**eMERGE Network Steering Committee Meeting**

September 17-18, 2015

Bethesda, MD

Attendance

Baylor College of Medicine Matthew Bainbridge

Baylor College of Medicine Richard Gibbs

Boston Children’s Hospital Ingrid Holm

CCHMC Armand Antommaria

CCHMC Beth Cobb

CCHMC John Harley

CCHMC Todd Lingren

CCHMC Bahram Namjou

CCHMC Cassandra Perry

CCHMC Cindy Prows

CHOP Berta Castillo

CHOP John Connolly

CHOP Hakon Hakonarson

CHOP/UPenn Brendan Keating

CHOP Michael March

CHOP Frank Mentch

CHOP Patrick Sleiman

Columbia Wendy Chung

Columbia George Hripcsak

Columbia Ali Gharavi

Columbia Ning (Sunny) Shang

Columbia Chunhua Weng

Geisinger Lindsay Bailey

Geisinger David Carey

Geisinger Dan Davis

Geisinger Alanna Kulchak Rahm

Geisinger Joseph Leader

Geisinger Raghu Metpally

Geisinger/U of Maryland Casey Overby

Geisinger/PSU Sarah Pendergrass

Geisinger/PSU Marylyn Ritchie

Geisinger Diane Smelser

Geisinger Susan Snyder

Geisinger Janet Williams

Geisinger Marc Williams

GH/UW David Carrell

GH/UW David Crosslin

GH/UW Adam Gordon

GH/UW Andrea Hartzler

GH/UW Eric Larson

GH/UW Dave Veenstra

Harvard Sandy Aronson

Harvard Emma Davenport

Harvard Robert Green

Harvard Maggie Helm

Harvard Elizabeth Karlson

Harvard Sekar Kathiresan

Harvard Shawn Murphy

Harvard Jordan Smoller

Harvard Scott Weiss

Mayo Mariza de Andrade

Mayo Adelaide Arruda-Olson

Mayo Suzette Bielinski

Mayo/JHU Chris Chute

Mayo Robert Freimuth

Mayo Iftikhar Kullo

Mayo Hongfang Liu

Mayo Jennifer McCormick

Mayo/Cornell Jyoti Pathak

Mayo Richard Sharp

Mayo Stephen Thibodeau

Mt. Sinai Aniwaa Owusu Obeng

Northwestern Rex Chisholm

Northwestern Geoff Hayes

Northwestern Laura Rasmussen-Torvik

Northwestern Megan Roy-Puckelwartz

Northwestern Maureen Smith

Northwestern Justin Starren

Partners/Broad/Harvard Birgit Funke

Partners/Broad Stacey Gabriel

Partners/Broad Niall Lennon

Partners/Broad/Harvard Heidi Rehm

NHGRI Jyoti Gupta

NHGRI Dave Kaufman

NHGRI Rongling Li

NHGRI Teri Manolio

NHGRI Ken Wiley

NHGRI Kira Wong

Vanderbilt Ellen Clayton

Vanderbilt Nancy Cox

Vanderbilt/CC Josh Denny

Vanderbilt Todd Edwards

Vanderbilt Tracy McGregor

Vanderbilt Josh Peterson

Vanderbilt Dan Roden

Vanderbilt Martha Shrubsole

Vanderbilt/CC Sarah Stallings

Vanderbilt Georgia Wiesner

CC Melissa Basford

CC Adam Hardebeck

CC Paul Harris

CC Jacqueline Kirby

CC/Vanderbilt Bradley Malin

|  |
| --- |
| **eMERGE Network*****Summary of the eMERGE Steering Committee***September 17-18th, 2015; Bethesda, MD |
| The initial Phase III eMERGE Steering Committee Meeting was held on September 17-18th, 2015 in Bethesda, MD. In order to ensure that the Network starts on a productive note as we embark on our initial year, please find highlights from the Steering Committee Meeting below. Presentation slides are [available here](https://emerge.mc.vanderbilt.edu/1694-2/) (login required).*Goals for the meeting:** Integrate new sites into the network and synergize site efforts
* Discuss sequencing proposal and pipeline
* Discuss clinic annotation and report of actionable genetic variants
* Discuss network data management, data sharing and genomic discovery
* Develop workgroup charters, milestones and timeline
 |
| **Day 1: Full-day Session** |
