eMERGE Network

Attendees of the eMERGE Steering Committee & ESP Meeting

December 4-5th, 2014; Bethesda, MD

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	Ellen Clayton
Vanderbilt-CC	Josh Denny
Vanderbilt	Nanibaa' Garrison
Vanderbilt-CC	Bradley Malin
Vanderbilt	Josh Peterson
Vanderbilt	Dan Roden
Vanderbilt-CC	Sarah Stallings
CC	Melissa Basford
CC	Adam Hardebeck
CC	Paul Harris
CC	Martha Shrubsole
CC-Case Western	Jonathan Haines
CC-U. Louisville	Kyle Brothers
CIDR	Elizabeth Pugh
CIDR	Kim Doheny
CIDR	Jane Romm

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

eMERGE Network

Summary of the eMERGE Steering Committee & ESP Meeting

December 4-5th, 2014; Bethesda, MD

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 (login required) and are also linked within the meeting summary below.

Goals for the meeting:

- Update achievements, focusing on scientific presentations
- Respond to ESP recommendations
- Workgroup updates
- Network projects updates
    - 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 plans for disseminating eMERGE II network-wide lessons learned and best practices to the greater scientific community.

ACTION ITEM: The Steering Committee will refine plans for the final 7 months of phase II, and will consider effective methods for completing outstanding work related to Phase II goals.

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

Scientific Presentations (selected from site submissions)

- Integrating Clinical Genomics into the EHR: Using Interface Terminologies (Intelligent Medical Objects, Inc.) 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)*
    - The Pediatric workgroup is focusing on an appendicitis phenotype, due to its applicability across both adult and pediatric sites and its age-dependent heritability. No GWAS has been completed for the condition, so there's much work to be done.
    - 272 cases with DNA samples are currently being genotyped at CCHMC.
    - The presentation provided an overview of three scenarios running the appendicitis algorithm.

- 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)*
    - This method that Mt. Sinai has been applying and developing is called Identity by Descent Mapping, which borrows ideas from linkage based pedigree analysis in a population-based way.
    - The Hispanic/Latino population was selected for this study because it is the most genetically diverse population available. A high amount of relatedness has also been identified in this population. Signals of abundant IBD sharing persists amongst this group, which opens up an opportunity to apply mapping approaches to this population.
    - Height was the selected trait for this study, and a molecular diagnosis using Pubmed data was performed to discover the high occurrence of (perceived) Steel Syndrome.
    - This method was identified as a good opportunity to ascertain the variability and penetrance of the Steel Syndrome phenotype.
- Implementation of Clinical Decision Support for Pharmacogenomics – *Pedro Caraballo (Mayo)*
    - The presentation provided an overview of Mayo's CDS pharmacogenomic CDS model, including obstacles and challenges.
    - Major challenges were identified:
        - Informatics/CDS professionals need to be more involved with determining clinical/PGx guidelines.
        - CDS standards are necessary.
        - EMR functionality needs to be improved to optimize CDS platforms.
        - CDS implementation/development is too resource-intensive in its current form. Better methods need to be identified to improve long-term sustainability.
- 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.
    - It was recommended that the Phenotyping and Genomics workgroups come up with a catalog of small projects in collaboration with ENCODE, 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 and cost-benefit analysis, the costs of preemptive genotyping are shown to be partially offset by improved outcomes related to more effective CYP2C19-tailored antiplatelet therapy.
- Post Mortem Whole Genome Sequencing: A Genomic Autopsy – *Murray Brilliant (Marshfield)*
    - Whole-genome sequencing was performed on 300 deceased subjects in the Personalized Medicine Research Project with complete health records. A number of functional variants were found that could have potentially been returnable, or had a clinical impact, prior to the subjects' deaths. An example subject was screened after being diagnosed with breast and ovarian cancers, and was found to have BRCA. The presenter suggested that preventative genetic screening could have saved resources in this situation, while some steering committee members dissented.

eMERGE PGx Network Project – Plenary Session Update – *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 the outline and timeline were reviewed.

- PGRNseq Platform paper is also out for review.
- MACE and Clopidogrel algorithm will be circulated to the network in January.
- Data from the October 2014 SPHINX update will be used in the paper.
- 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.
  - Members 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 PGx workgroup to think about curating variant classifications and keeping track of updates in the variant effect knowledge base as evidence for pathogenicity changes. For example, a variant classified as non-pathogenic for a specific phenotype (arrhythmia, for example), might be pathogenic in another phenotype context. In addition, as new data arrive and more genomic data are collected, evidence for variant effect will change.

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: The PGx workgroup should think about curating variant classifications and keeping track of updates in the variant effect knowledge base as evidence for pathogenicity changes. For example, a variant classified as non-pathogenic for a specific phenotype (arrhythmia, for example), might be pathogenic in another phenotype context. In addition, as new data arrive and more genomic data are collected, evidence for variant effect will change.

