed for discovery analysis for anthropometric traits from all BioMe participants and workgroup participation
2. **GIANT** (Genetic Investigation of Anthropometric Traits) – ExomeChip data contributed for discovery analysis for anthropometric traits from all BioMe participants and workgroup participation
3. **COGENT BP** (Continental Origins and Genetic Epidemiology Network) – GWAS data for BP from African American BioMe participants contributed for discovery analysis
4. **COGENT BP** (Continental Origins and Genetic Epidemiology Network) – data for BP from African American BioMe participants contributed for discovery analysis

5. **GHBP** (Genomics in Hispanics for Blood Pressure) – GWAS data for BP from Hispanics contributed for discovery analysis
6. **Massachusetts Institute of Technology** Computer Science and Artificial Intelligence Laboratory (John Guttag): Predictive Modelling and Personalized Health Decision Support Tools
7. **African American Anthropometrics Genetics Consortium** – GWAS data of BMI from African Americans BioMe participants contributed for discovery and follow up analysis and workgroup participation
8. **African American Type 2 Diabetes Genetics Consortium** – GWAS data of T2D from African American BioMe participants contributed for analysis
9. **CHARGE (Cohorts for Heart and Aging Research In Genomic Epidemiology)** Exome chip data contributed for analysis of BP from all BioMe participants
10. **CHARGE (Cohorts for Heart and Aging Research In Genomic Epidemiology)** Exome chip data contributed for analysis of blood traits from all BioMe participants
11. **CHARGE (Cohorts for Heart and Aging Research In Genomic Epidemiology) – Gene-Lifestyle interaction working group:** GWAS data contributed for analysis of lipids that examines interaction with smoking from all BioMe participants
12. **CKDGen (CKD Genetics Consortium)** – GWAS data contributed for discovery analysis and workgroup participation
13. **CKDGen (CKD Genetics Consortium)** Exome chip data contributed for analysis of BP from all BioMe participants
14. **GLGC** (Global Lipids Genetics Consortium) Exome chip data contributed for discovery analysis of all lipids from all BioMe participants
15. **GLGC** (Global Lipids Genetics Consortium) Exome chip data contributed for follow-up analysis of CAD from all African American BioMe participants
16. **ESP-LDL (Exome Sequencing Projects LDL Cholesterol)** Exome chip data contributed for follow-up analysis of LDL from all BioMe participants
17. **MAGIC (Meta-Analyses of Glucose and Insulin-related traits Consortium) Exome chip data** contributed for discovery analysis of HbA1c from all BioMe participants
18. **CHARGE (Cohorts for Heart and Aging Research In Genomic Epidemiology)** Exome chip data contributed for discovery analysis of Glycaemic traits from all BioMe participants
19. **CHARGE (Cohorts for Heart and Aging Research In Genomic Epidemiology)** Exome chip data contributed for follow-up analysis of Amyloidoses from all BioMe participants
20. **MEDIA MEta-analysis of type 2 DIabetes in African Americans (MEDIA) Consortium**
GWAS data contributed for discovery analysis of T2D from African American BioMe participants.
21. TranscenD (TRANS-ethnic Evaluation of vitamin D) GWAS data contributed for discovery analysis of Vitamin D from all BioMe participants.
22. **Lipids in HA** GWAS data contributed for discovery analysis of Lipids from all Hispanic American BioMe participants.
23. **AAGILE** (African American Glucose and Insulin Genetic Epidemiology (AAGILE) Consortium) GWAS data contributed for discovery analysis of HbA1c and glucose from all African American BioMe participants.
24. **DIAGRAM+ and GOT2D (Genetics of Type 2 Diabetes)** GWAS data contributed for follow-up analysis of T2D from all BioMe participants.
25. **BP in HA** GWAS data contributed for discovery analysis of BP from all Hispanic American BioMe participants.
26. **ICBP** (International Consortium for Blood Pressure) GWAS data contributed for discovery analysis of BP from all European American BioMe participants.
27. **HISLA study - Anthropometric Traits in HA** GWAS data contributed for discovery analysis of Anthropometric Traits from all Hispanic American BioMe participants.
28. **T2D Genes** Targeted sequencing for follow-up analysis of T2D from all European American BioMe participants.
29. **IGNITE Network (NHGRI U01)** Implementing Genomics in Practice
30. **Study of Latinos (SOL): Participation in lipids and antheropomtrics working group, replication with Bio*Me* HA.**
31. **TranscenD: GWAS data contributed for disocevry analyses for Vitamin D in all Bio*Me* participants**
32. **CARDIOGRAMPLUS4CD: GWAS data contributed for discovery and fine-mapping of genetic loci associated with CAD.**
33. **Runs of Homozygosity Consortium: Contributed GWAS data for a number of traits across all BioMe participants to test for runs of hozygosity.**

Northwestern University

1. **African American Type 2 Diabetes Genetics Consortium (MEDIA):** Don Bowden (Wake Forest) leads a collaborative group of ~12 sites including Northwestern and Vanderbilt. This consortium conducted a meta-analysis of genomic determinants of Type 2 Diabetes in African Americans. Paper is submitted to Nature Genetics, with multiple eMERGE investigators and the eMERGE consortium represented as authors.

2. **COGENT** (Continental Origins of Genetic Traits): New meta-analysis consortium focused on African-American and Latino populations; NU participates in trait-specific working group (for height & RBC phenotypes)
3. **GIANT** Consortium: Meta-analysis of height from >100,000 subjects across multiple studies. Awaiting draft of manuscript from consortium lead analysts.
4. Vanderbilt: Vitamin D Related Innate Immunity in Influenza
5. **PAGE** (Population Architecture using Genomics and Epidemiology) : inc. a PheWAS replication in eMERGE; including blood count data for AA pts
6. **CTSA project with Mayo** (neutropenia and thrombocytopenia)
7. **CAGE** (QRS in African Americans with VU)
8. Stanford University: **Extracting the Quality of Prostate Cancer Care from EHR**
9. **UCSF: Metformin and T2D Project**
10. **AAA Genotyping**– collaborating with Geisinger
11. **De-Id Project (Anonymization of clinical codes in support of genome-phenome association studies)** with Marshfield and Vanderbilt
12. **AAGILE/MEDIA: Fasting Glucose meta-analysis**
13. **Pharmacogenomics Research Network (PGRN):** associate member site
14. Imperial College London: **Genome-wide association meta-analysis of random glucose levels**
15. PheMA (Phenotype Modeling Architecture): Vanderbilt and Mayo

Vanderbilt University

1. **QRS GWAS Consortium**: across eMERGE; replication planned with CHARGE
2. QT Interval GWAS Consortium **(QT-IGC)**
3. **PRIMA** (CHARGE-led mega meta-analysis of PR interval)
4. Pharmacogenomics Research Network **(PGRN):** member site
5. **Kaiser Permanente, Marshfield Clinic, RIKEN** – statins & MI
6. **UCSF and RIKEN** – Metformin-related glycemic response
7. **Marshfield Clinic, Harvard Crimson, and Harvard Pilgrim** – Asthma response to inhaled steroids; methotrexate-induced liver injury
8. **RIKEN** – ACE inhibitor-associated angioedema Published, Asthma response to inhaled steroids
9. **Baylor, Marshfield Clinic, Mayo Clinic** – Amio-induced pulmonary toxicity and thyrotoxicosis.
10. **St Jude Children's Research Hospital** – Steroid-induced osteonecrosis
11. **Children's Hospital Oakland Research Institute (CHORI) and Marshfield Clinic** – Statin effects in asthma Published
12. **University of Florida** – Cerebrovascular disease and clopidogrel
13. **University of Colorado Boulder** – Seborrheic keratosis
14. **University of Maryland** – Association of CES1 G143E with bleeding and decreased events during clopidogrel therapy
15. **ICPC (International Clopidogrel Pharmacogenomics Consortium)**
16. **CPIC (Clinical Pharmacogenetics Implementation Consortium)**
17. Consensus Measures for Phenotypes and Exposures **(PhenX):** Network collaboration on eleMAP and other data standards projects
18. GANI_MED biorepository in Greifswald, Germany: biobank ELSI collaboration
19. Children's Hospital of Eastern Ontario (CHEO) / University of Ottawa to study re-identification risks in electronic medical records data tied to genomic records.
20. Sabanci University and Zirve University in Turkey to develop anonymization strategies for longitudinal records, which we evaluated with the Vanderbilt's QRS GWAS cohort.
21. **University of Illinois at Urbana Champaign** (Carl Gunter) and **Ecole Polytechnique Federale du Laussanne (EPFL)** around systematizing knowledge regarding identifiability attaches on research genomic data for the computer privacy and security community.
22. **University of South Australia** (Dr. Jiyong Li and Dr. Xiaofeng Ding) for de-identification policy search research
23. **GEMS (Genetics and Myopathy on Statins Consortium)**
24. **The Ohio State University** - Interaction model using two genetic variants in dopamine beta hydroxylase (DBH) to predict protection from myocardial infarction
25. **University of California at San Francisco** - Genomewide Meta-Analysis of Allopurinol Response in Patients with Gout or Hyperuricemia
26. PCORNet: use of PheKB phenotyping assistance
27. Qatar Biobank - EMR design

eMERGE Phase II Publications from June 2011 – October 2015

Digital Reference Library Available Here

Published/Accepted and Submitted Phase II Network Manuscripts

1. Nadkarni G, Gottesman O, Farouk S, Weng C, Peissig P, et al. Development and validation of an electronic phenotyping algorithm for chronic kidney disease. JAMIA. (Accepted)
2. Jeff JM, Brown-Gentry K, Goodloe R, Ritchie MD, Denny JC et al. Replication of SCN5A Associations with Electrocardiographic Traits in African Americans from Clinical and Epidemiologic Studies. Lecture Notes in Comput. Sci. 2014. (Accepted).
3. Overby CL, Rasmussen LV, Hartzler A, Connolly JJ, Peterson JF, et al. A Template for Authoring and Adapting Genomic Medicine Content in the eMERGE Infobutton Project. JAMIA (In Press)
4. Parihar A, Wood GC, Chu X, Jin Q, Argyropoulos G, et al. Extension of GWAS results for lipid-related phenotypes to extreme obesity using electronic health record (EHR) data and the Metabochip. Front Genet. 2014 Aug 5;5:222. **PMID: 25147553 PMCID: PMC4123014**
5. Arking DE, Pulit SL, Crotti L, van der Harst P, Munroe PB, et al. Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. Nat. Genet. 2014 Aug; 46(8):826-36, PMID: 24952745 **PMCID: PMC4124521**
6. Crosslin DR, Tromp G, Burt A, Kim DS, Verma SS, et al. Controlling for population structure and genotyping platform bias in the eMERGE multi-institutional biobank linked to electronic health records. Front Genet. 2014 Nov 4;5:352. **PMID: 25414722 PMCID: PMC4220165**
7. Namjou B, Marsolo K, Caroll R, Denny JC, Ritchie MD, et al. Phenome-wide association study (PheWAS) in EMER-linked pediatric cohorts, genetically links PLCL1 to speech language development and IL5-IL13 to Eosinophilic Esophagitis. Front. Genet., 18 November 2014, **PMID: 25477900 PMCID: PMC4235428**
8. Verma SS, de Andrade M, Tromp G, Kuivaniemi H, Pugh E, et al. Imputation and quality control steps for combining multiple genome-wide datasets. Front Genet. 2014 Dec 11;5:370, **PMID: 25566314 PMCID: PMC4263197**
9. Denny JC, Beilinski SJ, Basford MA, Bradford Y, Peissig PL, et al. Genetic variants associated with serum thyroid stimulating hormone (TSH) levels in European Americans and African Americans from the eMERGE Network. PLoS One. 2014 Dec 1;9(12):e111301, **PMID: 25436638 PMCID: PMC4249871**
10. Crosslin DR, Carrell DS, Burt A, Kim DS, Underwood JG, et al. Genetic variation in the HLA region is associated with susceptibility to herpes zoster. Genes Immun. 2014 Oct 9;0. **PMID: 25297839**
11. Verma SS, de Andrade M, Tromp G, Kuivaniemi H, Pugh E, et al. Imputation and quality control steps for combining multiple genome-wide datasets. Front Genet. 2014 Oct; 5:370.
12. Brothers KB, Lynch JA, Aufox SA, Connolly JJ, Gelb BD, et al. Practical Guidance on Informed Consent for Pediatric Participants in a Biorepository. Mayo Clin Proc. 2014 Sep 25; **PMID: 25264176**
13. Hayes G, Ng MCY, Shriner D, Chen BH, Li J, et al. Meta-analysis of type 2 Diabetes in African Americans Consortium. Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. PLoS Genet. 2014 Aug;10(8):e1004517. **PMID: 25102180 PMCID: PMC4125087**
14. Rasmussen-Torvik LJ, Stallings SC, Gordon AS, Almoguera B, Basford MA, et al. Design and Anticipated Outcomes of the eMERGE-PGx Project: A Multi-Center Pilot for Pre-Emptive Pharmacogenomics in Electronic Health Record Systems. Clin Pharmacol Ther. 2014 Jun 24; **PMID: 24960519**
15. Rasmussen LV, Thompson WK, Pacheco JA, Kho AN, Carrell DS, et al. Design patterns for the development of electronic health record-driven phenotype extraction algorithms. J Biomed Inform. 2014 Jun 21; **PMID: 24960203**
16. Almoguera B, Vazquez L, Connolly JJ, Bradfield J, Sleiman P, et al. Imputation of TPMT defective alleles for the identification of patients with high-risk phenotypes. Front. Genet. 2014 May 12;5:96. **PMID: 24860591 PMCID: PMC4026736**
17. Ye Z, Vasco DA, Carter T, Brilliant M, Schrodi SJ, et al. Genome wide association study of SNP-, gene-, and pathway-based approaches to identify genes influencing susceptibility to Staphylococcus aureus infections. Staphylococcus aureus. 2014 May 9;5:125. **PMID: 24847357 PMCID: PMC4023021**
18. Jarvik GP, Amendola LM, Berg JS, Brothers K, Clayton EW, et al. Return of Genomic Results to Research Participants: The Floor, the Ceiling, and the Choices In Between. Am J Hum Genet. 2014 May 7; **PMID: 24814192**
19. Sleiman P, Bradfiled J, Mentch F, Almoguera B, Connolly J, et al. Assessing the functional consequence of loss of function variants using electronic medical record and large-scale genomics consortium efforts. Front Genet. 2014 Apr 29;5:105. **PMID: 24808909 PMCID: PMC4010747**
20. Mitchell BD, Fornage M, McArdle PF, Cheng Y-C, Pulit S, et al. Using previously genotyped controls in genome-wide association studies (GWAS): application to the Stroke Genetics Network (SiGN). Front Genet. 2014 Apr 29;5:95. **PMID: 24808905 PMCID: PMC4010766**
21. Sun X, Lu Q, Mukheerjee S, Crane P, Elston RC, et al. Analysis pipeline for the epistasis search – statistical versus biological