| Welcome, Opening Remarks, General Updates – *Rongling Li** This is the first eMERGE III meeting, and also the 25th Steering Committee Meeting for the eMERGE Network.
* Jyoti Gupta, from the NHGRI sequencing center, is joining the NHGRI eMERGE team.
* Long-time NHGRI staff member, Dr. Elizabeth Thompson, 64 years old, passed away peacefully on July 14th at home. Elizabeth made a significant contribution to the NHGRI ethical legal and social implications research program.
* Brief summary of accomplishments from previous eMERGE phases.
	+ 531 total publications, and 198 network publications through July 2015
	+ eMERGE has created tools and software, including ELSI tools, like the model consent language, educational tools, genomic implementation tools, and phenotyping and genotyping discovery tools.
	+ eMERGE is growing:
		- eMERGE I had 5 sites and 18,500 individuals with GWAS data
		- eMERGE II had 53,000 total participants with GWAS data
		- eMERGE III begins with 71,600 total, including eMERGE 1 and 2 data imputed against 1000 genomes with > 55,000 genotyped and phenotyped individuals represented. That data has been deposited to dbGaP, thus contributing to the scientific community.

Network Overview – *Rex Chisholm** Rex will continue as PI Chair and Leadership WG Liaison.
* The network’s gene panel has been compiled ahead of schedule thanks to hard up-front work.
* The goal is to establish the 6 workgroups at this meeting and allowing time for WGs to work together on their charters and missions and report back.
* One of the exciting transitions in eMERGE III is moving from SNP level data to sequence level data. A great portion of this discussion focused on sequencing at both centers and how data will be returned to sites.

Sequencing Process and Sample Calling- *Richard Gibbs & Niall Lennon** Sequencing: Samples collected under CLIA conditions with annotated clinical records, will be transferred to the sequencing centers.
	+ Panel approach is more affordable and manageable at scale of 25,000.
	+ Baylor using Nimblegen capture and Harvard using Illumina – this slight difference will be studied and will be an output of this project helpful to the field.
	+ Harmonize 2 CSGs’ designs so processes return same .bed files – will do concordance analyses between 2 sites – the goal is to make any local differences invisible to the Network.
	+ Coverage depth limits target extent: 200-250X means only about 500kB.
	+ It will be 16-21 weeks before the assay is ready (won’t be ready for samples until April 2016).
	+ DNA sample size = 1ug.
	+ A clinical PM at each sequencing center will work with sites for sample transfer.
	+ Primary analysis will be SNVs, Indels, and structural variants.
		- Allelic coverage statistics will be available through DNA Nexus.
		- Any somatic variant analysis will be available where coverage is high enough.
* Variant interpretation/variant calling processes at the two sequencing centers will produce Patient-centered clinical CLIA reports which will flow through GeneInsight to the individual sites.
* Secondary, non-actionable findings will follow a different workflow and those more raw data will be collected in the Data Commons, which will contain functionality for investigators to do their own analyses. That will interact with other data storage sites like SPHINX.

