emerge network
ELECTRONIC MEDICAL RECORDS & GENOMICS

External Scientific Panel

Packet

April 17, 2017

TABLE *of* CONTENTS

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service

National Institutes of Health
National Human Genome
Research Institute
31 Center Drive MSC 2152
Building 31, Room 4B09
Bethesda, MD 20982-2152

April 3, 2017

Dear eMERGE External Scientific Panel members,

We are glad to let you know the eMERGE investigators have made significant progress after the October 7, 2016 ESP in-person meeting. Specifically, they have: 1) sequenced 8,161 samples and issued 1,257 clinical reports to study sites; 2) applied sequencing data transfer tools, and software to transfer and store sequencing data in DNAnexus and integrated clinical reports to EMRs; 3) re-imputed the eMERGE I-II data and eMERGE III genome-wide genotyping data that eMERGE sites contributed to the network before and after the eMERGE III awards using the Haplotype Reference Consortium (HRC version r1.1) panel and Michigan Imputation Server. The imputed datasets are under quality assurance (QA) assessment and will be submitted to dbGaP after QA; 4) completed PGRNseq multi-sample calling

We appreciate the expertise and effort you have devoted to eMERGE in the past, and we look forward to your continued input, especially at the ESP teleconference on April 17, 2017.

The eMERGE Coordinating Center (CC) has prepared this booklet in collaboration with the eMERGE investigators to ensure a productive teleconference. We would like to ask that you review these materials prior to the meeting.

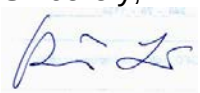
Within the booklet you will find the following important materials:

- Agenda for ESP Teleconference Call
- Network Overview
- Response to ESP Recommendations from October 2016
- Network Data Resources and Management Update
- eMERGE Workgroup progress report, and
- eMERGE Supplement Projects Report

If you have any questions, or would like more information, please do not hesitate to contact us or the CC program staff.

We look forward to seeing you at the meeting.

Sincerely,



Rongling Li, on behalf of the NHGRI eMERGE team

Rongling Li, MD, PhD, MPH
Project Director, eMERGE
Division of Genomic Medicine, NHGRI, NIH
lir2@mail.nih.gov

CC:
Jyoti Gupta
Teri Manolio
Sheethal Jose
Ken Wiley

AGENDA

eMERGE Network: External Scientific Panel Conference Call Agenda

Monday, April 17th, 2017 at 3:00-4:30pm EST (2:00-3:30pm CST; 12:00-1:30pm PST)

Dial-In Information: **1 888 936 7423** OR **+1 (510) 365-3332**

Access Code: **373-444-522**

GoTo Meeting Link: https://attendee.gototraining.com/r/2622182781112432130

Agenda Item	Discussant
• Welcome, Opening Remarks, General Updates (2 mins)	Rongling Li & Howard McLeod
• Network Introduction o Summary (5 mins) o Response to ESP Recommendations from Oct. 2016 (10 mins)	Rex Chisholm
• Sequencing and Genomics Data Update (15 mins)	Heidi Rehm, Richard Gibbs & David Crosslin
• Return of Results Workgroup Update (10 mins) o Accomplishments: IRB perspectives paper, HCP survey developed, & Participant Surveys coordinated across sites o Harmonizing activities: Cross Workgroup collaborations within eMERGE (Outcomes and Clinical Annotations WG) and externally (ClinGen, CSER) o Future work: Manuscripts, tracking return of results, surveying HCP and participants	Ingrid Holm & Iftikhar Kullo
• Discussion and Suggestions from ESP (20 mins)	ESP
• Executive Session (30 mins)	Rongling Li