filtering. Front Genet. 2014 Apr 30;5:106. **PMID: 24817878 PMCID: PMC4012196**
22. Kullo IJ, Haddad R, Prows CAM, Holm I, Sanderson SC, et al. Return of Genomic Results in the Genomic Medicine Projects of the eMERGE Network. Front Genet. 2014 Mar 26;5:50. **PMID: 24723935 PMCID: PMC3972474**
23. Connolly JJ, Glessner JT, Almoguera B, Crosslin DR, Jarvik GP, et al. Copy number variation analysis in the context of electronic medical records and large-scale genomics consortium efforts. Front Genet. 2014 Mar 18;5:51. **PMID: 24672537 PMCID: PMC3957100**
24. Jeff JM, Armstrong LL, Ritchie MD, Denny JC, Kho AN, et al. Admixture mapping and subsequent fine-mapping suggests a biologically relevant and novel association on chromosome 11 for type 2 diabetes in african americans. PLoS ONE. 2014 Mar 3;9(3):e86931. **PMID: 24595071 PMCID: PMC3940426**
25. Patel ZH, Kottyan LC, Lazaro S, Williams MS, Ledbetter DH, et al. The struggle to find reliable results in exome sequencing data: filtering out Mendelian errors. Front Genet. 2014 Feb 12;5:16. **PMID: 24575121 PMCID: PMC3921572**
26. Muthalagu A, Pacheco JA, Aufox S, Peissig PL, Fuehrer JT, et al. A Rigorous Algorithm To Detect And Clean Inaccurate Adult Height Records Within EHR Systems. Applied Clinical Informatics. 2014 Feb 19;5(1):118–126. **PMID: 24734128 PMCID: PMC3974252**
27. Crawford DC, Crosslin DR, Tromp G, Kullo IJ, Kuivaniemi H, et al. eMERGEing progress in genomics—the first seven years. Front Genet. 2014;5:184. **PMID: 24987407 PMCID: PMC4060012**
28. Cronin RM, Field JR, Bradford Y, Shaffer CM, Carroll RJ, et al. Phenome-wide association studies demonstrating pleiotropy of genetic variants within FTO with and without adjustment for body mass index. FTO. 2014;5:250.
29. Kullo IJ, Shameer K, Jouni H, Lesnick TG, Pathak J, et al. The ATXN2-SH2B3 locus is associated with peripheral arterial disease: an electronic medical record-based genome-wide association study. Front Genet. 2014;5:166. **PMID: 25009551 PMCID: PMC4070196**
30. Schrodi SJ, Mukherjee S, Shan Y, Tromp G, Sninsky JJ, et al. Genetic-based prediction of disease traits: prediction is very difficult, especially about the future. Front Genet. 2014;5:162. **PMID: 24917882 PMCID: PMC4040440**
31. Pathak J, Kho AN, Denny JC. Electronic health records-driven phenotyping: challenges, recent advances, and perspectives. J Am Med Inform Assoc. 2013 Dec;20(e2):e206–211. **PMID: 24302669 PMCID: PMC3861925**
32. Overby CL, Pathak J, Gottesman O, Haerian K, Perotte A, et al. A collaborative approach to developing an electronic health record phenotyping algorithm for drug-induced liver injury. J Am Med Inform Assoc. 2013 Dec;20(e2):e243–252. **PMID: 23837993 PMCID: PMC3861914**
33. Mosley JD, Van Driest SL, Larkin EK, Weeke PE, Witte JS, et al. Mechanistic Phenotypes: An Aggregative Phenotyping Strategy to Identify Disease Mechanisms Using GWAS Data. PLoS ONE. 2013 Dec 12;8(12):e81503. **PMID: 24349080 PMCID: PMC3861317**
34. Namjou B, Keddache M, Marsolo K, Wagner M, Lingren T, et al. EMR-linked GWAS study: investigation of variation landscape of loci for body mass index in children. Front Genet. 2013 Dec 3;4:268. **PMID: 24348519 PMCID: PMC3847941**
35. Overby CL, Pathak J, Gottesman O, Haerian K, Perotte A, et al. A collaborative approach to developing an electronic health record phenotyping algorithm for drug-induced liver injury. J Am Med Inform Assoc. 2013 Dec;20(e2):e243-52. doi: 10.1136/amiajnl-2013-001930. Epub 2013 Jul 9. **PMID: 23837993 PMCID: PMC3861914**
36. Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. Nat Biotech. 2013 Nov 24; **PMID: 24270849 PMCID: PMC3969265**
37. Overby CL, Kohane I, Kannry JL, Williams MS, Starren J, et al. Opportunities for genomic clinical decision support interventions. Genet Med. 2013 Oct;15(10):817–823. **PMID: 24051479 PMCID: PMC3858176**
38. Tarczy-Hornoch P, Amendola L, Aronson SJ, Garraway L, Gray S, et al. A survey of informatics approaches to whole-exome and whole-genome clinical reporting in the electronic health record. Genet Med. 2013 Oct;15(10):824–832. **PMID: 24071794 PMCID: PMC3951437**
39. Chute CG, Ullman-Cullere M, Wood GM, Lin SM, He M, et al. Some experiences and opportunities for big data in translational research. Genet Med. 2013 Oct;15(10):802–809. **PMID: 24008998 PMCID: PMC3906918**
40. Peterson JF, Bowton E, Field JR, Beller M, Mitchell J, et al. Electronic health record design and implementation for pharmacogenomics: a local perspective. Genet Med. 2013 Oct;15(10):833–841. **PMID: 24009000 PMCID: PMC3925979oc**
41. Kho AN, Rasmussen LV, Connolly JJ, Peissig PL, Starren J, et al. Practical challenges in integrating genomic data into the electronic health record. Genet Med. 2013 Oct;15(10):772–778. **PMID: 24071798**
42. Kannry JM, Williams MS. Integration of genomics into the electronic health record: mapping terra incognita. Genet Med. 2013 Oct;15(10):757–760. **PMID: 24097178**
43. Kannry J, Williams MS. The undiscovered country: the future of integrating genomic information into the EHR. Genet Med. 2013 Oct;15(10):842–845. **PMID: 24071799**
44. Hazin R, Brothers KB, Malin BA, Koenig BA, Sanderson SC, et al. Ethical, legal, and social implications of incorporating genomic information into electronic health records. Genet Med. 2013 Oct;15(10):810–816. **PMID: 24030434 PMCID:**

PMC3926430
45. Hartzler A, McCarty CA, Rasmussen LV, Williams MS, Brilliant M, et al. Stakeholder engagement: a key component of integrating genomic information into electronic health records. Genet Med. 2013 Sep 12; **PMID: 24030437 PMCID: PMC3909653**
46. McDavid A, Crane PK, Newton KM, Crosslin DR, McCormick W, et al. Enhancing the Power of Genetic Association Studies through the Use of Silver Standard Cases Derived from Electronic Medical Records. PLoS ONE. 2013 Jun 10;8(6):e63481. **PMID: 23762230 PMCID: PMC3677889**
47. Gottesman O, Kuivaniemi H, Tromp G, Faucett WA, Li R, et al. The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. Genet Med. 2013 Jun 6; **PMID: 23743551 PMCID: PMC3795928**
48. Chute CG, Kohane IS. Genomic medicine, health information technology, and patient care. JAMA. 2013 Apr 10;309(14):1467–1468. **PMID: 23571583 PMCID: PMC3959893**
49. Wei W-Q, Leibson CL, Ransom JE, Kho AN, Chute CG. The absence of longitudinal data limits the accuracy of high-throughput clinical phenotyping for identifying type 2 diabetes mellitus subjects. Int J Med Inform. 2013 Apr;82(4):239–247. **PMID: 22762862 PMCID: PMC3478423**
50. Jeff JM, Ritchie MD, Denny JC, Kho AN, Ramirez AH, et al. Generalization of Variants Identified by Genome-Wide Association Studies for Electrocardiographic Traits in African Americans. Ann Hum Genet. 2013 Mar 28; **PMID: 23534349 PMCID: PMC3743946**
51. Starren J, Williams MS, Bottinger EP. Crossing the Omic Chasm: A Time for Omic Ancillary Systems. JAMA. 2013 Mar 14;1–2. **PMID: 23494000 PMCID: PMC3857698**
52. Crosslin DR, McDavid A, Weston N, Zheng X, Hart E, et al. Genetic variation associated with circulating monocyte count in the eMERGE Network. Hum Mol Genet. 2013 Jan 12; **PMID: 23314186 PMCID: PMC3633369**
53. Kullo IJ, Jarvik GP, Manolio TA, Williams MS, Roden DM. Leveraging the electronic health record to implement genomic medicine. Genet Med. 2012 Sep 27; **PMID: 23018749**

In Process Phase II Network Manuscripts

16. Cognitive Interviews associated with developing a national survey on consent across a national network of genomic medicine sites. Lead Investigator: Melanie Myers (CCHMC)
17. Developing a national survey on consent across a national network of genomic medicine sites. Lead Investigator: Ingrid Holm & Maureen Smith
18. A Literature Review of U.S. Individuals' Perspectives on Privacy, Trust, and Perceived Risks and Benefits in Biobanking. Lead Investigator: Nanibaa' Garrison (VU)
19. A Review of U.S. Individuals' Perspectives on Governance and Consent in Biobanking. Lead Investigator: Nanibaa' Garrison
20. A Highly Accurate Electronic Algorithm for the Classification of Asthma Severity in Children. Lead Investigator: Erik Hysinger (CHOP)
21. Epistatic gene-based interaction analyses for glaucoma in eMERGE network and NEIGHBOR consortium. Lead Investigator: Shefali Verma (PSU)
22. Exploring the genetic architecture of Age-Related Macular Degeneration (AMD) in the eMERGE network. Lead Investigator: Molly Hall (PSU)
23. Genome-wide Association Study of Gastroesophageal Reflux Disease (GERD) in Adult and Pediatric Populations. Lead Investigator: Patrick Sleiman (CHOP)
24. Genome-wide Association Study of Atopic Dermatitis in Adult and Pediatric Populations. Lead Investigator: Berta Almoguera (CHOP)
25. Genome-wide Association Study of Asthma in Adult and Pediatric Populations. Lead Investigator: Berta Almoguera (CHOP)
26. Genome-wide Association Study of Attention Deficit Hyperactivity Disorder (ADHD). Lead Investigator: John Connolly (CHOP)
27. Locus-specific PheWas study : An investigation of size effect of IRF5 and STAT4 in various autoimmune diseases and other phenotypes in all available samples in eMERGE. Lead Investigator: Bahram Namjou (CHOP)
28. Investigation of PCSK9 SNPs with LDL-C, BMI and risk of T2D. Lead Investigator: Brendan Keating (CHOP)
29. Phenotype transportability across Electronic Health Records. Lead Investigator: Joshua Denny (VU)
30. Practical considerations for implementing genomic information resources: experiences from eMERGE and CSER. Lead Investigator: Luke Rasmussen (NU)
31. GWAS study on appendicitis in pediatric and adult population: using participants of the eMERGE Network. Lead Investigator: Bahram Namjou (CHOP)
32. The COGENT consortium meta-analysis of blood pressure African ancestry cohorts. Lead Investigator: Dinga Velez Edwards (VU)
33. GWAS study on LFT (liver function test) in pediatric population: comparison of size effect between adult and children using participants of the eMERGE Network. Lead Investigator: Bahram Namjou (CHOP)

34. Genetic variation among 84 pharmacogenes from the PGRNSeq in the eMERGE Network. Lead Investigators: Will Bush (Case) & David Crosslin (GH/UW)
35. PheWAS analysis of a functional variant in CDHR3. Lead Investigator: Michael March (CHOP)
36. PheWAS analysis of homozygous deletions in GWAS data. Lead Investigator: Michael March (CHOP)
37. A Systematic Literature Review of Individuals' Perspectives on Broad Consent and Data Sharing in the United States. Lead Investigator: Nanibaa' Garrison (VU)
38. Desiderata for Computable Representations of Electronic Health Records-Driven Phenotype Algorithms. Lead Investigator: Huan Mo (NU)
39. Using PheWAS to assess disease comorbidity and potential pleiotropy of genetic risk scores for rheumatoid arthritis. Lead Investigator: Robert Carroll (VU)
40. 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)
41. Exploring genetics and outcomes associated with acute kidney injury (AKI) using electronic health records and genomics. Lead Investigator: Girish N Nadkarni (Mt. Sinai)
42. Practical Considerations in Genomic Decision Support: The eMERGE Experience. Lead Investigator: Tim Herr (NU)
43. A Conceptual Model for Omic Data. Lead Investigator: Tim Herr (NU)
44. Novel NR3C1 and NR3C2 variants to improve genetic screening/testing for endocrine hypertension. Lead Investigator: Noura Abul-Husn (Mt. Sinai)
45. PGRNseq and GWAS predictors of Methylphenidate (MPH) response. Lead Investigator: Tanya Froelich (CCHMC)
46. Null Variant PheWAS. Lead Investigators: Marylyn Ritchie (MC/EIRH/PSU & CC)
47. Discovery, Replication and Clinical Associations of Pathway-Based Trans-eQTL. Lead Investigator: Laura Wiley (VU)
48. Multiscale Analysis Of Influenza Host-Pathogen Interactions: Fluomics. Lead Investigator: Ellie Sang Sukerman (Northwestern)
49. Practical Considerations in Genomic Decision Support: The eMERGE Experience. Lead Investigator: Tim Herr (VU)
50. Rare RYR1, CACNA1S variant annotation, exposure history, observed phenotypes in cases and controls. Lead Investigator: Senthilkumar Sadhasivam (CCHMC)
51. Variant Calling and Annotation for 82 known pharmacogenes (tentative). Lead Investigator: Will Bush (CC)
52. ePhenotyping for Abdominal Aortic Aneurysm: Algorithm development and KNIME workflow. Lead Investigator: Ken Borthwick (Geisinger)
53. Characterizing the individual and shared genetic components of pheWAS phenotypes. Lead Investigator: Jonathan Mosley (VU)
54. Examining gene variants in eMERGE samples for association with uterine fibroids. Lead Investigator: Todd Edwards (VU)
55. A Phenome-wide Survey of the Phenotypic Effects of Neanderthal Admixture. Lead Investigator: Corrine Simonti (VU)
56. Prospective participant selection and ranking to maximize actionable PGx variants and discovery in the eMERGE Network. Lead Investigator: David Crosslin (GroupHealth)
57. MVtest: a method to flexibly model the genetic determinants of trait variability. Lead Investigator: Todd Edwards (VU)
58. Autism Spectrum Disorders: Electronic Health Record Mining and Comorbidity Clustering. Lead Investigator: Todd Lingren (CCHMC)
59. Pediatric Providers are Poor at Identifying Severe Obesity in Young Children at Two Tertiary Pediatric Medical Centers. Lead Investigator: Cassandra Brady (CCHMC)
60. Developing an Algorithm to Detect Early Childhood Obesity Two Tertiary Pediatric Medical Centers. Lead Investigator: Todd Lingren, Vidhu Thaker (BCH)
61. Association of APOL1 G1/G2 risk alleles with metabolic and cardiovascular traits. Lead Investigator: Girish Nadkarni (Mt. Sinai) & Miriam Udler (BCH)
62. Rare variant annotation and observed phenotypes in SCN5A and KCNH2. Lead Investigator: Sara Van Driest (VU)
63. GWAS in cohort of eMERGE network subjects with ACE-inhibitor induced cough. Lead Investigator: Jonathan Mosely (VU)
64. Development of a dynamic XML event-driven ophthalmologic data capture framework. Lead Investigator: Peggy Peissig (Marshfield)
65. GWAS Consortium for QT Interval (QT-IGC). Lead Investigator: Chris Newton-Cheh (BCH)
66. Burden of structural variation and PheWAS. Lead Investigator: David Crosslin (GroupHealth)
67. Optimal management of different types of genetic information in the Electronic Medical Record. Lead Investigator: Brian Shirts (UW)
68. Evaluation of phenotype representation models and authoring tools using eMERGE algorithms. Lead Investigator: Jen Pacheco (NU)
69. Knowledge driven search for gene-gene interactions associated with hypothyroidism in the eMERGE network. Lead Investigator: Molly Hall (MC/EIRH/PSU)
70. Gene-Gene interactions associated with BMI - Replication in the eMERGE network. Lead Investigator: Marylyn Ritchie (MC/EIRH/PSU & CC)