Results Reporting- *Birgit Funke & Heidi Rehm** Clinical interpretation involves assessing the strength of gene-disease relationship, assessing the evidence linking the variant to disease, and interpreting the variant in the context of the patient’s phenotype to answer three questions: does the variant affect gene function; does this cause disease; does it fully explain the tested patients disease?
* Sites will access patient reports and interpreted variants from GeneInsight. They will access raw data and vcf files through Data Commons, and they will access discovery data through the CC.
* The CSGs will submit clinically relevant variants to ClinVAR.
* GeneInsight has two major components:
	+ GeneInsight lab (a case manager, all stored reports will be here, has a search program, all sites and staff will have logins to the system, patient identifying information will be de-identified).
	+ GeneInsight Clinic (9 systems for each of the 9 sites, only individual sites’ reports will be in it, patient identifying information will be preserved, alerts and clinical decision support functionality is available).
* GeneInsight has a knowledgebase which will be accessible and flow into GeneInsight Clinical so reports and alerts will be updated. GeneInsight knowledgebase is part of VariantWire, linking clinically validated, de-identified disease, gene, and variant data from all sites using the system. The hope is that the steering committee of VariantWire will approve eMERGE access.
* CSGs are developing the return of results process to be curated as variants are reinterpreted.
* Variant calls of genes on the eMERGE panel are being submitted to ClinVAR and any immediate discrepancies will be resolved. The Clinical Annotation and/or Return of Results workgroup can also be consulted on a case-by-case basis.
* An instance of GeneInsight will be provided to each site with an administrative login. In order for it to be utilized by a healthcare system, it would really need to be integrated with its EHR. The EHR Integration group could investigate current process flows and what would be needed to expand it. An alternative is that GeneInsight can put the information directly into the EHR.
* Close interaction with ClinGen will take place to engage subject matter experts in variant interpretation.

**ACTION ITEM: The CSGs will pre-curate common actionable variants.****ACTION ITEM: The CSGs will harmonize clinical variant classification between themselves.****ACTION ITEM: The CSGs will compare clinical report formats for harmonization.**Data Storage and Access –*David Crosslin** Raw data (bams, vcfs, FASTQs, and Recipe data) will be uploaded to Data Commons by sequencing centers, and annotated discovery analysis sets will be uploaded there by the CC for site access. This system has been used in the past, and was found to be a great tool for community engagement, it promoted discovery and was really about people, not computers.
* Imputed array data, multisample calling data from the PGx platform, will be pushed to Data Commons for site access. SPHINX will be re-evaluated with an eye to data asset assessment, future planning, use case exploration, privacy/policy specifications, and legal/technical safeguard descriptions.
* eMERGE will interface with the public through SPHINX, dbGAP, and possibly through other collaborations.

**ACTION ITEM: The CC will create a matrix of phenotype by participant and association results by phenotype for internetwork collaboration.**Genomic Data Management and Genomic Discovery Discussion- *Richard Gibbs, Niall Lennon, Birgit Funke, Heidi Rehm, & David Crosslin** Both array and sequence data will be requested for Pre III data.
* Cohort enrichment for specific disease indications (how many healthy participants vs how many participants have disease) was discussed.
* Default is that CSGs will only return pathogenic/likely pathogenic results. Including interpretation in light of indication would require discussion. This can be discussed on a site by site basis if interested. Having groups take turns presenting cases in a workgroup setting will also help inform processes.
* Documenting lessons learned is an important deliverable for the Network. There will not be one sequencing site for the world. The more we have projects that help us understand similarities and differences in processes, and learning from that, will add significantly from a scientific and implementation perspective. The different platforms being used by the two CSGs is point of least variability in the system compared to the rest of the interpretation pipeline.
* Getting different interpretations and advice from different sequencing providers is an issue, so modeling the real world and coming up with strategies for harmonization that can be broadly used is important. This is especially the case where the patient has had genetic testing before.
* Need to define context and how much ascertainment matters in the experiment. We will have biases due to recruitment, so recruitment strategy knowledge will be used to identify and correct for the biases we know. Recruitment will be driven by phenotype, and who to choose for sequencing (random or targeted sample) is still being determined by some sites. Thus adjustments may be possible.
* The ability to detect copy number variation is incomplete, but becoming robust. We are structured to further evolve the tools. Due diligence will be done to ensure the best coverage possible. Coverage reports will be readily available.
* Defining what to tell a patient when a mutation is discovered is the major challenge in implementation science, and eMERGE is poised to add to understanding in that area.