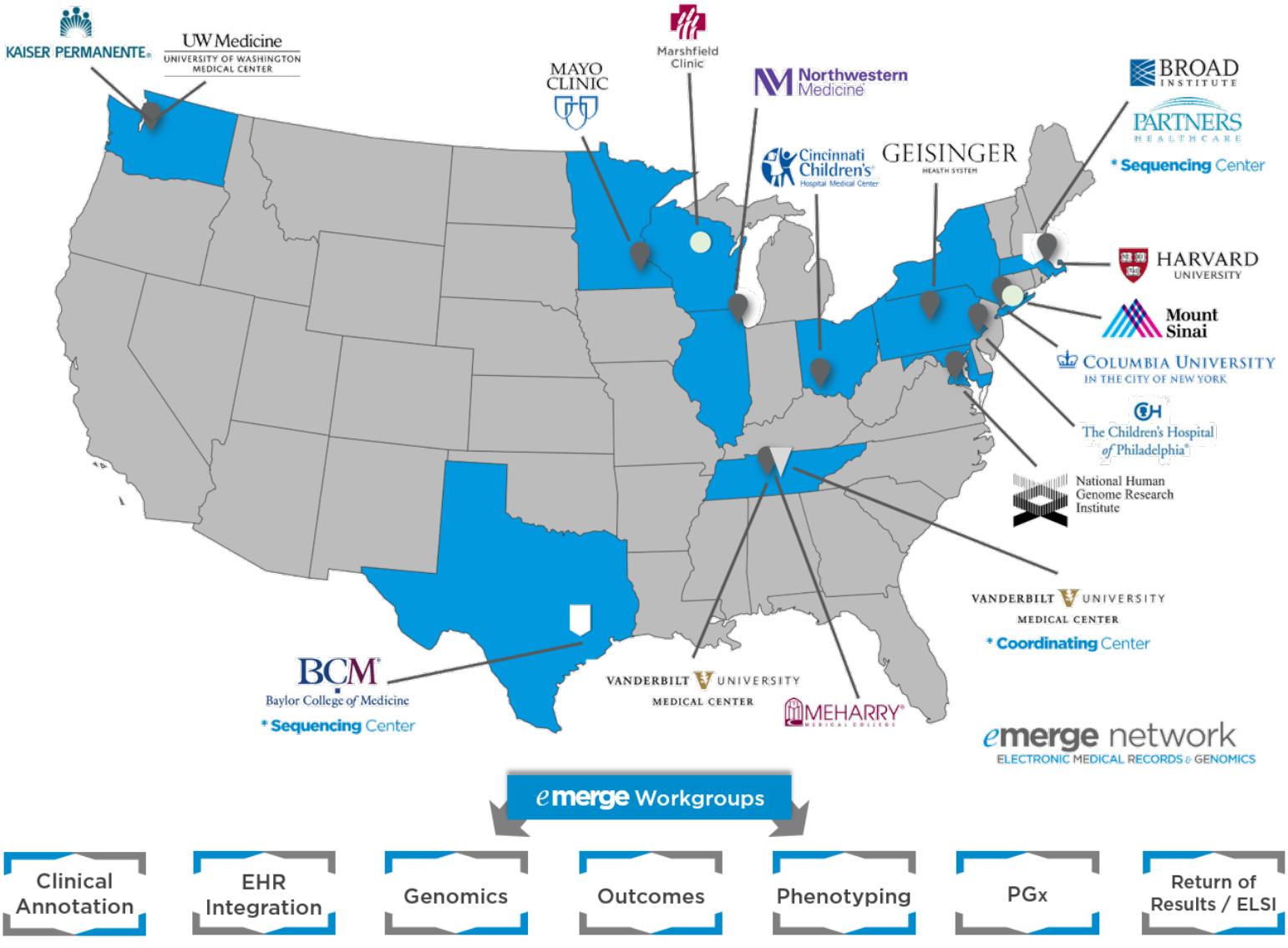
Next Meeting: Tuesday, October 10th, 2017 in Bethesda, MD

emerge network
ELECTRONIC MEDICAL RECORDS & GENOMICS

NETWORK OVERVIEW

eMERGE is a national consortium, organized by NHGRI, that conducts discovery and clinical implementation research in genomics and genomic medicine at medical research institutions across the country. eMERGE research combines DNA biorepositories with electronic health record (EHR) systems for large-scale, high-throughput genetics research with the ultimate goal of returning genomic testing results to patients in a clinical care setting. eMERGE researchers are experts in the diverse fields of genomics, statistics, ethics, informatics, and clinical medicine

During Phases I and II, the Network deployed 37 electronic phenotype algorithms across more than 58,000 subjects with dense genomic data, and more than 30 new phenotypes are prioritized for genomic and targeted sequencing data during eMERGE III. Genetic data housed in the eMERGE Network includes GWAS array, exome sequencing, whole genome sequencing, and pharmacogenomics panels. In addition to the 94,000 previous subjects enrolled, an additional 25,000 subjects are being sequenced during Phase III. We designed a new sequencing panel specific for eMERGE for these subjects. This panel includes 109 genes and 1551 SNVs of interest. Of these 68 genes and 14 SNVs are deemed actionable and will be returned to patients. Sites are completing enrollment, our two sequencing centers have begun sequencing samples, and return of genetic data has already commenced. Sites across the network have implemented institution-specific models of pharmacogenomics, returning drug metabolism information in the clinic. Themes of bioinformatics, genomic medicine, privacy, community engagement, and human subjects protections are of particular relevance to eMERGE.



emerge Workgroups

- Clinical Annotation
- EHR Integration
- Genomics
- Outcomes
- Phenotyping
- PGx
- Return of Results / ELSI

ESP RECOMMENDATIONS *from* OCTOBER 2016

To investigators:

1) The Network should explore best practices among the different sites and work together to resolve any conflicts so as to ensure consistency in the process of return of results.

a. The ROR/ELSI Workgroup has generated a table detailing site plans for return of results with the intention to summarize the data in order to assess similarities and differences across sites. In effect, the Workgroup aims to identify return of result pipelines and common deviations from those pipelines.
b. The ROR/ELSI Workgroup has a manuscript in process summarizing how IRBs at the various eMERGE sites reviewed each site's proposed return of results.
c. The ROR/ELSI Workgroup formed a Participant Survey subgroup aimed to harmonize domains in baseline, post-ROR and follow-up surveys across sites.
d. The ROR/ELSI Workgroup was funded for a supplement, Health Care Provider (HCP) Survey, aimed to survey preferences of HCPs who are receiving genomic information as part of eMERGE Phase III. This project will collect, analyze, and publish data surrounding the concerns and challenges of HCPs desirability, utility, actionability, and meaningfulness of incorporating results of genome sequencing (targeted sequencing of disease gene panels, WES, WGS) into clinical care.
e. The Clinical Annotation Workgroup developed a pipeline for discussing non-standard results to resolve any conflicts in a variant's interpretation/actionability. Based on that pipeline, they make recommendations to the ROR/ELSI group on returning results, adding to the Network's consistency in returning results.

2) The Network sites should create a timeline that details the period between receiving the clinical reports from the CSGs and returning results back to the patients.

The ROR/ELSI Workgroup will create a template that can be used at each site to follow the temporal course of clinical reports as these are created and eventually disclosed. Important dates include: draft report created; report finalized and 'released'; accessed by the site; reviewed; site approval of the report; contact to set up ROR visit; actual ROR visit takes place

3) The Network should address the ELSI issues as well as the scientific issues that come up when PIs decide what results they will return to the patients and they should publish on challenges faced, issues unresolved, lessons learned, and best practices developed.

An eMERGE Legal Considerations subgroup, comprised mostly of JDs within the Network, was formed. This group is focused on legal (HIPAA and related) considerations relevant to the identification of, and potential communication with, family members of patients for whom actionable genetic risk information has been identified. Specifically, the group will focus on identifying ramifications for health care systems surrounding these issues at each site.

(Continued on next page)

ESP RECOMMENDATIONS *from* OCTOBER 2016

To investigators:

4) The Network PIs need to document their decision-making process regarding which variants they choose to return to their patients, record the number of times PIs make changes in what was proposed to be returned, analyze these decisions systematically and then disseminate this analysis through publications.

a. The complete list of genes and SNPs on the eMERGEseq platform includes both clinically 'returnable' and 'not clinically returnable' genes and SNPs.
b. The Clinical Annotation Workgroup developed a consensus list of genes and SNVs included on the eMERGESeq panel that all of the sites agreed would generally be returnable. This is now called the "Consensus Actionable List".
c. However, because of study specific variables, some sites deviate from this default plan by either not returning all of the consensus list or adding non-consensus content to their reportable list. The sites have communicated to the CSGs their site-specific requirements for what is included on their reports and these have been tabulated.
d. In addition to these up-front agreed upon differences, some sites may not return all content on a case-by-case basis. We will request that sites track these deviations and report back to the CSGs.
e. The Clinical Annotation Workgroup also discusses all non-standard results, results with difficult interpretations and any variants that are unresolvable between the CSGs (see #5 below) in order to resolve those differences and/or track their occurrence if not resolvable. The CC is tracking all of these decisions through google notes and a google worksheet.

5) The sites should assess the impact of the inconsistency between the 2 CSGs' reports on the Network goals and attempt to minimize differences across clinical sites based on the CSG to which they're assigned.

The CSGs executed a variant harmonization effort to assess differences in approaches to variant interpretation. This showed remarkable similarity in approach and any major differences were resolved. Currently, all of Partners-Broad interpreted variants are automatically viewable in GeneInsight and BCM routinely uploads all variant reports to the GeneInsight database so all reported pathogenic and likely pathogenic variants seen in BCM-eMERGE samples can be viewed by Partners-Broad. Partners-Broad also sends Baylor, on a monthly basis, a full list of interpreted variants in all 109 eMERGE panel genes. Baylor compares this list to their internal database and reports out any interpretation discrepancies. This discrepancy list is checked by both CSGs before reporting any cases. Discrepancies in variants intended for reporting are discussed in a biweekly call to attempt resolution. Unresolvable variants would be presented on the Clinical Annotation WG calls (see #4 above).

6) The Network PIs should work with the HL7 standards group through the HL7 Clinical Genetics Working Group (WG) teleconferences.