71. Discovery and replication of genetic interactions for quantitative lipid traits. Lead Investigator: Marylyn Ritchie (MC/EIRH/PSU & CC)
72. PhenomE-wide Association Study in eMERGE Pediatric Cohorts. Lead Investigator: Bahram Namjou (CCHMC)
73. Copy Number Variation Burden Analysis on a Range of Phenotypes in the eMERGE Network. Lead Investigator: Marylyn Ritchie (MC/EIRH/PSU & CC)
74. Establishing a Genomic Medicine Content Collection Process: Progress within the eMERGE Network. Lead Investigator: Casey Overby (External Collaborator) & Luke Rasmussen (NU)
75. Development and Validation of an Electronic Phenotyping Algorithm for Chronic Kidney Disease. Lead Investigator: Girish Nadkarni (Mt. Sinai)
76. Genome-wide Association Study of Serum Creatinine Levels during Vancomycin Therapy. Lead Investigator: Sara Van Driest (VU)
77. PheKB.org: An Online Collaboration Tool for Phenotype Algorithm Development and Sharing. Lead Investigator: Jacqueline Kirby
78. Genetic Risk Scores for Complex Diseases in the eMERGE Network: Characterization and Predicative Abilities in Clinical Settings. Lead Investigator: Logan Dumitrescu (VU)
79. 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)
80. Chromosomal Anomalies that Affect Levels of White Blood Count (WBC) and its Differential. Lead Investigator: David Crosslin (UW)
81. Extracting the Quality of Prostate Cancer Care from Electronic Healthcare Records. Lead Investigator:
82. Big Data Needs in Clinical Genomics: Use Cases Defined by Medical Providers. Lead Investigator: Bryan Weichelt (Marshfield)
83. Practical Approaches to the Omic Chasm. Lead Investigator: Justin Starren (NU)
84. Developing Consents for Returning Pharmacogenomics Results: The eMERGE Experience. Lead Investigator: Maureen Smith (NU)
85. PCA Loadings. Lead Investigator: Gerard Tromp (Geisinger) and David Crosslin (UW)
86. Portable Applications for Implementing Multi-Site Clinical NLP Algorithms. Lead Investigator: David Carrell (GroupHealth)
87. Evaluation of a Secure Multiparty Computation Protocol to Enable Genome-Phenome Meta-Analysis in the Cloud. Lead Investigator: Wei Xie (VU)
88. Knowledge Driven Search for Gene-Gene Interactions Associated with Cataract in the eMERGE network. Lead Investigator: Marylyn Ritchie (MC/EIRH/PSU & CC)
89. Evaluation of a Differentially Private Top-k SNP Publication Strategy. Lead Investigator: Mehmet Kuzu (External Collaborator); Brad Malin (VU)
90. Effective Use of Electronic Health Records to Identify Venous Thromboembolism: Results from the eMERGE Network. Lead Investigator: Jyoti Pathak (Mayo)
91. GWAS of Infection or Colonization with Community Associated Methicillin-Resistant Staphylococcus Aureus (CA-MRSA). Lead Investigator: Abel Kho (NU)
92. A Trial of Pre-emptive Pharmacogenetic Genotyping: The PGx Project of the eMERGE Network. Lead Investigator: Laura Rasmussen-Torvik (NU)
93. Genetic Risk Factors for Development of Diverticulitis. Lead Investigator: Abel Kho (NU)
94. Diverticulosis. Lead Investigator: Will Thompson (NU)
95. The Geographic Distribution of Colon Polyps. Lead Investigator: Will Thompson (NU)
96. Colon Polyps. Lead Investigator: Abel Kho (NU)
97. Genome-wide Association Study of Extreme Obesity Defined by Electronic Medical Record Phenotyping. Lead Investigator: Glenn Gerhard (Geisinger)
98. A Collaborative Approach to Develop an Electronic Health Record Phenotyping Algorithm for Drug-Induced Liver Injury. Lead Investigator: Casey Overby and Chunhua Weng (Columbia)
99. Genetic Variants Associated with Response to Heart Failure Treatment: The Electronic Medical Records and Genomics (eMERGE) Network. Lead Investigator: Sue Bielinski (Mayo)
100. Genome Wide Association of Risk of Heart Failure: The Electronic Medical Records and Genomics (eMERGE) Network. Lead Investigator: Sue Bielinski (Mayo)
101. Using Electronic Health Records to Identify Heart Failure Cohorts with Differentiation for Preserved and Reduced Ejection Fraction in Primary Care and Biobank Populations. Lead Investigator: Sue Bielinski (Mayo)
102. Penetrance of Hemochromatosis (HFE) RS1799945 (H63D) and RS1800562 (C282Y) Homozygosity and Compound Heterozygosity. Lead Investigators: Carlos Gallego (UW), Daniel Kim (UW), Josh Denny (VU), and Maureen Smith (NU)
103. Meta-Analysis of Genome-Wide Association Studies for Abdominal Aortic Aneurysms. Lead Investigator: Greg Jones (External Collaborator) and Helena Kuivaniemi (Geisinger)

104. Genetic Variation that Predicts Susceptibility to Clostridium Difficile. Lead Investigator: Josh Denny (VU)
105. Genetic Variation that Predicts Susceptibility to Onychomycosis. Lead Investigator: David Carrell (GroupHealth)
106. GWAS of Venous Thromboembolism (VTE) among White Americans. Lead Investigator: John Heit (Mayo)
107. GWAS of Venous Thromboembolism (VTE) among African-Americans. Lead Investigator: John Heit (Mayo)
108. Genome-Wide Association Study of Abdominal Aortic Aneurysms with Electronic Medical Record Phenotyping. Lead Investigators: Helena Kuivaniemi and Gerard Tromp (Geisinger)

Published Phase I Manuscripts during Phase II

1. Ritchie MD, Verma SS, Hall MA, Goodloe RJ, Berg RL, et al. Electronic medical records and genomics (eMERGE) network exploration in cataract: Several new potential susceptibility loci. Molecular Vision 2014; 20:1281-1295
2. Shameer K, Denny JC, Ding K, Jouni H, Crosslin DR, et al. A genome- and phenome-wide association study to identify genetic variants influencing platelet count and volume and their pleiotropic effects. Hum Genet. 2013 Sep 12. **PMID: 24026423 PMCID: PMC3880605**
3. Ding K, de Andrade M, Manolio T, Crawford D, Rasmussen-Torvik L, et al. Genetic Variants that Confer Resistance to Malaria Are Associated with Red Blood Cell Traits in African Americans: An Electronic Medical Record-based Genome Wide Association Study. G3 (Bethesda). 2013 May 20. pii: g3.113.006452v1. doi: 10.1534/g3.113.006452. **PMID: 23696099 PMCID: PMC3704235**
4. Pacheco JA, Wilke RA, Thompson WK, Ritchie MD, Kho AN, et al. High Density GWAS for LDL Cholesterol in African Americans using Electronic Medical Records reveals a strong protective variant in APOE. Clin Transl Sci. 2012 Oct;5(5):394-399. doi: 10.1111/j.1752-8062.2012.00446.x. Epub 2012 Aug 23. **PMID: 23067351 PMCID: PMC3521536**
5. Ritchie MD, Denny JC, Zuvich RL, Crawford DC, Schildcrout JS, et al. Genome- and Phenome-Wide Analysis of Cardiac Conduction Identifies Markers of Arrhythmia Risk. Circulation. 2013 Mar 5. **PMID: 23463857 PMCID: PMC3713791**
6. Shameer K, Jouni H, Masys DR, Jarvik GP, Kho AN, et al. Genetic Loci implicated in erythroid differentiation and cell cycle regulation are associated with red blood cell traits. Mayo Clin Proc. 2012 May;87(5):461-74. **PMID: 22560525 PMCID: PMC3538470**
7. McDavid A, Weston N, Nelson SC, Zheng X, Hart E, et al. Genetic variants associated with the white blood cell count in 13,923 subjects in the eMERGE Network. Hum Genet. 2012 Apr;131(4):639-52. doi: 10.1007/s00439-011-1103-9. Epub 2011 Oct 30. **PMID: 22037903 PMCID: PMC3640990**
8. Hayes MG, Rasmussen-Torvik LJ, Pacheco JA, Armstrong LL, Denny JC, et al. Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study. J Am Med Inform Assoc. 2012 Mar-Apr;19(2):212-8. Epub 2011 Nov 19. **PMID: 22101970 PMCID: PMC3277617**
9. Rasmussen LV, Berg RL, Linneman JG, McCarty CA, Waudby C, et al. Importance of multi-modal approaches to effectively identify cataract cases from electronic health records. J Am Med Inform Assoc. 2012 Mar-Apr;19(2):225-34. **PMID: 22319176 PMCID: PMC3277618**
10. Leibson CL, Ransom JE, Kho AN, Caraballo PJ, Chai HS, et al. Impact of data fragmentation across healthcare centers on the accuracy of a high-throughput clinical phenotyping algorithm for specifying subjects with type 2 diabetes mellitus. J Am Med Inform Assoc. 2012 Mar-Apr;19(2):219-24. Epub 2012 Jan 16. **PMID: 22249968 PMCID: PMC3277630**
11. Armstrong LL, Bielinski SJ, Bradford Y, Carlson CS, Crawford DC, et al. Pitfalls of merging GWAS data: lessons learned in the eMERGE network and quality control procedures to maintain high data quality. Genet Epidemiol. 2011 Dec;35(8):887-98. doi: 10.1002/gepi.20639. **PMID: 22125226 PMCID: PMC3592376**
12. Crawford DC, Ritchie MD, Bielinski SJ, Basford MA, Bradford Y, et al. Variants Near FOXE1 Are Associated with Hypothyroidism and Other Thyroid Conditions: Using Electronic Medical Records for Genome- and Phenome-wide Studies. American Journal of Human Genetics. 2011 Oct:89(4):529-42. **PMID: 21981779 PMCID: PMC3188836**
13. Wang J, Kashyap S, Basford M, Li R, Masys DR, Chute CG. Mapping clinical phenotype data elements to standardized metadata repositories and controlled terminologies: the eMERGE Network experience. J Am Med Inform Assoc. 2011 Jul-Aug;18(4):376-86. **PMID: 21597104 PMCID: PMC3128396**
14. Basford M, Dressler LG, Fullerton SM, Koenig BA, Li R, et al. Ethical and practical challenges of sharing data from genome-wide association studies: The eMERGE Consortium experience. Genome Res. 2011 Jul;21(7):1001-7. Epub 2011 Jun 1. **PMID: 21632745 PMCID: PMC3129243**
15. Pacheco JA, Peissig PL, Rasmussen L, Newton KM, Weston N, et al. Electronic medical records for genetic research: results of the eMERGE consortium. Sci Transl Med. 2011 Apr 20;3(79):79re1. **PMID: 21508311 PMCID: PMC3690272**
16. Peissig P, Kho A, Bielinski S, Berg R, Choudhary V, et al. Validation of electronic medical record-based phenotyping algorithms: lessons learned from the eMERGE Network. J Am Med Inform Assoc doi:10.1136/amiajnl-2012-000896. **PMID: 23531748 PMCID: PMC3715338**
17. McDavid A, Crane P, Weston N, Ehrlich K, Newton K, et al. Confirmation of the Reported Association of Clonal Chromosomal Mosaicism with an Increased Risk of Incident Hematologic Cancer. PLoS ONE 8(3): e59823. doi:10.1371/journal.pone.0059823. **PMID: 23533652 PMCID: PMC3606281**

Published/Accepted Phase II Site Specific Manuscripts

BCH/CCHMC

1. Namjou B, Ni Y, Harley IT, Chepelev I, Cobb B, et al. The effect of inversion at 8p23 on BLK association with lupus in Caucasian population. PLoS One. 2014 Dec 29;9(12):e115614. **PMID: 25545785 PMCID: PMC4278715**
2. Kaufman KM...Harley JB, Nirmala NR, Grom AA. Whole-Exome Sequencing Reveals Overlap Between Macrophage Activation Syndrome in Systemic Juvenile Idiopathic Arthritis and Familial Hemophagocytic Lymphohistiocytosis. Arthritis & Rheumatology, 66: 3486–3495.
3. Holm IA. Clinical Management of Pediatric Genomic Testing. Current Genetic Medicine Reports, 2014, 2(4).
4. Li Q, Melton K, Lingren T, Kirkendall ES, Hall E. Phenotyping for patient safety: algorithm development for electronic health record based automated adverse event and medical error detection in neonatal intensive care. 2014, JAMIA 21(5)
5. Prows CA, Zhang X, Huth MM, Zhang K, Saldaña SN, Daraiseh NM, Esslinger HR, Freeman E, Greinwald JH, Martin LJ, Sadhasivam S. Codeine-related adverse drug reactions in children following tonsillectomy: a prospective study. Laryngoscope. 2014 May;124(5):1242–1250. **PMID: 24122716**
6. Zhai H, Brady P, Li Q, Lingren T, Ni Y, Wheeler DS, Solti I. Developing and evaluating a machine learning based algorithm to predict the need of pediatric intensive care unit transfer for newly hospitalized children. Resuscitation. 2014 Aug;85(8):1065–1071. **PMID: 24813568 PMCID: PMC4087062**
7. Morris DL, Fernando MMA, Taylor KE, Chung SA, Nititham J, Alarcón-Riquelme ME, Barcellos LF, Behrens TW, Cotsapas C, Gaffney PM, Graham RR, Pons-Estel BA, Gregersen PK, Harley JB, Hauser SL, Hom G, Langefeld CD, Noble JA, Rioux JD, Seldin MF, Systemic Lupus Erythematosus Genetics Consortium, Vyse TJ, Criswell LA. MHC associations with clinical and autoantibody manifestations in European SLE. Genes Immun. 2014 Apr;15(4):210–217. **PMID: 24598797 PMCID: PMC4102853**
8. 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; **PMID: 25034947**
9. Harley JB, Zoller EE. Editorial: What Caused All These Troubles, Anyway? Epstein-Barr Virus in Sjögren's Syndrome Reevaluated. Arthritis & Rheumatology (Hoboken, NJ). 2014 Sep;66(9):2328–2330. **PMID: 24891328**
10. Zhai H, Lingren T, Deleger L, Li Q, Kaiser M, Stoutenborough L, Solti I. Web 2.0-based crowdsourcing for high-quality gold standard development in clinical natural language processing. J Med Internet Res. 2013 Apr 2;15(4):e73. **PMID: 23548263 PMCID: PMC3636329**
11. Mangale D, Kariuki SN, Chrabot BS, Kumabe M, Kelly JA, Harley JB, James JA, Sivils KL, Niewold TB. Familial aggregation of high tumor necrosis factor alpha levels in systemic lupus erythematosus. Clin Dev Immunol. 2013 Sep 25;2013:267430. **PMID: 24187561 PMCID: PMC3800640**
12. Li Q, Zhai H, Deleger L, Lingren T, Kaiser M, Stoutenborough L, Solti I. A sequence labeling approach to link medications and their attributes in clinical notes and clinical trial announcements for information extraction. J Am Med Inform Assoc. 2013 Oct;20(5):915–921. **PMID: 23268488 PMCID: PMC3756265**
13. Li Q, Deleger L, Lingren T, Zhai H, Kaiser M, Stoutenborough L, Jegga AG, Cohen KB, Solti I. Mining FDA drug labels for medical conditions. BMC Med Inform Decis Mak. 2013 Apr 24;13:53. **PMID: 23617267 PMCID: PMC3646673**
14. Deleger L, Brodzinski H, Zhai H, Li Q, Lingren T, Kirkendall ES, Alessandrini E, Solti I. Developing and evaluating an automated appendicitis risk stratification algorithm for pediatric patients in the emergency department. J Am Med Inform Assoc. 2013 Dec;20(e2):e212–220. **PMID: 24130231 PMCID: PMC3861926**

CHOP

1. Sgariglia F, Pedrini E, Bradfield JP, Bhatti TR, D'Adamo P, Dormans JP, Gunawardena AT, Hakonarson H, Hecht JT, Sangiorgi L, Pacifici M, Enomoto-Iwamoto M, Grant SFA. The type 2 diabetes associated rs7903146 T allele within TCF7L2 is significantly under-represented in Hereditary Multiple Exostoses: Insights into pathogenesis. Bone. 2015 Mar 1;72:123–127. **PMID: 25498973**

2. Roy SM, Chesi A, Mentch F, Xiao R, Chiavacci R, Mitchell JA, Kelly A, Hakonarson H, Grant SFA, Zemel BS, McCormack SE. Body Mass Index (BMI) Trajectories in Infancy Differ by Population Ancestry and May Presage Disparities in Early Childhood Obesity. J Clin Endocrinol Metab. 2015 Jan 30;jc20144028. **PMID: 25636051**

3. Menezes MJ, Guo Y, Zhang J, Riley LG, Cooper ST, Thorburn DR, Li J, Dong D, Li Z, Glessner J, Davis RL, Sue CM, Alexander SI, Arbuckle S, Kirwan P, Keating BJ, Xu X, Hakonarson H, Christodoulou J. Mutation in mitochondrial ribosomal protein S7 (MRPS7) causes congenital sensorineural deafness, progressive hepatic and renal failure and lactic acidemia. Hum Mol Genet. 2015 Jan 2; **PMID: 25556185**

4. Mancini C, Nassani S, Guo Y, Chen Y, Giorgio E, Brussino A, Di Gregorio E, Cavalieri S, Lo Buono N, Funaro A, Pizio NR, Nmezi B, Kyttala A, Santorelli FM, Padiath QS, Hakonarson H, Zhang H, Brusco A. Adult-onset autosomal recessive ataxia associated with neuronal ceroid lipofuscinosis type 5 gene (CLN5) mutations. J Neurol. 2015 Jan;262(1):173–178. **PMID: 25359263**

5. Maier R, Moser G, Chen G-B, Ripke S, Cross-Disorder Working Group of the Psychiatric Genomics Consortium, Coryell W, Potash JB, Scheftner WA, Shi J, Weissman MM, Hultman CM, Landén M, Levinson DF, Kendler KS, Smoller JW, Wray NR, Lee SH, Cross-Disorder Working Group of the Psychiatric Genomics Consortium. Joint analysis of psychiatric disorders increases accuracy of risk prediction for schizophrenia, bipolar disorder, and major depressive disorder. Am J Hum Genet. 2015 Feb 5;96(2):283–294. **PMID: 25640677 PMCID: PMC4320268**