CERC Survey Update- *Maureen Smith & Ingrid Holm** The survey is now complete and the data are being analyzed. An overview of the survey development process was provided: survey design was aided by a systematic literature review and cognitive interviews; an oversampling strategy increased the data from generally unrepresented groups; the overall survey response rate was 15.8%.
* Preliminary analysis of respondent’s willingness to participate in biobank research showed little difference across the three broad consent scenario options randomized across the entire surveyed population. Factors affecting willingness to participate seemed to be affected only by race, education, and income. Site variables that might influence participants response was also investigated, and sites were surveyed to determine ascertainment strategies.
* One paper is being published, several others are in progress. There are also several posters and abstracts that have been presented and will be presented in the future. Next steps include analysis and manuscript development.
* The group discussed the Common Rule Notice of Proposed Rule Making (NPRM). The notice proposes to redefine human subjects to include all biospecimens regardless of identifiers, which is a major change. A major component of the change is a 15 element consent document that must be used by anyone wanting to use biospecimens. A brief consent form was proposed in the ANPRM. It must be written (rather than oral) for biospecimens. HHS will write a safe harbor template. There is the potential to waive consent, but very limited, and not applicable to academic researchers. In return, IRB oversight will be limited: they will only make sure consent process is ok, unless results are going to be returned. Individuals must be re-consented after 10 years or the samples and data must be destroyed. Deadline for expressing opinions is 12/7/15. The CERC group is preparing to respond/comment.

eMERGE Publication Policy Discussion – *Paul Harris** The current policy is designed to optimize transparency, inclusiveness, and simplicity for the Network.
* After some discussion from members about expanding the number of participating sites required for a paper to be considered a “Network publication”, the group agreed to keep the policy the same as it was during Phase II. Therefore, **2 or more sites participation** on a paper will be considered a “Network publication.”
 |
| **Day 2: Half-day Session** |
| Precision Medicine Initiative – Cohort Program Presentation – *Josh Denny** FY2016 Proposed Support for the PMI = $215 Million ($200 M for the NIH, $10M fir the FDA, and $5M for the Office of the National Coordinator for Health Information Technology.
* PMI Cohort will consist of 1 million newly recruited subjects and volunteers. Emphasis has been placed on broadly representing US population: all ages, all US regions, all stages of disease and variety of indications; special considerations for recruiting children, decisionally impaired and incarcerated people.
* All in cohort will have an initial health evaluation.
* Governance is still being worked out, but looking for a more simplified model of data access than dbGaP.
* Interactive participation model, with participant involvement at all levels, plus return of results and data access for participants.
* Single IRB; Broad consent sought.
* Operational structure will include participant health provider organization sites and a coordinating center for managing data flow between HPOs, participants, and researchers.
* Possible data sources include EHR data, biospecimen, self-report measures, baseline exam, and mHealth data.
* A core minimal data set will be defined, so that queries can be run on core data.
* The PMI group is trying to take lessons learned from the National Children’s Study.
* Method for determining site based application – Committee encouraged diverse application of participants. RFAs will likely be the method for applying.
* Each involved site would have deep EMR data provided for their participants and they can also perform site-based studies.

Return of Results / ELSI Workgroup Report *– Ingrid Holm & Iftikhar Kullo** The co-chairs presented the ROR/ELSI mission statement and charter. Overlaps with EHRI, Clinical Annotations, and Outcomes WG missions were noted, and a joint monthly call with the Clinical Annotation WG proposed
* Initial Plans:
	+ Develop and publish Network ROR standards within the first 6 months (March 2016).
	+ Develop jointly with the Clinical Annotation WG and ClinGen (several members overlap with that Network) the process and criteria for determining variant actionability for return.
	+ Survey sites for their ROR approach, including overall process, approach to consenting, patient/physician education plans, etc.