The Network will continue to leverage eMERGE investigators and staff directly involved with the HL7 Clinical Genomics Workgroup. Larry Babb, a member of the Partners-Broad CSG as well as the Sunquest liaison for the eMERGE GeneInsight instance is a member of the HL7 CG WG and Robert Freimuth, a member of Mayo's eMERGE site is co-chair of the HL7 Clinical Genomics WG. These two individuals enable coordination with any developing standards and provide assurance that all results reported to eMERGE participants will conform to HL7 standards for transmission to EMRs. Larry and Robert, who are members of eMERGE's EHR Workgroup, also facilitate feedback from eMERGE investigators to enable eMERGE input to the evolving HL7 standards.

ESP RECOMMENDATIONS *from* OCTOBER 2016

To investigators:

7) The Network should take the opportunity to describe and analyze the process of switching EMR systems and its impact on eMERGE research and implementation.

The EHRI group is actively monitoring EMR transitions at Network sites. The Workgroup tracks many barriers to implementation and plans to write a manuscript and present best practices after data has been collected at all the sites.

8) The Network members should work together to create more collaborative, Network-wide products.

a. The ROR/ELSI Workgroup is developing survey instruments to assess the impact of ROR on providers and patients across all Network sites.
b. The Return of Results and Outcomes Workgroups are collaborating to explore Network-wide familial implications of ROR.
c. The ROR/ELSI and Outcomes Workgroups are collaborating on an outcomes data collection database, wherein ROR/ELSI is developing a return-of-result-information capture form to be triggered on a per-record basis for outcomes data capture.
d. Network-wide geocoding efforts are underway to assess the impact of environmental and genetic factors on phenotypic expression of disease.
e. The EHRI Workgroup, through its Infobutton Subgroup, is leading an effort to coordinate Infobutton and information resource activities across eMERGE as well as collaboratively with other consortia. The group aims to expand the development and evaluation of a content management system (CMS) to optimize the reuse of quality information on genetic screening and testing. In addition, the group aims to develop a plan for sustainability for the adoption of Infobuttons across the eMERGE sites.
f. The Clinical Annotations Workgroup is collaborating with ROR/ELSI and the clinical sequencing centers on a manuscript that focuses on incidental and secondary findings from the first 10,000 eMERGESeq participants.
g. Workgroups have overlapping membership across subject, representatives are actively engaged in multiple topics and report back, streamlining work and ensuring input is gathered from appropriate members.

9) The investigators are encouraged to consider the 'ancillary study' mechanism as a collaborative approach in order to broaden access to the national resource represented by the eMERGE Network.

a. The Heath Care Provider Survey sub-group collectively applied for a RO1 grant in order to implement the survey at the individual sites.
b. The CC receives email alerts from the NIH for Request For Applications and also tracks the NIHFunding Twitter account on a weekly basis. The CC forwards any relevant requests out to the Network.

ESP RECOMMENDATIONS *from* OCTOBER 2016

To investigators:

10) The Network should actively disseminate lessons learned, best practices, experiences in conducting research, and results from the genomic discovery and clinical implementation research using EHR and biorepositories to the scientific community.

a. The Network disseminates findings through a variety of mechanisms including social media, the eMERGE website, and through its many publications.
b. The CSGs are planning on writing a manuscript based on their platform validation of the eMERGE sequencing panel.
c. CERC survey paper recently disseminated the findings of their survey: Public Attitudes toward Consent and Data Sharing in Biobank Research: A Large Multi-site Experimental Survey in the US.
d. We currently have 496 published manuscripts, 99 in development, and have had 62 presentations at over the last six months at the American Medical Informatics Association (AMIA), American College of Medical Genetics and Genomics (ACMG), and American Society of Human Genetics (ASHG) conferences.

(NHGRI responses continued on next page)

ESP RECOMMENDATIONS *from* OCTOBER 2016

To NHGRI:

1) NHGRI Program Staff should make sure that the sites continue to collect outcomes and study the impact on return of results for the PGx project. NHGRI should ensure that any newly discovered risks are reported to variant carriers in the PGx participants.

The NHGRI staff emphasized the importance of collecting outcomes in the PGx project during the February Steering Committee meeting as well as during the various working group calls. The progress of the PGx workgroup (WG) is listed below:

a. The PGx WG is working with the Clinical Annotations workgroup to develop a pipeline for return of PGx results in eMERGE III and they have been collecting additional outcomes data.
b. They worked with the Phenotyping workgroup to integrate and prioritize PGx-related outcomes such as drug response in the Network-wide phenotype prioritization list.
c. The PGx WG continues to discuss how best to collect and disseminate outcomes from the PGx study. They are working to move several papers in progress through to publication, including a project examining outcomes assessed through collection of data related to CDS activations (first author Tim Herr). They hope to use a recently published Translational Pharmacogenetics Program (TPP) paper (first author Luzum, pmid 28090649) as an example to create a new survey on intermediate-range outcomes to disseminate to the group (to collect data for a future publication).
d. Unfortunately, the PGRNSeq data was generated on research, not clinical, DNA samples. If new rare genetic variants are detected in PGRNSeq data, most sites are not able to return this information to participants. However, if CPIC guidelines are updated for genetic results already returned, the PGx group will work with the EHRI group to make sure that clinical decision support (CDS) is updated to reflect these new guidelines. To facilitate this, the PGx working group have recently created a grid where they track the dates of activity for all the CDS implemented about eMERGE-PGx.