6. Guo Y, Menezes MJ, Menezes MP, Liang J, Li D, Riley LG, Clarke NF, Andrews PI, Tian L, Webster R, Wang F, Liu X, Shen Y, Thorburn DR, Keating BJ, Engel A, Hakonarson H, Christodoulou J, Xu X. Delayed diagnosis of congenital myasthenia due to associated mitochondrial enzyme defect. Neuromuscul Disord. 2015 Mar;25(3):257–261. **PMID: 25557462**

7. Zamzow J, Culnan E, Spiers M, Calkins M, Satterthwaite T, Ruparel K, Abrams D, Chiavacci R, Hakonarson H, Gur R. B-37The Relationship between Body Mass Index and Executive Function from Late Childhood through Adolescence. Arch Clin Neuropsychol. 2014 Sep;29(6):550. **PMID: 25176787**

8. White PS, Xie HM, Werner P, Glessner J, Latney B, Hakonarson H, Goldmuntz E. Analysis of chromosomal structural variation in patients with congenital left-sided cardiac lesions. Birth Defects Res Part A Clin Mol Teratol. 2014 Jul 26; **PMID: 25066379**

9. Swerdlow DI, Preiss D, Kuchenbaecker KB, Holmes MV, Engmann JEL, et al. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. Lancet. 2014 Sep 24; **PMID: 25262344**

10. St Pourcain B, Haworth CMA, Davis OSP, Wang K, Timpson NJ, Evans DM, Kemp JP, Ronald A, Price T, Meaburn E, Ring SM, Golding J, Hakonarson H, Plomin R, Davey Smith G. Heritability and genome-wide analyses of problematic peer relationships during childhood and adolescence. Hum Genet. 2014 Dec 17; **PMID: 25515860**

11. Soleimanpour SA, Gupta A, Bakay M, Ferrari AM, Groff DN, Fadista J, Spruce LA, Kushner JA, Groop L, Seeholzer SH, Kaufman BA, Hakonarson H, Stoffers DA. The diabetes susceptibility gene Clec16a regulates mitophagy. Cell. 2014 Jun 19;157(7):1577–1590. **PMID: 24949970 PMCID: PMC4184276**

12. Sleiman PMA, Wang M-L, Cianferoni A, Aceves S, Gonsalves N, Nadeau K, Bredenoord AJ, Furuta GT, Spergel JM, Hakonarson H. GWAS identifies four novel eosinophilic esophagitis loci. Nat Commun [Internet]. 2014 Nov 19 [cited 2015 Mar 27];5. Available from: http://www.nature.com/ncomms/2014/141119/ncomms6593/full/ncomms6593.html

13. Scott D. Cook-Sather JL. Modulatory effects of TAOK3 variants on morphine requirement in acute postoperative pain: An early genome wide association study contribution to the field of pediatric pain. Pain. 2014;155(11).

14. Robinson EB, Kirby A, Ruparel K, Yang J, McGrath L, Anttila V, Neale BM, Merikangas K, Lehner T, Sleiman PMA, Daly MJ, Gur R, Gur R, Hakonarson H. The genetic architecture of pediatric cognitive abilities in the Philadelphia Neurodevelopmental Cohort. Mol Psychiatry. 2014 Jul 15; **PMID: 25023143**

15. Prokudin I, Li D, He S, Guo Y, Goodwin L, Wilson M, Rose L, Tian L, Chen Y, Liang J, Keating B, Xu X, Jamieson RV, Hakonarson H. Value of whole exome sequencing for syndromic retinal dystrophy diagnosis in young patients. Clin Experiment Ophthalmol. 2014 Jul 24; **PMID: 25060287**

16. Pizzino A, Pierson TM, Guo Y, Helman G, Fortini S, et al. TUBB4A de novo mutations cause isolated hypomyelination. Neurology. 2014 Sep 2;83(10):898–902. **PMID: 25085639 PMCID: PMC4153852**

17. Picard M, Zhang J, Hancock S, Derbeneva O, Golhar R, et al. Progressive increase in mtDNA 3243A>G heteroplasmy causes abrupt transcriptional reprogramming. Proc Natl Acad Sci USA. 2014 Sep 23;111(38):E4033–4042. **PMID: 25192935 PMCID: PMC4183335**

18. Pellegrino R, Kavakli IH, Goel N, Cardinale CJ, Dinges DF, et al. A novel BHLHE41 variant is associated with short sleep and resistance to sleep deprivation in humans. Sleep. 2014;37(8):1327–1336. **PMID: 25083013 PMCID: PMC4096202**

19. Naj AC, Jun G, Reitz C, Kunkle BW, Perry W, et al. Effects of multiple genetic Loci on age at onset in late-onset Alzheimer disease: a genome-wide association study. JAMA Neurol. 2014 Nov 1;71(11):1394–1404. **PMID: 25199842**

20. Moore TM, Reise SP, Gur RE, Hakonarson H, Gur RC. Psychometric Properties of the Penn Computerized Neurocognitive Battery. Neuropsychology. 2014 Sep 1; **PMID: 25180981**

21. Mitchell LE, Agopian AJ, Bhalla A, Glessner JT, Kim CE, et al. Genome-wide association study of maternal and inherited effects on left-sided cardiac malformations. Hum Mol Genet. 2014 Aug 18; **PMID: 25138779**

22. Mestdagh P, Hartmann N, Baeriswyl L, Andreasen D, Bernard N, et al. Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study. Nat Methods. 2014 Aug;11(8):809–815. **PMID: 24973947**

23. Matsunami N, Hensel CH, Baird L, Stevens J, Otterud B, et al. Identification of rare DNA sequence variants in high-risk autism families and their prevalence in a large case/control population. Mol Autism. 2014 Jan 27;5(1):5. **PMID: 24467814**

24. Holmes MV, Dale CE, Zuccolo L, Silverwood RJ, Guo Y, et al. Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. BMJ. 2014;349:g4164. **PMID: 25011450 PMCID: PMC4091648**

25. Hadley D, Wu Z-L, Kao C, Kini A, Mohamed-Hadley A, et al. The impact of the metabotropic glutamate receptor and other gene family interaction networks on autism. Nat Commun. 2014;5:4074. **PMID: 24927284 PMCID: PMC4059929**

26. Guo Y, Kartawinata M, Li J, Pickett HA, Teo J, et al. Inherited bone marrow failure associated with germline mutation of ACD, the gene encoding telomere protein TPP1. Blood. 2014 Oct 30;124(18):2767–2774. **PMID: 25205116 PMCID: PMC4215308**

27. Glessner JT, Bick AG, Ito K, Homsy JG, Rodriguez-Murillo L, et al. Increased frequency of de novo copy number variants in congenital heart disease by integrative analysis of single nucleotide polymorphism array and exome sequence data. Circ Res. 2014 Oct 24;115(10):884–896. **PMID: 25205790 PMCID: PMC4209190**

28. Escott-Price V, Bellenguez C, Wang L-S, Choi S-H, Harold D, et al. Gene-wide analysis detects two new susceptibility genes for Alzheimer's disease. PLoS ONE. 2014;9(6):e94661. **PMID: 24922517 PMCID: PMC4055488**

29. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature. 2014 Nov 13;515(7526):209–215.

30. Culnan E, Zamzow J, Spiers M, Calkins M, Satterthwaite T, et al. B-36Relationships between Body Mass Index and Social Cognition among 8-19 Year-Olds. Arch Clin Neuropsychol. 2014 Sep;29(6):550. **PMID: 25176786**

31. Cook-Sather SD, Li J, Goebel TK, Sussman EM, Rehman MA, Hakonarson H. TAOK3, a novel genome-wide association study locus associated with morphine requirement and postoperative pain in a retrospective pediatric day surgery population. Pain. 2014 Sep;155(9):1773–1783. **PMID: 24909733 PMCID: PMC4157963**

32. Connolly JJ, Hakonarson H. Etiology of autism spectrum disorder: a genomics perspective. Curr Psychiatry Rep. 2014 Nov;16(11):501. **PMID: 25212713**

33. Calkins ME, Moore TM, Merikangas KR, Burstein M, Satterthwaite TD, et al. The psychosis spectrum in a young U.S. community sample: findings from the Philadelphia Neurodevelopmental Cohort. World Psychiatry. 2014 Oct;13(3):296–305. **PMID: 25273303**

34. Banerjee A, Wang H-Y, Borgmann-Winter KE, MacDonald ML, Kaprielian H, et al. Src kinase as a mediator of convergent molecular abnormalities leading to NMDAR hypoactivity in schizophrenia. Mol Psychiatry. 2014 Oct 21; **PMID: 25330739**

35. Almoguera B, He S, Corton M, Jose PF-S, Blanco-Kelly F, et al. Expanding the phenotype of PRPS1 syndromes in females: neuropathy, hearing loss and retinopathy. Orphanet Journal of Rare Diseases. 2014 Dec 10;9(1):190.

36. St Pourcain B, Skuse DH, Mandy WP, Wang K, Hakonarson H et al. Variability in the common genetic architecture of social-communication spectrum phenotypes during childhood and adolescence. Mol Autism. 2014 Feb 24;5:18. **PMID: 24564958PMCID: PMC3940728**

37. Purkey MT, Li J, Mentch F, Grant SFA, Desrosiers M, Hakonarson H, Toskala E. Genetic Variation in Genes Encoding Airway Epithelial Potassium Channels Is Associated with Chronic Rhinosinusitis in a Pediatric Population. PLoS One. 2014 Mar 3;9(3). **PMID: 24595210PMCID: PMC3940609**

38. Matsunami N, Hensel CH, Baird L, Stevens J, Otterud B, et al. Identification of rare DNA sequence variants in high-risk autism families and their prevalence in a large case/control population. Mol Autism. 2014;5(1):5. **PMID: 24467814**

39. Keller M, Glessner J, Resnick E, Perez E, Chapel H, et al. Burden of Copy Number Variation in 14 Common Variable Immunodeficiency. Clin Exp Immunol. 2013 Dec 13; **PMID: 24329717** 3. Cardinale CJ, Kelsen JR, Baldassano RN, Hakonarson H. Impact of exome sequencing in inflammatory bowel disease. World J Gastroenterol. 2013 Oct 28;19(40):6721–6729. P**MID: 24187447, PMCID: PMC3812471**

40. Kasperaviciute D, Catarino CB, Matarin M, Leu C, Novy J, et al. Epilepsy, hippocampal sclerosis and febrile seizures linked by common genetic variation around SCN1A. Brain. 2013 Oct;136(Pt 10):3140–3150. **PMID: 24014518 PMCID: PMC3784283**

41. St Pourcain B, Whitehouse AJO, Ang WQ, Warrington NM, Glessner JT, et al. Common variation contributes to the genetic architecture of social communication traits. Mol Autism. 2013 Sep 18;4(1):34. **PMID: 24047820 PMCID: PMC3853437**

42. Ong BA, Li J, McDonough JM, Wei Z, Kim C, et al. Gene network analysis in a pediatric cohort identifies novel lung function genes. PLoS ONE. 2013 Sep 2;8(9):e72899. **PMID: 24023788 PMCID: PMC3759429**

43. Mortensen LJ, Kreiner-Møller E, Hakonarson H, Bønnelykke K, Bisgaard H. The PCDH1-gene and asthma in early childhood. Eur Respir J. 2013 Aug 29; **PMID: 23988763**

44. Hamshere ML, Langley K, Martin J, Agha SS, Stergiakouli E,et al. High loading of polygenic risk for ADHD in children with comorbid aggression. Am J Psychiatry. 2013 Aug 1;170(8):909–916. **PMID: 23599091 PMCID: PMC3935265**

45. Monda KL, Chen GK, Taylor KC, Palmer C, Edwards TL, et al. A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. Nat Genet. 2013 Jun;45(6):690–696. PMID: 23583978 PMCID: PMC3694490

46. Shi L, Zhang X, Golhar R, Otieno FG, He M, Hou C, Kim C, Keating B, Lyon GJ, Wang K, Hakonarson H. Whole-genome sequencing in an autism multiplex family. Mol Autism. 2013 Apr 18;4(1):8. **PMID: 23597238 PMCID: PMC3642023**

47. March ME, Sleiman PM, Hakonarson H. Genetic polymorphisms and associated susceptibility to asthma. Int J Gen Med. 2013 Apr 17;6:253–265. **PMID: 23637549 PMCID: PMC3636804**

48. JA, Hakonarson H, Rebbeck TR, Grant SFA. Obesity-susceptibility loci and the tails of the pediatric BMI distribution. Obesity (Silver Spring). 2013 Feb 14; **PMID: 23408508 PMCID: PMC3661695**

49. Glessner JT, Li J, Hakonarson H. ParseCNV integrative copy number variation association software with quality tracking. Nucleic Acids Res. 2013 Jan 4; **PMID: 23293001 PMCID: PMC3597648**

50. Guo Y, Lanktree MB, Taylor KC, Hakonarson H, Lange LA, Keating BJ. Gene-centric meta-analyses of 108 912 individuals confirm known body mass index loci and reveal three novel signals. Hum Mol Genet. 2013 Jan 1;22(1):184–201. **PMID: 23001569 PMCID: PMC3522401**

51. Asselbergs FW, Guo Y, van Iperen EPA, Sivapalaratnam S, Tragante V, et al. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. Am J Hum Genet. 2012 Nov 2;91(5):823–838. **PMID: 23063622 PMCID: PMC3487124**

Geisinger

1. Williams, MS. Genomic Medicine Implementation: Learning by Example. Am J Med Genet C Semin Med Genet. 2014 166:8-14.

2. Wade JE, Ledbetter DH, Williams MS. Implementation of Genomic Medicine in a Health Care Delivery System: A Value Proposition? Am J Med Genet C Semin Med Genet 2014 166:112-116

3. Brownstein CA, …, Williams MS, … and Margulies DM. An international effort towards developing standards for best practices in analysis, interpretation and reporting of clinical genome sequencing results in the CLARITY Challenge. Genome Biol 2014 Mar 25;15(3):R53. **PMID: 24667040 PMCID: PMC4073084**

4. Schully SD, Dotson WD, Lam TK, Chang CQ, Aronson N, et al. Evidence Synthesis and Guideline Development in Genomic Medicine: Current Status and Future Prospects. Genet Med 2014 Jun 19. **PMID: 24946156**

5. Segal MM, Abdellateef M, El-Hattab AW, Hilbush BS, De La Vega FM. Clinical pertinence metric enables hypothesis-independent genome-phenome analysis for neurological diagnosis. J Child Neurol published online 24 August 2014. **PMID: 25156663**

6. Snyder SR, Mitropoulou C, Patrinos GP, Williams MS. Economic evaluation of pharmacogenomics: A value-based approach to pragmatic decision-making in the face of complexity. Public Health Genomics 2014 Sep 26. **PMID: 25278172**

7. Bock JA, Fairley KJ, Smith RE, Maeng DD, Pitcavage JM, et al. Cost-Effectiveness of IL28B Genotype-Guided Protease Inhibitor Triple Therapy versus Standard of Care Treatment in Patients with Hepatitis C Genotypes 2 or 3 Infection. Public Health Genomics. 2014 Sep 18; **PMID: 25247313**

8. Williams JL, Faucett WA, Smith-Packard B, Wagner M, Williams MS. An Assessment of Time Involved in Pre-test Case Review and Counseling for a Whole Genome Sequencing Clinical Research Program. J Genet Couns. 2014 Feb 27; **PMID: 24573557**

9. Caudle KE, Klein TE, Hoffman JM, Müller DJ, Whirl-Carrillo M, et al. Incorporation of Pharmacogenomics into Routine Clinical Practice: the Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline Development Process. Curr Drug Metab. 2014 Jan 30; **PMID: 24479687 PMCID: PMC3977533**

10. Lipsky EB, King BR, Tromp G. Node-Oriented Workflow (NOW): A Command Template Workflow Management Tool for High Throughput Data Analysis Pipelines. Journal of Datamining in Genomics & Proteomics. 2014;5(155).