**ACTION ITEM: The eMERGE ROR/ELSI Workgroup will be invited to CSER ROR monthly calls.**Clinical Annotation Workgroup Report – *Heidi Rehm** This workgroup defined their role as ‘focusing on activities that build consistency of approaches to the gene and variant interpretation across the eMERGE sequencing centers and study sites, as well as support contribution to public knowledge bases.’
* The group will apply the ClinGen approach to gene-disease validity assessment to all genes on the panel. The Gene-Disease Validity classification was reviewed as well.
* A collaboration with the ROR/ELSI WG was proposed so that the groups can gather feedback and develop consensus on standard language used in clinical reports from the CSGs.

PGx Update – *Laura Rasmussen-Torvik** Six of 10 sites have returned all results from the project. Barriers for returning results were identified by the WG, and included physician pushback, manual data entry, and operational issues related to timely CLIA validation.
* The 5 PGx phenotypes are in varying stages of completion. MACE/Clopidogrel and Lipids have completed site validation, and are undergoing analysis. Intractable Epilepsy is almost finished with secondary validation, and Methylphenidate is still in algorithm development as additional patients were added to the project at a late stage.
* Moving forward into Phase III, the PGx workgroup will be folded into the new Outcomes WG as a standing agenda item. Former PGx workgroup members were encouraged to join the Outcomes WG.

EHRI Workgroup Report – *Sandy Aronson** Specific task areas for this group are:
	+ Engineering: Establish, document and seek to continuously improve process flows for delivery of eMERGE reports and data.
		- Significant engagement from all sites was stressed as being a requirement for this task area.
	+ Science: Experiment with innovative approaches that go beyond core requirements and evaluate their effectiveness.
	+ Community: Liaise with other groups, engage in collaborative projects, and disseminate learning and best practices.
* The group wants to serve as a resource for the Network to help determine feasibility with current EHR infrastructure.
* Sandy Aronson is a developer of GeneInsight, and will be very helpful in understanding GeneInsight to specific EHRI workflow questions
* The workgroup will have process flow documentation and a list of strategic projects by the Jan. 2016 SC Meeting.

Columbia Site Presentation – *Chunhua Weng, Ali Gharavi, & George Hripcsak** Columbia provided an overview of their site, which is new to eMERGE Phase III. Their program, called GENIE (Genomic Integration with EHR), will work to fulfill the following site aims:
	+ Advance next-generation phenotyping
	+ Perform genetic association studies of rare variants
	+ Develop practical, scalable learning mechanisms for returning results
	+ Provide genomic decision support
	+ Disseminate data, tools, algorithms, and best practices
* Columbia’s biobank contains over 26,000 samples.
* Of GENIE’s submissions to dbGaP, 40% are related to kidney disease phenotypes, 12% for congenital heart defects, and 12% for Alzheimer’s disease.
* They have extensive experience with phenotyping, as well as the ability to draw phenotyping standards from OHDSI (Observations Health Data Sciences and Informatics) as a unique resource for the Network.
* For ROR, Columbia used a dashboard-based approach, a tool called Harvest. The tool highlights keywords related to the phenotype being queried, and clicking on the keyword will then highlight a timestamped portion of the patient’s timeline.
* Contributions of genomic discovery from Columbia were also highlighted.
* Educational resources from Columbia were highlighted. The site currently holds a monthly seminar on ELSI issues with leading experts, as well as offering weekly year-long courses in precision medicine. Columbia also has 5 post-doctoral fellows annually that specifically investigate ELSI and precision medicine issues.

Outcomes Workgroup Report – *Hakon Hakonarson & Josh Peterson** The co-chairs reviewed carryover process outcomes measure from Phase II. All process outcomes projects are wrapping up data collected, and some have already working on publications.
* The Outcomes workgroup will develop cross-site metrics to track the implementation and impact of eMERGE III. The workgroup will focus on answering the overarching question of whether eMERGE III generated genomic results changes health care utilization and impacts patients and families.
* The following objectives for the workgroup were developed:
	+ Define and prioritize eMERGE III outcomes.
	+ Designate mandatory vs optional outcomes.
	+ Develop a framework to guide outcome assessment at all sites.
	+ Develop a reporting mechanism and schedule.
	+ Follow through on eMERGE PGx evaluation plans.
* The group will focus primarily on process outcomes, health outcomes, and economic outcomes.
* Prior to the January 2016 SC Meeting, the group will have:
	+ Identified an outcomes representative at each site.
	+ Created an Outcomes Map (outcomes impact analysis).
	+ Drafted a prioritization survey.
	+ Established sub working groups.