2) NHGRI Program Staff should encourage more collaborative projects among the eMERGE Network members.

As mentioned in recommendation #8 to the Network above, the NHGRI Program Staff has encouraged network collaborations. There have been several ongoing collaborative projects since the last ESP meeting in October 2016. Examples of collaboration among eMERGE working groups are listed below:

a. The Data Integration team, that developed a multi-lab to multi-site network for sharing variant knowledge and de-identified clinical cases in a manner that maintains CLIA compliance, has merged with the Electronic Health Record Integration (EHRI) Workgroup. Together, they continue mapping clinical report elements to a harmonized XML structure that can be feed into EHRs.
b. The Outcomes WG is collaborating with the Return-of-Results/Ethical Legal Social Implications (ROR/ELSI) WG to survey patients and address the timing of ROR, healthcare costs and utilization, and costs to participants. They will also work with the Familial Implications of ROR subgroup to address familial implications from the participant perspective.
c. The Clinical Annotations workgroup is collaborating with the ROR/ELSI workgroup to address how patients and providers react to variant reclassification.
d. Investigators from both eMERGE and CSER consortia discussed opportunities for collaboration, such as developing tools/best practices and to implement successful family cascade projects; and creating a road map/process for coordinating family genetic testing (see details in the joint meeting minutes). The meeting materials can be found at the NHGRI website (https://www.genome.gov/27567557).

*e*MERGE-Seq OVERVIEW & CLINICAL REPORTING

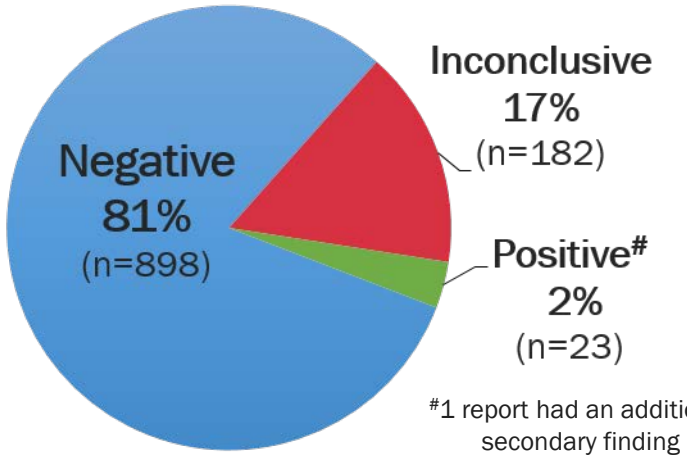
PARTNERS-BROAD STATUS UPDATE: Sequencing

	Samples received			Sequencing status and performance				Reporting
	Round 1	Round 2	**Total**	Completed Sequencing	Mean Coverage	Bases @ 20X (%)	Targets covered (%)	Reports issued
UW/Group Health	1221	1178	**2399**	1221	436	99.6	99.8	375
Geisinger	1250	0	**1250**	1250	449	99.8	99.8	0
CCHMC	1494	0	**1494**	544	**Will be calculated after Round 1**			0
Harvard	1269	1268	**2537**	0	N/A	N/A	N/A	N/A
Total	5234	2446	**7680**	3015				375

Data provided accurate as of 3-24-2017
N/A indicates items have not been received or fully processed at this time

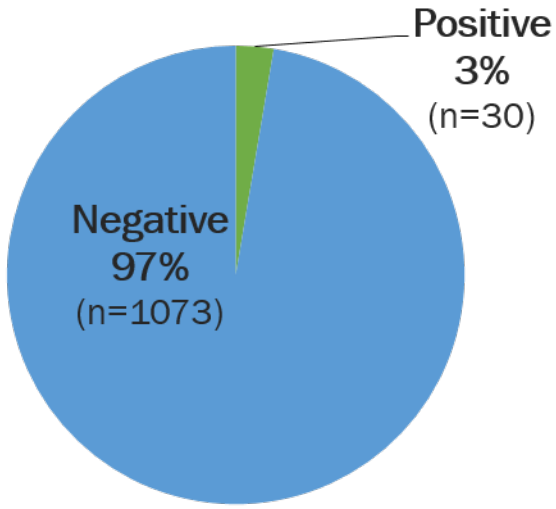
PARTNERS-BROAD STATUS UPDATE: Interpretation & Reporting