11. Muir AJ, Gong L, Johnson SG, Lee MTM, Williams MS, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for IFNL3 (IL28B) Genotype and PEG Interferon-α-Based Regimens. Clin Pharmacol Ther. 2013 Oct 4; **PMID: 24096968 PMCID: PMC3904555**

12. Sun X, Elston R, Morris N, Zhu X. What Is the Significance of Difference in Phenotypic Variability across SNP Genotypes? Am J Hum Genet. 2013 Jul 31; **PMID: 23910463 PMCID: PMC3738833**

13. Jones GT, Bown MJ, Gretarsdottir S, Romaine SPR, Helgadottir A, et al. A sequence variant associated with sortilin-1 (SORT1) on 1p13.3 is independently associated with abdominal aortic aneurysm. Hum Mol Genet. 2013 Apr 9; **PMID: 23535823 PMCID: PMC3690970**

14. Wang X, Morris NJ, Zhu X, Elston RC. A variance component based multi-marker association test using family and unrelated data. BMC Genet. 2013 Mar 4 ;14:17. **PMID: 23497289 PMCID: PMC3614458**

15. Zhu X, Feng T, Elston RC. Linkage-disequilibrium-based binning misleads the interpretation of genome-wide association studies. Am J Hum Genet. 2012 Nov 2; 91(5):965–968; author reply 969–970. **PMID: 23122590 PMCID: PMC3487138**

16. Lu Q, Wei C, Ye C, Li M, Elston RC. A likelihood ratio-based Mann-Whitney approach finds novel replicable joint gene action for type 2 diabetes. Genet Epidemiol. 2012 Sep; 36(6):583–593. **PMID: 22760990 PMCID: PMC3634342**

17. Wang X, Morris NJ, Schaid DJ, Elston RC. Power of single- vs. multi-marker tests of association. Genet Epidemiol. 2012 Jul; 36(5):480–487. **PMID: 22648939 PMCID: PMC3708310**

GroupHealth/UW

1. Escott-Price, V., et al. (2014) Gene-wide analysis detects two new susceptibility genes for Alzheimer's disease. 2014. PLOS One (In Press)

2. Naj, A.C., et al. Age-at-Onset in Late Onset Alzheimer Disease is Modified by Multiple Genetic Loci. 2014. JAMA Neurol (In Press)

3. Nelson Peter T., et al. GWAS indicates that ABCC9 -- a drug target -- is linked to hippocampal sclerosis of aging, which is a prevalent cause of dementia. Acta Neuropathologica 2014. (In Press)

4. Ruiz, A., et al. International Genomics of Alzheimer's Disease Project identifies TRIP4 as a novel susceptibility gene. 2014. Translational Psychiatry (In Press).

5. Benitez BA, Jin SC, Guerreiro R, Graham R, Lord J, et al. Missense variant in TREML2 protects against Alzheimer's disease. Neurobiol Aging. 2014 Jun; 35(6):1510.e19–26. **PMID: 24439484 PMCID: PMC3961557**

6. Burke W, Evans BJ, Jarvik GP. Return of results: Ethical and legal distinctions between research and clinical care. Am J Med Genet C Semin Med Genet. 2014 Mar;166(1):105–111. **PMID: 24616381**

7. Sherva R, Tripodis Y, Bennett DA, Chibnik LB, Crane PK, de Jager PL, Farrer LA, Saykin AJ, Shulman JM, Naj A, Green RC, GENAROAD Consortium, Alzheimer's Disease Neuroimaging Initiative, Alzheimer's Disease Genetics Consortium. Genome-wide association study of the rate of cognitive decline in Alzheimer's disease. Alzheimers Dement. 2014 Jan; 10(1):45–52. **PMID: 1623535033 PMCID: PMC3760995**

8. 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; **PMID: 24357849**

9. Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013 Dec; 45(12):1452–1458. **PMID: 24162737 PMCID: PMC3896259**

10. Burke W, Matheny Antommaria AH, Bennett R, Botkin J, et al. Recommendations for returning genomic incidental findings? We need to talk! Genet Med. 2013 Nov; 15(11):854–859. **PMID: 23907645 PMCID:PMCID: PMC3832423**

11. 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**

12. Ridge PG, Mukherjee S, Crane PK, Kauwe JSK, Alzheimer's Disease Genetics Consortium. Alzheimer's disease: analyzing the missing heritability. PLoS ONE. 2013; 8(11):e79771. **PMID: 24244562 PMCID: PMC3820606**

13. Reitz C, Tosto G, Vardarajan B, Rogaeva E, Ghani M, et al. Independent and epistatic effects of variants in VPS10-d receptors on Alzheimer disease risk and processing of the amyloid precursor protein (APP). Transl Psychiatry. 2013; 3:e256. **PMID: 23674637 PMCID: PMC3669917**

14. Reitz C, Mayeux R, Alzheimer's Disease Genetics Consortium. TREM2 and neurodegenerative disease. N Engl J Med. 2013 Oct 17; 369(16):1564–1565. **PMID: 24131184 PMCID: PMC3980568**

15. Lin C-F, Valladares O, Childress DM, Klevak E, Geller ET, et al. DRAW+SneakPeek: analysis workflow and quality metric management for DNA-seq experiments. Bioinformatics. 2013 Oct 1; 29(19):2498–2500. **PMID: 23943636 PMCID: PMC3777113**

16. Cruchaga C, Kauwe JSK, Harari O, Jin SC, et al. GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. Neuron. 2013 Apr 24; 78(2):256–268. **PMID: 23562540 PMCID: PMC3664945**

17. Miyashita A, Koike A, Jun G, Wang L-S, Takahashi S, et al. SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians. PLoS ONE. 2013; 8(4):e58618. **PMID: 23565137 PMCID: PMC3614978**

18. Whitcomb DC, LaRusch J, Krasinskas AM, Klei L, Smith JP, et al. Common genetic variants in the CLDN2 and PRSS1-PRSS2 loci alter risk for alcohol-related and sporadic pancreatitis. Nat Genet. 2012 Dec; 44(12):1349–1354. **PMID: 23143602 PMCID: PMC3510344**

19. Coppola G, Chinnathambi S, Lee JJ, Dombroski BA, Baker MC, et al. Evidence for a role of the rare p.A152T variant in MAPT in increasing the risk for FTD-spectrum and Alzheimer's diseases. Hum Mol Genet. 2012 Aug 1;21(15):3500–3512. **PMID: 22556362 PMCID: PMC3392107**

20. Allen M, Zou F, Chai HS, Younkin CS, Crook J, et al. Novel late-onset Alzheimer disease loci variants associate with brain gene expression. Neurology. 2012 Jul 17;79(3):221–228. **PMID: 22722634 PMCID: PMC3398432**

21. Schellenberg GD, Montine TJ. The genetics and neuropathology of Alzheimer's disease. Acta Neuropathol. 2012 Sep;124(3):305–323. **PMID: 22618995 PMCID: PMC3708460**

22. Floyd JS, Heckbert SR, Weiss NS, Carrell DS, Psaty BM. Use of administrative data to estimate the incidence of statin-related rhabdomyolysis. JAMA. 2012 Apr 18;307(15):1580–1582. **PMID: 2251168117 PMCID: PMC3483067**

Marshfield/Essentia/PSU

1. Verma SS, Peissig P, Cross D, Waudby C, Brilliant M et al. Benefits of accurate imputations in genome-wide association studies. EvoBio. 2014. (Accepted)

2. McCarty CA, Berg R, Waudby C, Foth W, Kitchner T, Cross D. Long-term recall of elements of informed consent: a pilot study comparing tranditional and computer-based consenting. IRB: Ethics & Hum Research. (Accepted)

3. Krauss, McCarty CA, Berg R, Rotter, Ritchie MD et al. Dose-response curves extracted from comprehensive electronic medical records identify the SORT1 gene locus as a novel determinant of atorvastatin potency (ED50) in a clinical practice-based setting. JAMA (Accepted)

4. Barrie ES, Weinshenker D, Verma A, Pendergrass S, Lange L, et al. Regulatory Polymorphisms in Human DBH Affect Peripheral Gene Expression and Sympathetic Activity. Circ Res. 2014 Oct 17. pii: CIRCRESAHA.114.304398. **PMID: 25326128**

5. SEQC/MAQC-III Consortium, A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. Nat Biotechnol. 2014 Sep;32(9):903-14

6. Pasquale LR, Loomis SJ, Weinreb RN, Kang JH, Yaspan BL, et al. Estrogen pathway polymorphisms in relation to primary open angle glaucoma: an analysis accounting for gender from the United States. Mol Vis. 2013 Jul 12;19:1471-81. Print 2013. **PMID: 23869166 PMCID: PMC3712669**

7. Hebbring SJ, Schrodi SJ, Ye Z, Zhou Z, Page D. A PheWAS approach in studying HLA-DRB1*1501. Genes Immun. 2013 Apr;14(3):187-91. **PMID: 23392276 PMCID: PMC3637423**

8. Ye Z, Mayer J, Ivacic L, Zhou Z, He M, Schrodi SJ, Page D, Brilliant MH, Hebbring SJ. Phenome-wide association studies (PheWASs) for functional variants. Eur J Hum Genet. 2014 Jul 30; **PMID: 25074467**

9. McCarty CA, Huggins W, Aiello AE, Bilder RM, Hariri A, Jernigan TL, Newman E, Sanghera DK, Strauman TJ, Zeng Y, Ramos EM, Junkins HA, PhenX RISING network. PhenX RISING: real world implementation and sharing of PhenX measures. BMC Med Genomics. 2014;7:16. **PMID: 24650325 PMCID: PMC3994539**

10. Land ME, Cooper RF, Young J, Berg E, Kitchner T, Xiang Q, Szabo A, Ivacic LC, Stepien KE, Page CD, Carroll J, Connor T, Brilliant M. Cone structure in subjects with known genetic relative risk for AMD. Optom Vis Sci. 2014 Aug;91(8):939–949. **PMID: 25014365 PMCID: PMC4111779**

11. Hall MA, Dudek SM, Goodloe R, Crawford DC, et al. Environment-wide association study (EWAS) for type 2 diabetes in the Marshfield Personalized Medicine Research Project Biobank. Pac Symp Biocomput. 2014;200–211. **PMID: 24297547**

12. Mahnke AN, Plasek JM, Hoffman DG, Partridge NS, Foth WS, et al. A rural community's involvement in the design and usability testing of a computer-based informed consent process for the personalized medicine research project. Am J Med Genet A. 2013 Nov 22; **PMID: 24273095**

13. Maenner MJ, Baker MW, Broman KW, Tian J, Barnes JK, et al. FMR1 CGG expansions: prevalence and sex ratios. Am J Med Genet B Neuropsychiatr Genet. 2013 Jul;162B(5):466–473. **PMID: 23740716 PMCID: PMC3885228**

14. Samwald M, Freimuth R, Luciano JS, Lin S, Powers RL, et al. An RDF/OWL knowledge base for query answering and decision support in clinical pharmacogenetics. Stud Health Technol Inform. 2013;192:539–542. **PMID: 23920613**

15. Pendergrass SA, Verma SS, Holzinger ER, Moore CB, Wallace J, et al. Next-generation analysis of cataracts: determining knowledge driven gene-gene interactions using biofilter, and gene-environment interactions using the phenx toolkit. Pac Symp Biocomput. 2013;147–158. **PMID: 23424120 PMCID: PMC3615413**

16. Pendergrass SA, Frase A, Wallace J, Wolfe D, Katiyar N, et al. Genomic analyses with biofilter 2.0: knowledge driven filtering, annotation, and model development. BioData Min. 2013;6(1):25. **PMID: 24378202 PMCID: PMC3917600**

17. Foth W, Waudby C, Brilliant MH. Certificates of confidentiality and the Marshfield Clinic's Personalized Medicine Research Project. Virtual Mentor. 2012;14(8):653–656. **PMID: 23351322 PMCID: PMC3698606**

Mayo

1. Kullo IJ, Leeper N. The Genetic Basis of Peripheral Arterial Disease: Current Knowledge, Challenges and Future Directions. Circ Res. (Accepted)

2. Ridgeway JL, Han LC, Olson JE, Lackore KA, Koenig BA, , et al. Demographic characteristics distinguished participants and non-participants of a clinic-based biobank and reasons for non-participation differed between refusers and non-responders. (Submitted to Journal of Clinical Epidemiology)

3. Ostrem MR, Formea CM, McCormick J, Abdalrhim AD, Han LC, et al. A patient-centered approach to the development and pilot of a warfarin pharmacogenomics patient education tool for health professionals. Curr Pharm Teach Learn. 2015 Mar-Apr;7(2):249-25

4. Peterson KJ, Pathak J. Scalable and High-Throughput Execution of Clinical Quality Measures from Electronic Health Records using MapReduce and the JBoss Drools Engine. AMIA Annual Symposium Proceedings. 2014:250-259

5. Jadhav A, Sheth A, Pathak J. Online Information Searching for Cardiovascular Diseases: A Quantitative Analysis of Mayo Clinic Search Query Logs. AMIA Annual Symposium Proceedings. 2014:151-160

6. Bobo WV, Pathak J, Kremers HM, Yawn BP, Brue SM, Stoppel CJ, Croarkin PE, St Sauver J, Frye MA, Rocca WA. An electronic health record driven algorithm to identify incident antidepressant medication users. J Am Med Inform Assoc. 2014 Oct;21(5):785–791. **PMID: 24780720 PMCID: PMC4147111**

7. Khaleghi M, Isseh IN, Bailey KR, Kullo IJ. Family history as a risk factor for peripheral arterial disease. Am J Cardiol. 2014 Sep 15;114(6):928–932. **PMID: 25107577 PMCID: PMC4206951**

8. Khaleghi M, Isseh IN, Jouni H, Sohn S, Bailey KR, Kullo IJ. Family history as a risk factor for carotid artery stenosis. Stroke. 2014 Aug;45(8):2252–2256. **PMID: 25005442**

9. Olson JE, Bielinski SJ, Ryu E, Winkler EM, Takahashi PY, Pathak J, Cerhan JR. Biobanks and personalized medicine. Clin Genet. 2014 Jul;86(1):50–55. **PMID: 24588254**

10. Bielinski SJ, Olson JE, Pathak J, Weinshilboum RM, Wang L, Lyke KJ, Ryu E, Targonski PV, Van Norstrand MD, Hathcock MA, Takahashi PY, McCormick JB, Johnson KJ, Maschke KJ, Rohrer Vitek CR, Ellingson MS, Wieben ED, Farrugia G, Morrisette JA, Kruckeberg KJ, Bruflat JK, Peterson LM, Blommel JH, Skierka JM, Ferber MJ, Black JL, Baudhuin LM, Klee EW, Ross JL, Veldhuizen TL, Schultz CG, Caraballo PJ, Freimuth RR, Chute CG, Kullo IJ. Preemptive genotyping for personalized medicine: design of the right drug, right dose, right time-using genomic data to individualize treatment protocol. Mayo Clin Proc. 2014 Jan;89(1):25–33. **PMID: 24388019 PMCID: PMC3932754**

11. Kullo IJ, Shameer K, Jouni H, Lesnick TG, Pathak J, Chute CG, de Andrade M. The ATXN2-SH2B3 locus is associated with peripheral arterial disease: an electronic medical record-based genome-wide association study. Front Genet. 2014;5:166. **PMID: 25009551 PMCID: PMC4070196**

12. Jadhav A, Andrews D, Fiksdal A, Kumbamu A, McCormick JB, Misitano A, Nelsen L, Ryu E, Sheth A, Wu S, Pathak J. Comparative analysis of online health queries originating from personal computers and smart devices on a consumer health information portal. J Med Internet Res. 2014;16(7):e160. **PMID: 25000537 PMCID: PMC4115262**

13. Bielinski SJ, Olson JE, Pathak J, Weinshilboum RM, Wang L, et al. Preemptive genotyping for personalized medicine: design of the right drug, right dose, right time-using genomic data to individualize treatment protocol. Mayo Clin Proc. 2014 Jan;89(1):25–33. **PMID: 24388019 PMCID:PMC3932754**

14. Pathak J, Kiefer RC, Chute CG. Mining Anti-coagulant Drug-Drug Interactions from Electronic Health Records Using Linked Data. In: Baker CJO, Butler G, Jurisica I, editors. Data Integration in the Life Sciences. Springer Berlin Heidelberg; 2013. p. 128–140.

15. Pathak J, Bailey KR, Beebe CE, Bethard S, Carrell DC, et al. Normalization and standardization of electronic health records for high-throughput phenotyping: the SHARPn consortium. J Am Med Inform Assoc. 2013 Nov 4; **PMID: 24190931 PMCID: PMC3861933**

16. Jouni H, Shameer K, Asmann YW, Hazin R, Andrade M de, Kullo IJ. Clinical Correlates of Autosomal Chromosomal Abnormalities in an Electronic Medical Record–Linked Genome-Wide Association Study A Case Series. Journal of Investigative Medicine High Impact Case Reports. 2013 Oct 1;1(4):2324709613508932.