Genomics Workgroup Report – *Sekar Kathiresan & Megan Roy-Puckelwartz** The WG will identify best practices and facilitate analyses to assess the phenotypic impact of common and rare variant data arising from eMERGE II and III.
* The group will focus on additional validation/replication/analysis of Phase I & II GWAS datasets. They will also coordinate the integration of GWAS from the new sites in Phase III (this process will likely take 6-9 months).
* Co-chairs discussed their intent to interact with the CC and SC to identify and test possible QC and analysis pipelines for rare variant association testing.
* Collaboration will take place with the Phenotyping WG to:
	+ Identify/compile existing phenotype data.
	+ Systematically evaluate where data can be enhanced.
	+ Prioritize data points that would be most powerful for both eMERGE II and eMERGE III data.
	+ Implement processes to procure highest priority data and hasten experimental progress.
	+ These initiatives will be prioritized over the next few months.
* The SC discussed the request to implement all 41 eMERGE phenotypes across all pre-e3 data. NHGRI has provided sites with a priority list for running Phase II phenotypes –Rongling commented that the sites committed to completing their lists within 6 months- 1 year. There was hesitation about the feasibility of running all 41 phenotypes in that time frame. Harvard suggested case counts for a few ICD9 codes per phenotype first to determine which phenotypes can be run.

Harvard Site Report – *Scott Weiss** Harvard’s biobank is comprised of 50,000 patients (with a goal of 75,000 by 2018), and they also have access to 4 million patients through the Partners Health Biobank Portal.
* Harvard is submitting 25,000 subjects linked to all phenotypes that can be run. 5,000 subjects have been genotyped and deposited to dbGaP, and have also been sent to UW for imputation. Children are not consented in Harvard’s biobank, so they are not part of the eMERGE study set.
* An overview and workflow of GeneInsight was provided. GeneInsight is an IT platform developed at Partners that helps streamline EHR data.

Phenotyping Workgroup Report – *Josh Denny & George Hripcsak** The phenotyping workgroup will carry out core functions on phenotyping in eMERGE III, and will develop next-generation phenotyping methods.
* For phase III, the WG will expand beyond clean case/controls, and will also include ‘gray’ cases in the middle.
* The WG will work to identify what factors will help get Phase II phenotypes to GWAS status.
 |
| **Summary of Action Items:**1. The CSGs will pre-curate common actionable variants.
2. The CSGs will harmonize clinical variant classification between themselves.
3. The CSGs will compare report formats.
4. The CC will create a matrix of phenotype by participant and association results by phenotype for internetwork collaboration.
5. The CSG/CC Group will discuss creating a matrix of the final gene/SNP list.
6. The CC will facilitate Network efforts to annotate the gene/SNP list to include information about phenotypes of each and sample ascertainment.
7. The CC will work with the Network to catalog planned eMERGE 3 genomic submission data and provide directions for submission.
8. The CC will work with the Network to expand the existing combined data set with case/control status. Details on how this data can be made available to the Network to be discussed on the October PI call.
9. The CC will provide targeted information to the Network to explain what phenotypic data is currently available and the existing workflow to access this data based on the Publications Policy.
10. The CC will work with co-chairs to begin scheduling monthly workgroup calls.
11. The eMERGE ROR/ELSI Workgroup will be invited to CSER ROR monthly calls.
 |
| **Next Meeting: January 25-26th, 2016; Nashville, TN** |
|  |