Diagnostic/Primary Findings



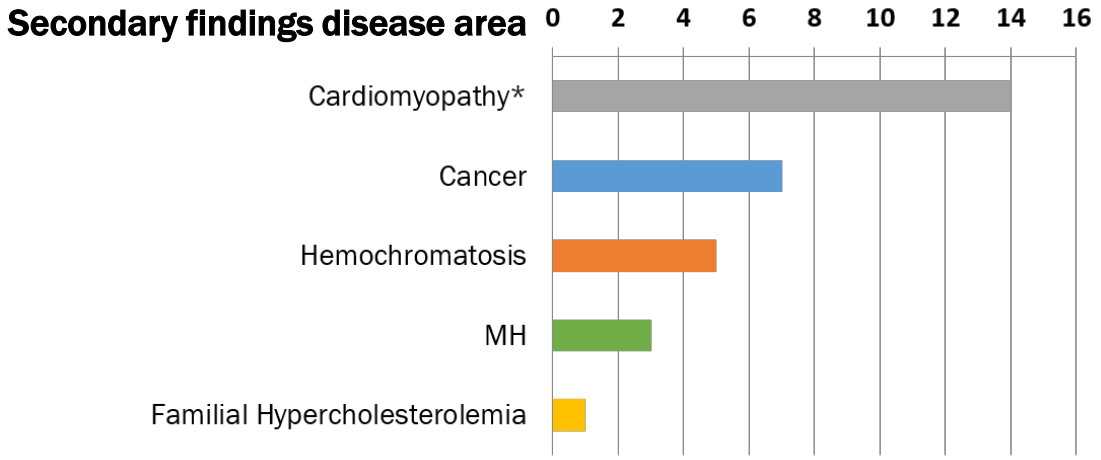
#1 report had an additional secondary finding

Additional/Secondary Findings



Case Review *of* 1103/1221 UW Samples

- Reportable Content: from Clinical Annotation Workgroup developed consensus actionable list only
- 375 reported
- 29 awaiting sign out (positive + inconclusive)
- 69 awaiting variant confirmation and drafting (positive + inconclusive)
- 630 negative reports awaiting drafting and/or sign out



* Also has path HFE variant

PARTNERS-BROAD STATUS UPDATE: Other

REPORTING: PGx

- Plan to issue batched report along with accompanying excel document for further parsing and consumption by sites, validation strategy underway.
- Report content: CPIC level A genes, diplotype and metabolizer phenotype. Supplemental tables: * allele translation tables, CPIC dosing recommendations, etc.
- CSGs discussed reporting plans in PGx WG, current efforts directed at content harmonization to ensure reporting consistency.
- Mock report draft with all possible diplotype combinations was presented by CSGs to PGx WG for further input

NON-STANDARD RESULTS

- **Trisomy 12 detected and confirmed in CRC case**
  - Discussed in Clinical Annotation WG, deemed reportable
  - Will issue non-standard site report with possible explanations/associations (mosaicism, CML)

BAYLOR HGSC STATUS UPDATE: Sequencing

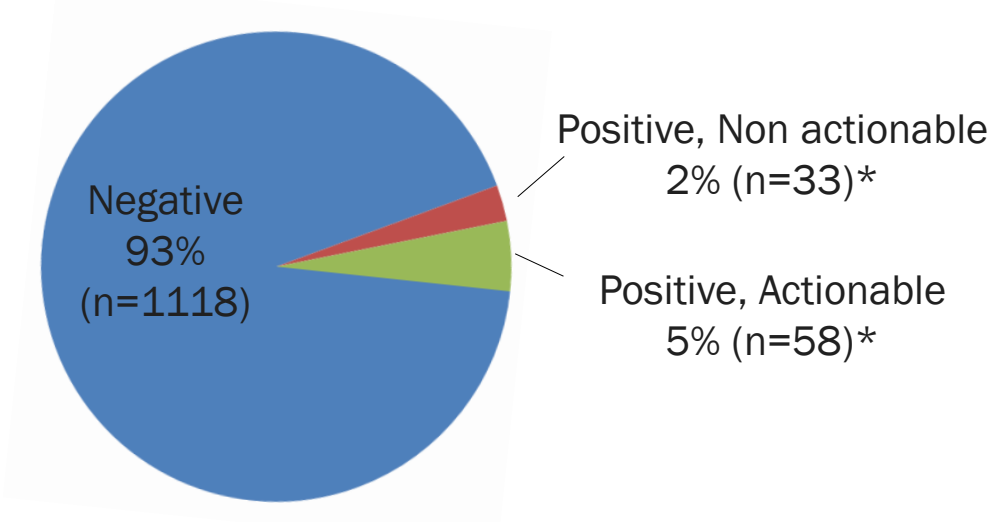
	Samples received			Sequencing status and performance				Reporting
	Round 1	Round 2	**Total**	Completed Sequencing	Mean Coverage	Bases @ 20X (%)	Targets covered (%)	Reports issued
Northwestern	1504	0	**1504**	1400	336	99.66	99.90	882
Mayo	2538	0	**2538**	2347	342	99.75	99.95	0
CHOP	1410	0	**1410**	844	333	99.90	99.98	0
Vanderbilt	1058	0	**1058**	468	342	99.90	99.98	0
Columbia	396	0	**396**	87	336	99.91	99.94	0
Meharry	0	0	**0**	0	NA	NA	NA	NA
Total	6906	0	**6906**	5146				882