17. Takahashi PY, Ryu E, Olson JE, Anderson KS, Hathcock MA, et al. Hospitalizations and emergency department use in Mayo Clinic Biobank participants within the employee and community health medical home. Mayo Clin Proc. 2013 Sep;88(9):963–969. **PMID: 24001488**

18. Olson JE, Ryu E, Johnson KJ, Koenig BA, Maschke KJ, et al. The Mayo Clinic Biobank: a building block for individualized medicine. Mayo Clin Proc. 2013 Sep;88(9):952–962. **PMID: 24001487**

19. Ludvigsson JF, Pathak J, Murphy S, Durski M, Kirsch PS, et al. Use of computerized algorithm to identify individuals in need of testing for celiac disease. J Am Med Inform Assoc. 2013 Aug 16; **PMID: 23956016 PMCID: PMC3861918**

20. Tao C, Li D, Shen F, Lian Z, Pathak J, Liu H, Chute CG. Phenotyping on EHR Data Using OWL and Semantic Web Technologies. Proceedings of the 2013 International Conference on Smart Health. Berlin, Heidelberg: Springer-Verlag; 2013. p. 31–32. 9. McGuire AL, Joffe S, Koenig BA, Biesecker BB, McCullough LB, et al. Ethics and Genomic Incidental Findings. Science. 2013 May 31;340(6136):1047–1048. **PMID: 23686340 PMCID: PMC3772710**

21. Wooten EC, Hebl VB, Wolf MJ, Greytak SR, Orr NM, et al. Formin homology 2 domain containing 3 variants associated with hypertrophic cardiomyopathy. Circ Cardiovasc Genet. 2013 Feb;6(1):10–18. **PMID: 23255317 PMCID: PMC3578062**

22. Zhu Q, Freimuth RR, Pathak J, Chute CG. Using Clinical Element Models for Pharmacogenomic Study Data Standardization. AMIA Summits Transl Sci Proc. 2013;2013:292–296. **PMID: 24303283 PMCID: PMC3845787**

23. Sohn S, Ye Z, Liu H, Chute CG, Kullo IJ. Identifying Abdominal Aortic Aneurysm Cases and Controls using Natural Language Processing of Radiology Reports. AMIA Summits Transl Sci Proc. 2013;2013:249–253. **PMID: 24303276 PMCID: PMC3845740**

24. Seedorff M, Peterson KJ, Nelsen LA, Cocos C, McCormick JB, et al. Incorporating expert terminology and disease risk factors into consumer health vocabularies. Pac Symp Biocomput. 2013;421–432. **PMID: 23424146 PMCID: PMC3587774**

25. Ridgeway JL, Han LC, Olson JE, Lackore KA, Koenig BA, et al. Potential bias in the bank: what distinguishes refusers, nonresponders and participants in a clinic-based biobank? Public Health Genomics. 2013;16(3):118–126. **PMID: 23595106 PMCID: PMC3821039**

26. Rea S, Bailey KR, Pathak J, Haug PJ. Bias in Recording of Body Mass Index Data in the Electronic Health Record. AMIA Summits Transl Sci Proc. 2013;2013:214–218. **PMID: 24303267 PMCID: PMC3845755**

27. Pathak J, Kiefer RC, Chute CG. Using linked data for mining drug-drug interactions in electronic health records. Stud Health Technol Inform. 2013;192:682–686. **PMID: 23920643 PMCID: PMC3909652 19**

28. McCormick, J. Whole Genome Sequencing. Lahey Clinic Medical Ethics Journal. 2013;20(1):1–2. 18. Li D, Simon G, Chute CG, Pathak J. Using Association Rule Mining for Phenotype Extraction from Electronic Health Records. AMIA Summits Transl Sci Proc. 2013;2013:142–146. **PMID: 24303254 PMCID: PMC3845788**

29. Kiefer RC, Freimuth RR, Chute CG, Pathak J. Mining Genotype-Phenotype Associations from Public Knowledge Sources via Semantic Web Querying. AMIA Summits Transl Sci Proc. 2013;2013:118–122. **PMID: 24303249; PMCID: PMC3845769**

30. Tao C, Jiang G, Oniki TA, Freimuth RR, Zhu Q, et al. A semantic-web oriented representation of the clinical element model for secondary use of electronic health records data. J Am Med Inform Assoc. 2012 Dec 25; **PMID: 23268487 PMCID: PMC3628064**

31. Zhu Q, Freimuth RR, Lian Z, Bauer S, Pathak J, et al. Harmonization and semantic annotation of data dictionaries from the Pharmacogenomics Research Network: A case study. J Biomed Inform. 2012 Nov 29; **PMID: 23201637 PMCID: PMC3606279**

32. Tao C, Pathak J, Solbrig HR, Wei W-Q, Chute CG. Terminology representation guidelines for biomedical ontologies in the semantic web notations. J Biomed Inform. 2012 Sep 28; **PMID: 23026232 PMCID: PMC3563768**

33. Clifford L, Singh A, Wilson GA, Toy P, Gajic O, et al. Electronic health record surveillance algorithms facilitate the detection of transfusion-related pulmonary complications. Transfusion. 2012 Aug 31; **PMID: 22934792**

34. Pathak J, Kiefer RC, Chute CG. Using Semantic Web Technologies for Cohort Identification from Electronic Health Records for Clinical Research. AMIA Summits Transl Sci Proc. 2012 Mar 19;2012:10–19. **PMID: 22779040 PMCID: PMC3392057**

35. Pathak J, Kiefer RC, Bielinski SJ, Chute CG. Mining the Human Phenome using Semantic Web Technologies: A Case Study for Type 2 Diabetes. AMIA Annu Symp Proc. 2012;2012:699–708. **PMID: 23304343 PMCID: PMC3540447**

36. Pathak J, Kiefer RC, Bielinski SJ, Chute CG. Applying semantic web technologies for phenome-wide scan using an electronic health record linked Biobank. J Biomed Semantics. 2012;3(1):10. **PMID: 23244446 PMCID: PMC3554594**

37. Pathak J, Kiefer R, Freimuth R, Chute C. Validation and discovery of genotype-phenotype associations in chronic diseases using linked data. Stud Health Technol Inform. 2012;180:549–553. **PMID: 22874251**

38. Li D, Endle CM, Murthy S, Stancl C, Suesse D, et al. Modeling and Executing Electronic Health Records Driven Phenotyping Algorithms using the NQF Quality Data Model and JBoss® Drools Engine. AMIA Annu Symp Proc. 2012;2012:532–541. **PMID: 23304325 PMCID: PMC3540464**

Mount Sinai

1. Huopaniemi I, Nadkarni G, Nadukuru R, Lotay V, Ellis S, et al. Disease progression subtype discovery from longitudinal EMR data with a majority of missing values and unknown initial time points. Proceedings Annual Symposium of the American Medical Informatics Association 2014; (In Press)

2. Bottinger E. Disease progression subtype discovery from longitudinal EMR data with a majority of missing values and unknown initial time points. (Submitted)

3. Fort D, Wilcox AB, Weng C. Could Patient Self-reported Health Data Complement EHR for Phenotyping? Proc of AMIA 2014 Fall Symp, Nov 15-19, 2014, Washington DC, 1738-1747

4. Overby CL, Kohane I, Kannry JL, Williams MS, Starren J, et al. Opportunities for genomic clinical decision support interventions. Genet Med. 2013 Oct;15(10):817-23. doi: 10.1038/gim.2013.128. Epub 2013 Sep 19. **PMID: 24051479 PMCID: PMC3858176**

5. Overby CL, Pathak J, Gottesman O, Haerian K, Perotte A, Murphy S, Bruce K, Johnson S, Talwalkar J, Shen Y, Ellis S, Kullo I, Chute C, Friedman C, Bottinger E, Hripcsak G, Weng C. A collaborative approach to developing an electronic health record phenotyping algorithm for drug-induced liver injury. J Am Med Inform Assoc. 2013 Dec;20(e2):e243-52. doi: 10.1136/amiajnl-2013-001930. Epub 2013 Jul 9. **PMID: 23837993 PMCID: PMC3861914**

6. Overby CL, Weng C, Haerian K, Perotte A, Friedman C, Hripcsak G. Evaluation considerations for EHR-based phenotyping algorithms: A case study for drug-induced liver injury. AMIA Summits Transl Sci Proc. 2013 Mar 18;2013:130–134. **PMID: 24303321 PMCID: PMC3814479**

Northwestern

1. Rasmussen LV. The electronic health record for translational research. J Cardiovasc Transl Res. 2014 Aug;7(6):607–614. PMID: 25070682 PMCID: PMC4147395

2. Gawron AJ, Kho A, Roberts A, Keswani RN, Muthalagu A, Thompson WK. Tu1332 Anatomic Adenoma Detection Rates Determined via Natural Language Processing: a Refined Quality Metric for Colonoscopy? Gastrointestinal Endoscopy. 2013 May 1;77(5):AB502–AB503.

3. Smith ME, Aufox S. Biobanking: The Melding of Research with Clinical Care. Curr Genet Med Rep. 2013 Jun 1;1(2):122–128. **PMID: 24159428 PMCID: PMC3804314**

4. Chisholm RL. The opportunities and challenges of implementing genomics-informed personalized medicine. Clin. Pharmacol. Ther. 2013 Aug;94(2):181–182. **PMID: 23872829**

5. Thompson WK, Rasmussen LV, Pacheco JA, Peissig PL, Denny JC, Kho AN, et al. An Evaluation of the NQF Quality Data Model for Representing Electronic Health Record Driven Phenotyping Algorithms. AMIA Annu Symp Proc. 2012;2012:911–920. **PMID: 23304366 PMCID: PMC3540514**

Vanderbilt

1. Dumitrescu L, Goodloe R, Farber-Eger E, Boston J, Crawford DC. The effects of electronic medical record phenotyping details on genetic association studies: HDL-C as a case study. (Submitted to Pacific Symposium on Biocumputing)

2. Muhammad N, Ayday E, Clayton EW, Glunter C, et al. Systematization of Knowledge: Privacy and Security in the Genomic Era. (Submitted to IEEE Security & Privacy Conference)

3. Xu H, Aldrich MC, Chen Q, Liu H, Peterson NB, et al. Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality. J Am Med Inform Assoc. 2014 Jul 22. pii: amiajnl-2014-002649. **PMID: 25053577**

4. Weeke P, Denny JC, Bastarache L, Shaffer C, Bowton E, et al. Examining rare and Low-Frequency Genetic Variants Previously Associated with Lone or Familial Forms of Atrial Fibrillation in an Electronic Medical Record System: A Cautionary Note. Circulation Cardiovascular Genetics. (Accepted)

5. The Myocardial Infarction Genetics Consortium Investigators. Inactivating Mutations in NPC1L1 and Protection from Coronary Heart Disease. N Engl J Med 2014; 371:2072-2082November 27, 2014DOI: 10.1056/NEJMoa1405386. **PMID: 25390462**

6. Kawai VK, Cunningham A, Vear SI, Van Driest SL, Oginni A, et al. Genotype and risk of major bleeding during warfarin treatment. The Pharmacogenomics Journal. (Accepted)

7. Karnes JH, Cronin RM, Rollin J, Teumer A, Pouplard C, et al. A genome-wide association study of heparin-induced thrombocytopenia using an electronic medical record. Thrombosis and Hemostasis (Accepted).

8. Wan Z, Vorobeychik Y, Xia W, Clayton EW, Kantarcioglu M, A Game Theoretic Framework for Analyzing Re-identification Risk. PLoS ONE (Accepted)

9. Stitziel NO, Won HH, Morrison AC, Peloso GM, Do R, et al. Inactivating Mutations in NPC1L1 and Protection from Coronary Heart Disease. New England Journal of Medicine. 2014 Nov 12;371(22):2072–2082. **PMID: 25390462**

10. Kawai VK, Cunningham A, Vear SI, Van Driest SL, Oginni A, et al. Genotype and risk of major bleeding during warfarin treatment. Pharmacogenomics. 2014 Dec;15(16):1973-83. doi: 10.2217/pgs.14.153. **PMID: 25521356 PMCID: PMC4304738**

11. Xie W, Kantarcioglu M, Bush W, Crawford D, Denny J, et al. SecureMA: Protection Participant Privacy in Genetic Association Meta-Analysis. Bioinformatics. Bioinformatics. 2014 Aug 21. **PMID: 25147357**

12. Ramsey LB, Johnson SG, Caudle KE, Haidar CE, Voora D, et al. The Clinical Pharmacogenetics Implementation Consortium Guideline for SLCO1B1 and Simvastatin-Induced Myopathy: 2014 Update. Clin Pharmacol Ther. 2014 Oct;96(4):423–428. **PMID: 24918167 PMCID: PMC4169720**

13. Li M, Carrell D, Aberdeen J, Hirschman L, Malin BA. De-identification of clinical narratives through writing complexity measures. Int J Med Inform. 2014 Oct;83(10):750–767. **PMID: 25106934**

14. Sattar AHMS, Li J, Liu J, Heatherly R, Malin B. A probabilistic approach to mitigate composition attacks on privacy in non-coordinated environments. Knowledge-Based Systems. 2014; **PMCID: PMC3914710**

15. Weeke P, Roden DM. Applied pharmacogenomics in cardiovascular medicine. Annu Rev Med. 2014;65:81–94. **PMID: 24111889**

16. Ogbogu U, Burningham S, Ollenberger A, Calder K, Du L, et al. Policy recommendations for addressing privacy challenges associated with cell-based research and interventions. BMC Med Ethics. 2014;15(1):7. **PMID: 24485220 PMCID: PMC3914710**

17. Wei W-Q, Cronin RM, Xu H, Lasko TA, Bastarache L, Denny JC. Development and evaluation of an ensemble resource linking medications to their indications. J Am Med Inform Assoc. 2013 Sep 1;20(5):954–961. **PMID: 23576672 PMCID: PMC3756263**

18. Rosenbloom ST, Madison JL, Brothers KB, Bowton EA, Clayton EW, Malin BA, et al. Ethical and practical challenges to studying patients who opt out of large-scale biorepository research. J Am Med Inform Assoc. 2013 Jul 25; **PMID: 23886923 PMCID: PMC3861935**

19. Brothers KB, Westbrook MJ, Wright MF, Myers JA, Morrison DR, Madison JL, et al. Patient awareness and approval for an opt-out genomic biorepository. Personalized Medicine. 2013 Jun;10(4):349–359. **PMID: 24416062 PMCID: PMC3882901**

20. Oetjens MT, Denny JC, Ritchie MD, Gillani NB, Richardson DM, Restrepo NA, et al. Assessment of a pharmacogenomic marker panel in a polypharmacy population identified from electronic medical records. Pharmacogenomics. 2013 May;14(7):735–744. **PMID: 23651022 PMCID: PMC3725600**

21. Altman RB, Clayton EW, Kohane IS, Malin BA, Roden DM. Data re-identification: societal safeguards. Science. 2013 Mar 1;339(6123):1032–1033. **PMID: 23449577 PMCID: PMC3740512**

22. Atreya RV, Smith JC, McCoy AB, Malin B, Miller RA. Reducing patient re-identification risk for laboratory results within research datasets. J Am Med Inform Assoc. 2013 Jan 1;20(1):95–101. **PMID: 22822040 PMCID: PMC3555327**

23. Xia W, Heatherly R, Ding X, Li J, Malin B. Efficient discovery of de-identification policy options through a risk-utility frontier. Proceedings of the third ACM conference on Data and application security and 21 privacy. New York, NY, USA: ACM; 2013. p. 59–70.