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

- Review of Progress on Previous 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)*
  - Geisinger has collaborated with Regeneron with the goal of whole-exome sequencing 100,000 patients over the next 5 years. The project has currently sequenced >15,000 samples at Geisinger. The long-term goal is to have >500,000 patients enrolled.
  - The project's overarching goal is to create a translational genomics pipeline.
  - Early data for GWAS of quantitative lab values for lipids was demonstrated.
  - Geisinger currently has several thousand pediatric samples, and has plans to sequence those in the future.
  - Geisinger is free to use the data for non-commercial purposes. Data is considered "pre-competitive." The organization is also working to identify the best way to create mutually-beneficial data sharing methods.
- 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 the CERC Survey group 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 CERC Survey group 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 was presented in its final format. The site has had 1,900 visitors to this point. Additional content is being developed, including a platform for Infobutton support, and the group is also working to incorporate search engine optimization to increase web traffic.
  - Collaborations
    - Joint CERC/ROR workgroup – joint meetings covering common issues related to implementation projects, CERC survey, and other ongoing projects. The groups meet once every 2 months.
    - Infobutton Project – contributing educational content in relation to Infobutton integration onto Myresults.org.
    - PGx Education Outcomes – the group is consulting on educational metrics for PGx.
  - 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 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 workgroup.
      - Content sharing infrastructure is currently being investigated.
      - The project is on track to develop an Infobutton platform to be used on Myresults.org.
    - CDS Repository
      - Proposed a new project to 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 – 55 null variants have been identified for PheWAS, and the project is on track to be completed by the end of Phase II.
    - Structural Variation Analysis – *ongoing*
- Collaborations
    - ENCODE Collaboration – NHGRI's ENCODE consortium is currently completing an Affiliate Membership Application to become a collaborator with eMERGE.
- 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

- **<u>Pediatrics</u>**
    - Projects
        - A PheWAS Analysis has been completed at all three pediatric sites.
        - Pediatric Algorithms
            - Atopic Dermatitis: 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 optimized for CNV calls.

- **<u>PGx</u>**
    - 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. New data will be uploaded to SPHINX in Feb. 2015.
        - 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.* This paper is targeting a high-impact journal, and will serve as the foundation for multiple eMERGE publications moving forward.

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

- **<u>Phenotyping</u>**
    - Projects
        - Phase II Phenotype Implementation
            - Current Status: 20 phenotypes have been completed, 7 are in progress. 3 extra phenotypes have been completed, and 6 extra are in progress.
            - There were 15 completed during Phase I.
        - The group has begun looking at geo-mapping for eMERGE PGx samples, and will continue to expand that during the remainder of Phase II.
        - Extension of PheKB to become data repository w/ validation tools
        - The workgroup is expanding PheKB to become a data repository with data validation tools. A secure data transfer method for large files has also recently been created in PheKB.
    - 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 and the American Society for Clinical Pharmacology & Therapeutics (ASCPT) to identify what data and phenotypes they would like to see added in SPHINX.
5. The PGx workgroup should think about curating variant classifications and keeping track of updates in the variant effect knowledge base as evidence for pathogenicity changes. For example, a variant classified as non-pathogenic for a specific phenotype (arrhythmia, for example), might be pathogenic in another phenotype context. In addition, as new data arrive and more genomic data are collected, evidence for variant effect will change.
6. For Phase III, it was mentioned that the Network could explore identifying external funding sources (industry collaborations, philanthropy, etc.) to increase the long-term sustainability of eMERGE research.
7. The CERC and ROR workgroups should consider drafting a manuscript documenting challenges/lessons learned/IRB approval for the survey project.
8. The PGx workgroup was encouraged to begin planning future directions of SPHINX.
9. The ROR workgroup should consider how to address variances of pharmacogenetic markers.
10. The EHRI workgroup in conjunction with the CC will explore creation of an online resource or repository for clinical decision support related to genomic medicine.
11. The Phenotyping workgroup should work to capture downstream products of PheKB use and connect to PhenX to benefit from context of prospective data capture (PhenX) as compared with existing data re-use.
12. As a method of gaining national recognition for the eMERGE informatics group, a connection could be made with the Healthcare Information Management Systems Society (HIMSS).
13. Sites should be encouraged to continue to work on key projects, especially the eMERGE PGx, even if they might not be funded in eMERGE III.
14. In future presentations, the CERC Survey Group should focus on the scientific content, design and implementation of the survey.
15. If warranted by the results of the CERC Survey, consideration should be given in Phase III to deploying it again in Spanish and potentially other languages, particularly Asian languages as recruitment seems to be very low in these populations.
16. The Network should generate and conduct analysis on consortium-level data on clinical economic impact, and potentially engage government agencies (e.g., HRSA, AHRQ) on economics-based projects.
17. The Network should identify groups that could pose barriers to implementation (e.g., hospital administrators, hospital IT personnel) and invite them to attend SC meetings to see the Network's successes and help them to better understand the benefit and feasibility of implementation.

18. The Network should continue to engage with stakeholders and EHR vendors.
19. The Network needs to fully use the existing genotype and phenotype data for genomic discovery research.
20. The Network should explore conducting a content analysis study to evaluate the use of eMERGE web resources and assess utility for users at different sites.
21. The Network should focus on its inter-site efforts, among and between the adult and pediatric sites.

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