Data provided accurate as of 3-24-2017
N/A indicates items have not been received or fully processed at this time

BAYLOR HGSC STATUS UPDATE: Interpretation & Reporting

Positive Findings

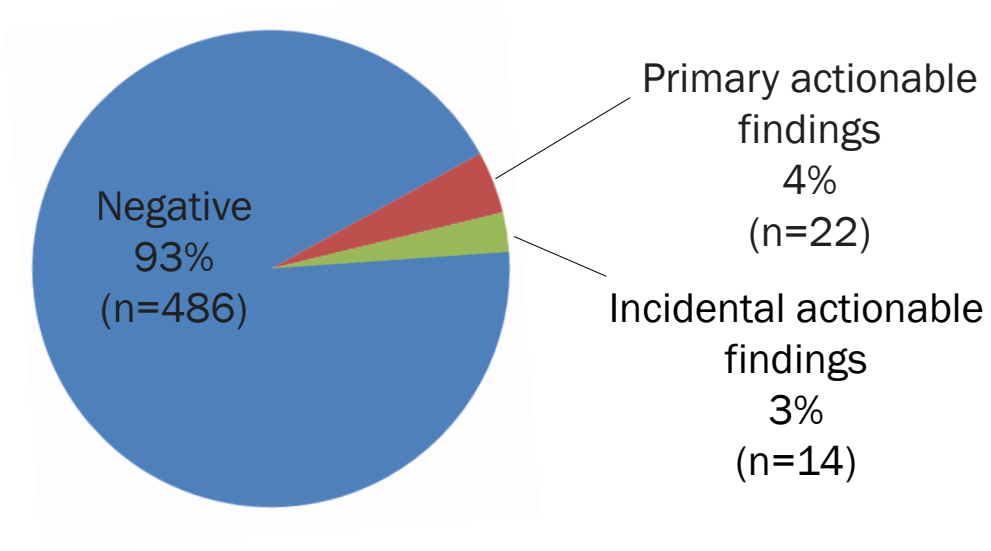
1,211 NU Samples



*3 patients had one actionable and one non actionable variant
5 mosaic findings not included