24. Heatherly RD, Loukides G, Denny JC, Haines JL, Roden DM, Malin BA. Enabling genomic-phenomic association discovery without sacrificing anonymity. PLoS ONE. 2013;8(2):e53875. **PMID: 23405076 PMCID: PMC3566194**

25. Van Driest S, Shi Y, Bowton E, Schildcrout J, Peterson J, et al. Clinically Actionable Genotypes Among 10,000 Patients With Preemptive Pharmacogenomic Testing. Clin Pharmacol Ther. 2013 Nov 19; **PMID: 24253661**

26. Westbrook MJ, Wright MF, Van Driest SL, McGregor TL, Denny JC, Zuvich RL, et al. Mapping the incidentalome: estimating incidental findings generated through clinical pharmacogenomics testing. Genet. Med. 2012 Nov 29; **PMID: 23196672 PMCID: PMC3648626**

27. Wilke RA, Ramsey LB, Johnson SG, Maxwell WD, McLeod HL, Voora D, et al. The clinical pharmacogenomics implementation consortium: CPIC guideline for SLCO1B1 and simvastatin-induced myopathy. Clin. Pharmacol. Ther. 2012 Jul;92(1):112–117. **PMID: 22617227 PMCID: PMC3384438**

28. A, Loukides G, Nergiz ME, Saygin Y, Malin B. Anonymization of Longitudinal Electronic Medical Records. IEEE Transactions on Information Technology in Biomedicine. 2012 May;16(3):413 –423. **PMID: 22287248 PMCID: PMC3779068**

29. Clayton EW. Sharing individual research results with biospecimen contributors: counterpoint. Cancer Epidemiol. Biomarkers Prev. 2012 Feb;21(2):260–261. **PMID: 22313940 PMCID: PMC3815582**

30. Schildcrout JS, Denny JC, Bowton E, Gregg W, Pulley JM, Basford MA, et al. Optimizing Drug Outcomes Through Pharmacogenetics: A Case for Preemptive Genotyping. Clinical Pharmacology & Therapeutics. 2012;92(2):235–242. **PMID: 22739144 PMCID: PMC3785311**

31. Pulley JM, Denny JC, Peterson JF, Bernard GR, Vnencak-Jones CL, Ramirez AH, et al. Operational Implementation of Prospective Genotyping for Personalized Medicine: The Design of the Vanderbilt PREDICT Project. Clinical Pharmacology & Therapeutics. 2012;92(1):87–95. **PMID: 22588608 PMCID: PMC358130**

eMERGE Workgroup Charters

Consent, Education, Regulation, and Consultation (CERC) Workgroup

Co-Chairs: Maureen Smith & Ingrid Holm

The eMERGE II consortium supports research by existing biorepositories with linked electronic health records to incorporate current genomic knowledge into clinical research and ongoing clinical care. In particular, this workgroup will:

- Share and compare methods for obtaining ethical consent in recruitment of participants based on genotype
- Evaluate recommendations from L. Beskow workshop on genotype driven research recruitment (in press)
- Assess the educational needs of patients and physicians in returning genomic research results
- Collect existing methods of risk communication
- Create a resource of the CLIA/CAP regulations for clinical investigators by linking with the relevant experts
- Establish liaisons with other new and ongoing projects, including the Return of Results Consortium, Biobank subcommittee of the CTSAs, and others as identified, to better address the issues associated with returning genomic research results through EHR/decision support programs
- Explore the role and impact of personal utility on returning result decisions

EHRI Workgroup

Co-Chairs: Justin Starren & Marc Williams

The EMR Integration workgroup will develop eMERGE II consensus and concepts for EMR integration of genomic information and delivery of clinical genomic decision support utilizing EMR.

1. Delineate common and distinct approaches and challenges for EMR integration of relevant genomic information for clinical pharmacogenomic, monogenic disorders, and common disease risk applications.
2. Develop concepts and tools for clinical genomic decision support implementations in EMR.
3. Establish and apply features and benchmarks for clinical genomic decision support implementation research, including data ascertainment and outcomes analysis.
4. Address challenges and approaches for utilization of whole genome/exome sequence-associated information in EMR/decision support.
    a) Several publications from the EHRI have outlined challenges and approaches
    b) Implementation was not pursued as a network wide activity. Some eMERGE sites have moved forward with next generation sequencing implementation with development of CDS.

5. Establish and maintain dialog with main EMR vendors.
6. Support the use and evaluation of CDS tools in eMERGE II clinical implementation projects

Genomics Workgroup

Co-Chairs: David Crosslin & Patrick Sleiman

- Workgroup activities for eMERGE II centered on the following topics:
    - Continued guidance for the eMERGE Genomics CC
        - Quality Control and data cleaning for GWAS and other high-density arrays
        - Imputation analysis/methods for GWAS
    - Genotype/phenotype inventory by study site
    - Genetic association study designs
        - including pharmacogenomics
    - Co-Chairs: David Crosslin & Gerard Tromp

Pediatrics Workgroup

Co-Chairs: Hakon Hakonarson & John Harley

The Pediatric Workgroup was formed to provide a forum to find solutions for the scientific, public policy, ethical, and legal issues confronting eMERGE that have a uniquely pediatric component.

Examples include the vagaries of human subject consent in pediatrics, the complexities of the return of results to pediatrics patients and their guardians, and the phenotypes that are different from those found at adult institutions, including pediatric-specific diseases, growth and developmental milestones. Also, coordinating phenotypes and data collection will constitute a special opportunity for this workgroup.

The Pediatric Workgroup will strive to minimize the duplication of the work being done by the other workgroups in eMERGE and endeavor to focus its attention on the pediatric component in instances where this will be helpful.

Phenotyping Workgroup

Co-Chairs: Josh Denny & Peggy Peissig

1) Coordinate and complete network phenotypes and support covariates for analysis

- Create, execute, validate, and share phenotype algorithms
- Coordinate and prioritize phenotype algorithms
- Develop best practices for phenotype development, sharing, validation, execution
- Facilitate privacy-preserving deposit and efficient reuse of data in dbGaP

- Advance the science of de-identification

2) Efficient, effective, transportable phenotyping methods, structure and standards

- Quantify portable components of algorithms and methods (in progress)
- Develop structured templates or frameworks for representation of phenotypes (in progress
- Share algorithms via open libraries that extend beyond eMERGE
- Promote beyond eMERGE – PGRN, SHARP, Beacon

3) Coordination with all eMERGE workgroups

4) Coordinate with other networks

Return of Results Workgroup

Co-Chairs: Gail Jarvik, & Iftikhar Kullo

1. Define an initial set of variants that are potentially useful in clinical practice for purposes such as assessment of genetic risk for complex disorders or selection or dosing of drugs. This initial set will focus on common disease risk variants and pharmacogenetic variants for which we expect to have data. We will assess the levels of evidence supporting these variants and consider the cost and benefit of incorporating them into patient care. To do this we will interact with the larger eMERGE II community and other NHGRI funded ROR initiatives.
2. Assess ways to address the dynamic nature of genetic risk, ie potential change in risk as additional susceptibility variants are identified.
3. Periodic review of topics of interest to the group to be conducted on the monthly teleconference calls.

eMERGE Network Steering Committee
December 4-5, 2014
Bethesda, MD

Attendance

CCHMC/BCH	Armand Antommaria
CCHMC/BCH	Ariel Chandler
CCHMC/BCH	Beth Cobb
CCHMC/BCH	John Harley
CCHMC/BCH	Ingrid Holm
CCHMC/BCH	Todd Lingren
CCHMC/BCH	Bahram Namjou
CCHMC/BCH	Yizhao Ni
CCHMC/BCH	Cindy Prows
CCHMC/BCH	Wendy Wolf
CHOP	Berta Castillo
CHOP	John Connolly
CHOP	Joseph Glessner
CHOP	Hakon Hakonarson
CHOP	Brendan Keating
CHOP	Frank Mentch
CHOP	Patrick Sleiman
CHOP	Lyam Vazquez
Geisinger	Kenneth Borthwick
Geisinger	David Carey
Geisinger	Helena Kuivaniemi
Geisinger	Joseph Leader
Geisinger	David Ledbetter
Geisinger/U. Maryland	Casey Overby
Geisinger	Gerard Tromp
Geisinger	Marc Williams
GH/UW	David Carrell
GH/UW	David Crosslin
GH/UW	Andrea Hartzler
GH/UW	Gail Jarvik
GH/UW	Brian Shirts
GH/UW	Susan Trinidad
Marsh/Essentia/PSU	Murray Brilliant
Marsh/Essentia/PSU	Molly Hall
Marsh/Essentia/PSU	Scott Hebbring
Marsh/Essentia/PSU	Terrie Kitchner
Marsh/Essentia/PSU	Peggy Peissig
Marsh/Essentia/PSU/CC	Marylyn Ritchie
Marsh/Essentia/PSU/CC	Shefali Setia
Marsh/Essentia/PSU/CC	John Wallace
Mayo	Pedro Caraballo
Mayo	Mariza de Andrade
Mayo	Robert Freimuth
Mayo	Iftikhar Kullo
Mayo	Jen McCormick
Mayo	Jyoti Pathak
Mt. Sinai	Noura Abul-Husn
Mt. Sinai	Erwin Bottinger
Mt. Sinai	Stephen Ellis
Mt. Sinai	Genevieve Galarneau
Mt. Sinai	Carol Horowitz
Mt. Sinai	Eimear Kenny
Mt. Sinai	Ana Meijia
Mt. Sinai	Girish Nadkarni
Mt. Sinai	Aniwaa Owusu Obeng
Mt.Sinai/Columbia	Chunhua Weng
Northwestern	Rex Chisholm
Northwestern	Geoff Hayes
Northwestern	Jennifer Pacheco
Northwestern	Laura Rasmussen-Torvik
Northwestern	Luke Rasmussen
Northwestern	Justin Starren
Northwestern	Maureen Smith
NHGRI	Steve Benowitz
NHGRI	Rongling Li
NHGRI	Teri Manolio
NHGRI	Jackie Odgis
NHGRI	Mike Pazin
NHGRI	Bob Wildin
NHGRI	Ken Wiley
Vanderbilt/Louisville/CC	Kyle Brothers
Vanderbilt	Ellen Clayton
Vanderbilt/CC	Josh Denny
Vanderbilt	Nanibaa' Garrison
Vanderbilt	Josh Peterson
Vanderbilt	Dan Roden
Vanderbilt/CC	Sarah Stallings
CC	Melissa Basford
CC	Adam Hardebeck
CC	Paul Harris
CC-Case Western	Jonathan Haines
CC/Vanderbilt	Bradley Malin
CC/Vanderbilt	Martha Shrubsole

External Scientific Panel

U. of Alabama	Eta Berner
UNC – Chapel Hill	Gerardo Heiss
Moffitt Cancer Center	Howard McLeod
U. of Pittsburgh	Lisa Parker

Network Invitees and Guests

Aurora Research Institute	Michael Michalkiewicz
Complete Genomics, Inc.	Raith Erickson
CIDR	Elizabeth Pugh
CIDR	Kim Doheny
CIDR	Jane Romm

The eMERGE Steering Committee & ESP Meeting was held on December 4-5, 2014 in Bethesda, MD. In order to ensure that the Network remains productive as we continue through our final year, please find highlights from the Steering Committee/ESP Meeting below.

Presentation slides are available here and are also linked within the meeting summary below.

Goals from the meeting include:

- Update achievements by focusing on scientific presentations
- Response to ESP recommendations
- Workgroup updates
- Network projects update
  - PGx
  - CERC Survey
- Products dissemination update
- External collaboration update

Day 1: Full Session (Opening Remarks, Science Presentations)

- Welcome, Opening Remarks, General Updates – *Rongling Li (NHGRI)*
  - The Steering Committee was encouraged to refine plans for the final 7 months of phase II, and to consider effective methods of completing outstanding workgroup issues in relation to Phase II goals.
  - The Steering Committee was also encouraged to refine their plans for disseminating network-wide lessons learned and best practices in eMERGE II to the greater scientific community.

ACTION ITEM: The Steering Committee was encouraged to refine plans for the final 7 months of phase II, and to consider effective methods of completing outstanding workgroup issues in relation to Phase II goals.

ACTION ITEM: The Steering Committee was also encouraged to refine their plans for disseminating network-wide lessons learned and best practices in eMERGE II to the greater scientific community.

- Integrating Clinical Genomics into the EHR: Using Interface Terminology (IMO) and Interoperable Standards (SNOMED, LOINC, FHIR) – *Jennifer Pacheco (Northwestern)*
  - Following the presentation, the group was asked to begin thinking about collaboration opportunities to improve clinical genomic result representation, thereby accelerating dissemination of those results.
  - Groups (genomics, phenotyping, EHRI) could begin to identify opportunities as Phase II wraps up, and real work to address this issue could start during eMERGE III.
- Appendicitis - *John Harley (CCHMC)*
- New Method to Link Gene Discovery to Genomic Medicine in EHR-linked Biobanks: Uncovering Surprisingly High Incidence of Steel Syndrome in Puerto Ricans – *Eimear Kenny (Mt. Sinai)*
- eMERGE PGx Plenary Session – *Dan Roden (VU), Josh Denny (VU) & Laura Rasmussen-Torvik (Northwestern)*
  - The PGx project was discussed in depth, including a review of timelines and progress made by the group. The group is currently on pace to complete the project by June.

- An update was provided on the network-wide variant paper (led by David Crosslin & Will Bush) and discussed an outline and timeline.
  - PGRNseq Platform paper is also out for review.
  - MACE and Clopidogrel algorithm will be circulated to the network in January.
  - The paper will use data from the October 2014 SPHINX update.
  - The goal is to submit paper to the American Journal of Human Genetics by Dec. 20, 2014.
- The dbGaP submission plan and timeline was reviewed.
  - NHGRI urged the group to identify what data to include in the upcoming dbGaP submission, and to ask PharmGKB for advice on additional data pieces to include.
  - Sites were also encouraged to start working with PGRN sites to identify what phenotypes they would like to see in SPHINX.
- SPHINX public site is undergoing modifications, and those improvements were reviewed.
- An update on CDS metrics data and collection was provided, as well as the proposed Infobutton project timeline. The project is on schedule to be completed by July.
- Site initiated analyses were reviewed
  - Bob Wildin (NHGRI) asked the group to think about how to reclassify non-pathogenic variants when excluding symptomatic ICD-9 codes.

ACTION ITEM: The workgroup was asked to consider what data to include in the upcoming dbGaP submission, and to work with PharmGKB on identifying other relevant data pieces.

ACTION ITEM: Sites were encouraged to work with PGRN to identify what data and phenotypes they would like to see added in SPHINX.

ACTION ITEM: Bob Wildin (NHGRI) asked the group to consider ways of reclassifying non-pathogenic variants when excluding symptomatic ICD-9 codes.

- Implementation of Clinical Decision Support for Pharmacogenomics – *Pedro Caraballo (Mayo)*
- Guest Presentation: Using ENCODE Data to Interpret Disease-associated Genetic Variation – *Mike Pazin (NHGRI)*
  - Dr. Pazin presented an overview of NGHRI's Encyclopedia of DNA Elements (ENCODE) as a resource for illuminating the role of genetic variation in human diseases. The discussion highlighted ENCODE's aspirational goals of cataloging all functional elements in the genome, as well as building maps that can be used to make predictions about genome function.
  - Rex recommended that the Phenotyping and Genomics workgroups come up with a catalog of small projects to collaborate with ENCODE on, and report back to the network with proposals on ways to expand the collaboration network-wide.
  - The revised affiliate membership document has been accepted by the network, and ENCODE will begin moving forward with their application for membership.
- Simulation of the Clinical and Economic Impact of Preemptive, Multiplexed Pharmacogenomic Testing – *Josh Peterson (VU)*

- Using a predictive model, the costs of preemptive genotyping is offset by improved outcomes related to CYP2C19-tailored antiplatelet therapy.

- Post Mortem Whole Genome Sequencing: A Genomic Autopsy – *Murray Brilliant (Marshfield)*

Day 2: Full Session (ESP Commentary, Science Presentations)

- Review of Progress of Prior ESP Recommendations - *Rex Chisholm*
- Optimal Display of Different Types of Genetic Information in the EHR: An eMERGE-CSER Collaboration – *Casey Overby (UMD)*
    - Following the overview, NHGRI and ESP members asked whether end users have been involved with this analysis yet. As of now, eMERGE has not focused on user perception of genetic data in the EHR, but more on where this data is being held/displayed in EHRs. Both the location and type of genetic data being displayed were identified as focus areas moving forward.
- Initial Analysis of Whole Exome Sequence Data from 10,000 Geisinger Patients: Implications & Opportunities – *David Carey (Geisinger)*
- Workgroup Timelines and Ongoing Projects
    - In addition to the presentations and discussions mentioned above, all eMERGE workgroups presented updates. Further details can be found in the next section.
- ESP Closing Comments
    - The ESP noted that eMERGE should attempt to publish (or at least document internally) the differences between sites/lessons learned during Phase II.
    - The panel also mentioned that in future grant phases, the network should work to identify external funding sources (industry collaborations, philanthropy, etc.) to increase the long-term sustainability of eMERGE research.

ACTION ITEM: The network should attempt to publish (or at least document internally) the differences between sites/lessons learned during Phase II.

ACTION ITEM: For Phase III, the network could identify external funding sources (industry collaborations, philanthropy, etc.) to increase the long-term sustainability of eMERGE research.

Workgroup Presentations

- **CERC**
    - Projects
        - Patient Perspectives on Broad Consent in Biobank Research in the eMERGE Network (Survey Project)
            - Current status: cognitive interviews and REDCap development complete, pilot survey in process, survey on schedule to be distributed in 2 waves (Wave 1: Feb-Mar, Wave 2: Apr-May)
        - Myresults.org (designed by John Connolly, CHOP)

- Collaborations
  - Joint CERC/ROR workgroup – ongoing joint meetings
  - CSER ELSI Consortium – ongoing joint meetings
  - Infobutton Project – contributing educational content
  - PGx Education Outcomes
- Publications
  - Pediatric Biobank Model Consent Language (Kyle Brothers) – *published*
  - Age of Majority and Consent (Kyle Brothers) – *in process*
  - Return of Research Results (Gail Jarvik) – *published*

ACTION ITEM: The ESP asked the workgroup to consider drafting at least a manuscript documenting challenges/lessons learned for CERC Survey.

- **EHRI**
  - Projects
    - Data collection related to genomic EMR CDS implementation challenges is underway
      - Data submitted from all sites, first draft manuscript available by end of December
    - Infobutton Project
      - Preliminary analysis of content has been completed and distributed to EHRI workgroup.
      - Content sharing infrastructure is currently being investigated.
    - IOM Action Collaborative
    - CDS Repository
      - This new project will collect each site's CDS rules and contain them on a site hosted by the CC.
  - Publications
    - Conceptual Model of Omic Data – *in process*
    - Practical Considerations in Genomics Decision Support – *in process*
    - EHRI-CSER White Paper – *in process*
- **Genomics**
  - Projects
    - eMERGE Imputation
      - All adult and pediatric samples are completed (55,289 total samples)
    - PGRNseq Multisample
      - Halfway to the enrollment target (5,249 currently enrolled)
    - Null Variant Analysis – *ongoing*
    - Structural Variation Analysis – *ongoing*
  - Collaborations
    - ENCODE Collaboration – ENCODE is currently completing Affiliate Membership Application
  - Publications
    - Frontiers in Genetics Special Issue – Genetics Research in Electronic Health Records Linked to DNA Biobanks – *published*
      - This issue currently has over 27,000 online views
- **Pediatrics**
  - Projects
    - PheWAS Analysis
    - Pediatric Algorithms
      - Atopic Dermatitis: primary and secondary validation complete; 1,695 cases, 8,072 controls.
      - ADHD: primary and secondary validation complete; more samples desired.

        - CNVs
            - Future directions: currently working on Phase I data, Phase II data will soon be used; review of significant genes underway. PennCNV is being developed for optimizing CNV calls.
        - PGRNseq
        - Rare Disease Discovery and Return of Results
- **PGx**
    - Projects
        - Network-wide Implementation: 2 sites complete.
        - UW Recalling Project: recalling 5,000 BAMs using most recent human genome reference.
        - Process Outcomes: currently assessing provider and patient education.
        - SPHINX: public and private site updates are ongoing.
        - Network phenotypes selected: Major adverse cardiac events while using Clopidogrel (adult sites), Methylphenidate and Tacrolimus (pediatric sites).
        - Lipids: aims to analyze sequence data modulations of lipid levels. Data dictionary available on PheKB.
    - Publications
        - Design and Anticipated Outcomes of the eMERGE PGx Project: A Multicenter Pilot for Preemptive Pharmacogenomics in Electronic Health Record Systems – *published*
        - Network Variant Paper – *in process*

ACTION ITEM: The ESP encouraged the PGx workgroup to begin planning future directions of SPHINX.

- **Phenotyping**
    - Projects
        - Phase II Phenotype Implementation
            - Current Status: 20 completed, 7 in progress, 3 extra completed, 6 extra in progress; 15 completed during Phase I.
        - Geocoding
        - Extension of PheKB to become data repository w/ validation tools
    - Publications
        - Desiderata for Computable Representations of Electronic Health Records-Driven Phenotype Algorithms – *in process*
        - Portable applications for implementing multi-site clinical NLP algorithms – *in process*
        - Modular phenotyping – *in process*
        - PheKB.org: An Online Collaboration Tool for Phenotype Algorithm Development and Sharing – *in process*
        - When Phenotypes Aren't Transportable (the story of RHTN) – *in process*
        - Codes do not always cut it: comparison of using coded data vs. more complex algorithms in defining accurate phenotypes – in process
        - Development and Validation of an Electronic Phenotyping Algorithm for Chronic Kidney Disease – *published* (AMIA 2014 Distinguished Paper Winner)
        - SOEMPI: A Secure Open Enterprise Master Patient Index Software Toolkit for Private Record Linkage – *published* (AMIA 2014 Distinguished Paper Nominee)
- **Return of Results**
    - Projects
        - Genomic Medicine Pilots are investigating genetic risk scores, SNPs, whole-genome sequencing, and preemptive pharmacogenetics.
    - Publications
        - Return of Results in the Genomic Medicine Projects of the eMERGE Network – *published*

- Return of Genomic Results to Research Participants: The Floor, the Ceiling, and the Choices in Between – *published* (Named in 11 Best AJHG Papers of 2012-14)

ACTION ITEM: The ESP asked the ROR workgroup to consider how to address variances of pharmacogenetic markers.

Summary of Action Items:

1. The Steering Committee was encouraged to refine plans for the final 7 months of phase II, and to consider effective methods of completing outstanding workgroup issues in relation to Phase II goals.
2. The Steering Committee was also encouraged to refine their plans for disseminating network-wide lessons learned and best practices in eMERGE II to the greater scientific community.
3. The PGx workgroup was asked to consider what data to include in the upcoming dbGaP submission, and to work with PharmGKB on identifying other relevant data pieces.
4. PGx sites were encouraged to work with PGRN to identify what data and phenotypes they would like to see added in SPHINX.
5. Bob Wildin (NHGRI) asked the PGx group to consider ways of reclassifying non-pathogenic variants when excluding symptomatic ICD-9 codes.
6. For Phase III, the ESP mentioned that the network could identify external funding sources (industry collaborations, philanthropy, etc.) to increase the long-term sustainability of eMERGE research.
7. The ESP asked the CERC workgroup to consider drafting at least a manuscript documenting challenges/lessons learned for the survey project.
8. The ESP encouraged the PGx workgroup to begin planning future directions of SPHINX.
9. The ESP asked the ROR workgroup to consider how to address variances of pharmacogenetic markers.

Next Meeting: March 30-31st, 2015; Bethesda, MD

Meeting Summary

eMERGE Network External Scientific Panel

Conference Call – 5/8/2014

Attendance

ESP Attendees:

University of Alabama-Birmingham: Eta Berner; **Univeristy of North Carolina**: Gerardo Heiss; **Moffitt Cancer Center:** Howard McLeod; **University of Pittsburg**h: Lisa Parker; **InterMountain Healthcare**: Stan Huff

Network Attendees:

CCHMC/BCH: John Harley, Ingrid Holm, Zak Kohane**; CHOP:** Hakon Hakonarson, John Connolly; **Geisinger:** Marc Williams, David Carey**; GroupHealth/UW:** Gail Jarvik, Eric Larson, David Crosslin, Andrea Hartzler, Aaron Scrol; **Marshfield/Essentia/PSU**: Murray Brilliant, Marylyn Ritchie, Terrie Kitchner; **Mayo**: Iftikhar Kullo; **Mount Sinai**: Erwin Bottinger; **Northwestern**: Rex Chisholm, Maureen Smith; **Vanderbilt**: Dan Roden, Josh Denny; **NHGRI**: Rongling Li, Teri Manolio, Jackie Odgis, Ken Wiley; **CC**: Melissa Basford, Paul Harris, Dana Crawford, Brandy Mapes, Lauren Melancon, Sarah Stallings

Decisions and discussion

Opening Remarks/Welcome (Rongling Li & Howard McLeod)

Rongling welcomed the group and thanked the ESP for their service to the Network. She continued to thank the eMERGE investigators for their hard work and progress. Rongling referenced the eMERGE Workshop held in January 2014 and reminded the group the goal of this workshop was for the upcoming presentation of concept clearance at the May Council meeting. Three questions were circulated to the ESP prior to the call, these three topics reflect specific areas eMERGE investigators would like ESP insight. Howard also thanked the ESP members and the Network sites for their serious response to previous ESP recommendations along with great progress.

Network Overview and Response to ESP recommendations (Rex Chisholm)

An overview of past ESP recommendations and the Network's response were presented. Highlights included publications, including a Network citation analysis and an update on page views and citations from the Genetics in Medicine Special Issue led by the EHRI workgroup. The Network continues to publish a large number of papers each year, and the Phase II publication trajectory is strong. eMERGE has been working to better utilize their dissemination of tools and knowledge to the broader scientific community, efforts have included: the updating and expansion of gwas.org, added functionality to PheKB and the Record Counter, and the creation of myresults.org and

SPHINX. The ESP asked the group if there were any industry standards that could be used to compare eMERGE's dissemination efforts to other NIH consortium. While no exact method was apparent the eMERGE CC believed this may be achievable through a strategic PubMed search along with other searchable databases. The National Center for Biomedical Ontology (NCBO) consortium has also considered a method for comparing use of tools. The Network was encouraged to reach out to NCBO to begin the comparison between these two groups.

Discussion of Network Workgroup Updates

EHR Implementation (Marc Williams)

Marc gave an overview of the current and proposed future EHRI efforts including:

- The Genetics in Medicine Special Issue has been well received by the community and visibility highlights were shared. Select topics from this special issue have also been presented at each AMIA meeting since the issue's release.
- An infobutton project is currently underway and has focused its efforts on utilizing infobutton standards to provide genomics education within the EMR.
- A joint CSER/EHRI paper is being led by the workgroup with a focus on identifying a location for genomic results in the EMR.
- All sites are in different phases of implementation for the eMERGE PGx project as well as local pharmacogenomic information. All groups have experienced some challenges and delays; these challenges has created fruitful discussions among group members.
- Proposed future directions for the remainder of eMERGE Phase II and the potential eMERGE Phase II were shared with the ESP. The group is working to prioritize these tasks and facilitate collaboration within eMERGE and with external groups.

The ESP noted that GH/UW was listed as on hold for implementation due to an institutional perspective that evidence for implementation was not strong enough. Gail and Eric spoke to this point expressing the rigorous guidelines used by their committee. Carbamazepine has already been implemented at their site and other drug gene pairs are under review with the hope that some will be approved. Each drug gene pair is review separately and approved on an individual basis. It was also mentioned that GH/UW is working closely with their implementation team; this team will be available for implementation once additional drug gene pairs are approved.

The ESP commented that even though the initial decision makers may have a high standards the clinicians may have an even higher standard as this will no doubt change their practice.

The future directions comment: *Bring clinical practice leaders into the fold to ensure CDS integrates in their specific clinical workflow* was highlighted. This may be a larger issue and the group may want to think about how to not only engage clinical practice leaders but also decision makers who

may or may not be scientists. This allows decision makers to have an opportunity to be fully engaged and gain a better understanding of CDS implementation for genomic medicine.

eMERGE investigators agreed that these conversations must be bidirectional so all groups can learn from each other. The genomic community would benefit from these discussions as it would allow for the group to anticipate challenges and information that would be helpful for decision making committees and in CDS.

The EHRI workgroup will work to expand this future direction bullet to extend past clinical workflow and include decision makers.

The ESP suggested that the EHRI workgroup to reflect and publish on the various challenges associated with CDS Implementation. This would aide other groups and also helps facilitate bidirectional conversations. This discussion in the literature would be of great value. The EHRI workgroup is currently collecting information from each site on their implementation challenges and barriers that will lead the group to a lessons learned and best practices type of paper.

Pediatrics (Hakon Hakonarson & John Harley)

Hakon outlined a new initiative by the Pediatric workgroup focused on heritability using genome-wide complex trait analysis (GCTA). This assessment will look both at social communication and developmental traits. A similar study has been done by CHOP for autism. The pediatric workgroup believes that GCTA can be adopted by all pediatric and adult sites and will be able to be parlayed into any phenotype selected by the Network by using EMR longitudinal data. The group will continue to create an action plan to move forward with heritability and plans to solicit phenotypes from the larger eMERGE Network.

The pediatric sites are actively working to create and validate their proposed Network phenotypes. Asthma is complete while others are in various stages of validation and development. The pediatric sites have specifically worked to create algorithms that not only perform well at the pediatric sites but can also be transported to adult populations. The pediatric sites are participating in as many adult led phenotypes as possible but some phenotypes are not relevant to their population.

Future directions of the group were presented and included projects that could be achieved during the final year of Phase II along with Phase III proposals. In the remaining time in Phase II the pediatric workgroup plans to use existing data to expand the current copy number variant effort. They also plan to impute drug-gene interactions from existing GWAS data. Hakon stated that 1/10 individuals have at least one risk allele and it can be of great importance to return these loss of function variants to patients. Phase III efforts may be expanded by utilizing a low cost custom chip to add to the number of pediatric samples currently available for analysis. These additional

samples would allow for additional pediatric activities within the Network along with additional pediatric publication opportunities.

The ESP had no comments at this time for the Pediatric workgroup.

General Discussion and Recommendations from the ESP

The group reviewed the three specific questions posed to the ESP.

- For the items that are listed to be accomplished by the workgroups during the current eMERGE funding cycle, what are the projects to which you would give the highest priority? Are there specific items or projects that require additional information or clarification?
    - The ESP noted multiple areas for focus:
        - Continued development of tools and processes
        - Collaborations between workgroups
        - Implementation and practice
    - Of these three the ESP felt implementation and practice was of the highest priority
- The eMERGE Network continues to disseminate tools and knowledge through publications, presentations, and online media. What other mechanisms can eMERGE use to further disseminate eMERGE knowledge and tools (specific examples)? Or is the Network casting a broad enough net at present?
    - The ESP felt the Network was going a good job disseminating tools and knowledge and offered no additional suggestions for improvement.
- Beyond current network collaborations, are there other consortia or individuals who eMERGE should look to for guidance or potential partnerships?
- Many sites have mentioned collaborations with outside groups such as PCORI and CTSA but the ESP could like to see these collaborations become more of a Network effort.
- eMERGE commented that some workgroups such as EHRI and CERC are doing this well but the Network will continue to pursue additional Network wide collaborations.

Meeting Summary

eMERGE Network – ESP Teleconference

Executive Session – 5/8/2014

ESP	Eta Berner (UAB)	**NHGRI**	Rongling Li
	Gerardo Heiss (UNC)		Teri Manolio
	Stan Huff (Intermountain Healthcare)		Jackie Odgis
	Howard McLEod (UNC, Chair)		Ken Wiley
	Lisa Parker (Pittsburgh)		

The External Scientific Panel (ESP) met with members of NHGRI staff in Executive Session after the ESP teleconference on May 8, 2014.

- The ESP was impressed by the scope of activity and pleased with productivity numbers.
- The ESP noted that the Coordinating Center seems motivated to find uses for their software.
- Visibility among the broader informatics community has increased tremendously; eMERGE has generated more interest by improving its dissemination of information about the network, its mission and progress.
- The group agreed that adult sites are working together well, but ESP raised concerns that the pediatric sites seem isolated. To encourage interactions, the ESP members recommended that the Pediatric WG develop some specific projects for pediatric sites and possibly adult sites to collaborate. Pediatric groups should also identify any special challenges they have faced in working together and in sharing data between other groups, pediatric and adult,to develop more efficient methods of analyzing data for joint projects. All sites operate under a network-wide data sharing and consent agreement, so this should not deter collaboration efforts.
- The ESP suggested that the wording of the future directions listed by each working group be made more consistent. The priorities of these tasks should reflect the views of the network as a whole.
- With regard to future directions on discovery vs. implementation, although these two entities are not entirely opposed, implementation was identified as a higher priority. eMERGE would be in a unique position to address implementation, particularly in using data as part of CDS. If the focus is on implementation, then it might be possible to determine where more work needs to be done on the level of discovery as well as determine which discoveries would be more feasible to implement and which would be more appealing to policymakers.

ESP Recommendations

1) The Network should continue to disseminate its products and best practices to the broader scientific community to increase visibility.
2) Pediatric sites need to work together to take advantage of this network, such as data sharing, and network project development.
3) The Pediatric Workgroup should identify and document the challenges preventing collaboration between pediatric sites.
4) The Pediatric Workgroup should design specific, multi-site projects on which to collaborate.
5) Future directions stated by all workgroups should have wording that is consistent and reflects the priorities of the network as a whole.
6) The Network should consider having more of a focus on implementation in the future.