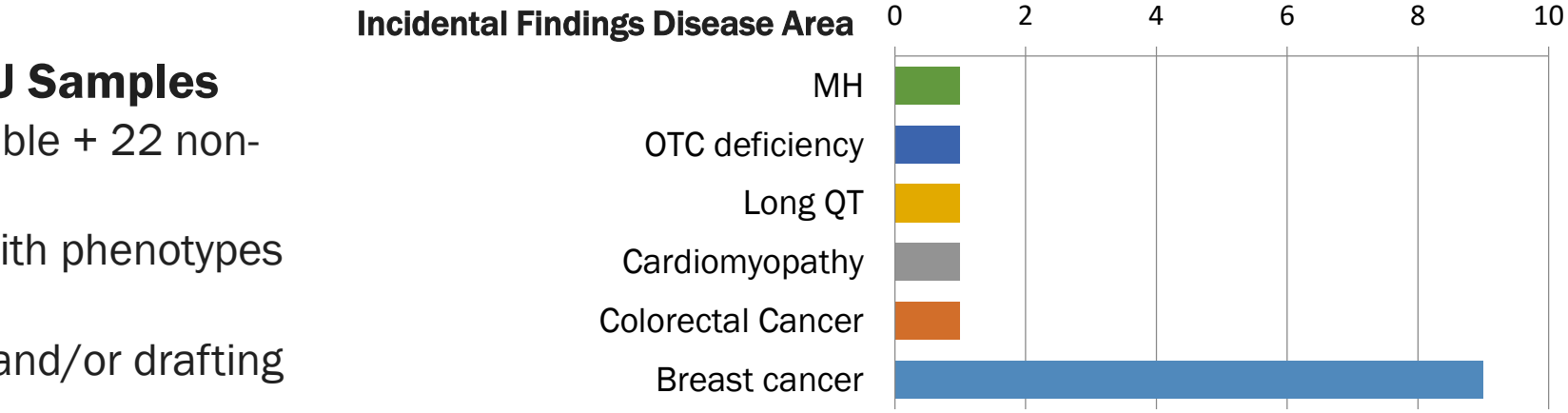
Primary and Incidental Findings

522 NU Samples with phenotype



Case Review *of* 1211 NU Samples

- Reporting for 90 genes (68 actionable + 22 non-actionable) and 28 SNPs
- 1211 NU patients, including 522 with phenotypes
- 882 reported
- 329 awaiting variant confirmation and/or drafting including 5 mosaic variants



BAYLOR HGSC STATUS UPDATE: Other

SEQUENCING

- Low input DNA protocol validated (<500ng)

REPORTING

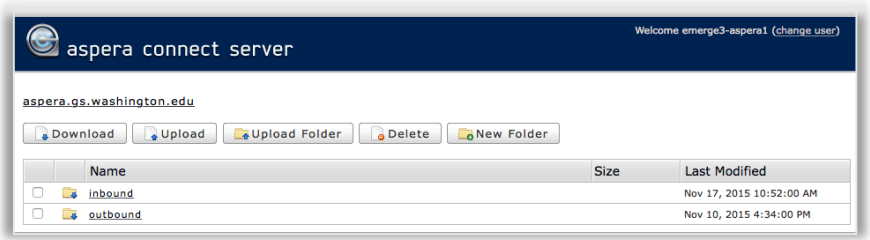
- Most effort clarifying 'report content issues'
- Additional tables will be added to the report for e.g. variants that lack disease association
- PGx Reporting:
  - PGx diplotypes table added to the clinical report with interpretation
  - Determining other site-specific PGx requirements

NON-STANDARD RESULTS

- Mosaic findings in cancer-related genes
  - Discussed with Northwestern University: Two mosaic TP53 variants confirmed by Sanger sequencing in older patients, likely somatic. A decision was made not to report.
  - 5 additional mosaic variants awaiting Sanger confirmation

NETWORK DATA RESOURCES & MANAGEMENT

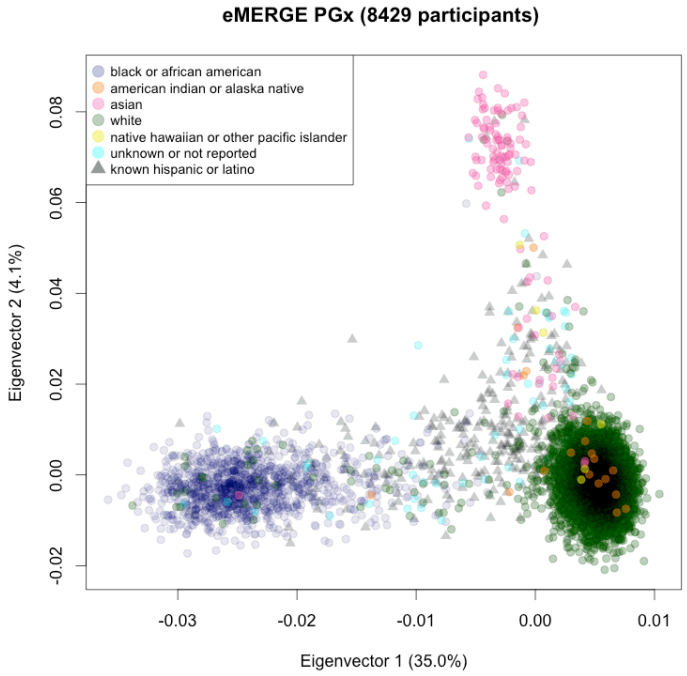
Aspera High-Speed File Transfer Software *at the* University *of* Washington



- At the beginning of eMERGE 3, the Coordinating Center invested in an Aspera server in the Department of Genome Sciences.
- This Aspera server has a dedicated 10Gb/sec Science DMZ/I2 network link for data dissemination and acquisition
- Over the last six months, eMERGE sites have deposited **111.3 terabytes** of data and have downloaded **130 terabytes**.
- DNAnexus staff also have an account, so all data available to the Commons.
- All legacy and analysis data are available for download including the following:
  - All eMERGE I, 2, and 3 array data
  - All imputed data, individual files
  - Multisample PGRNseq VCF (will soon be updated with new file)
  - We will also store eMERGEseq data for the duration of the project

GENOMIC DATA: PGRNseq Multi-sample Calling

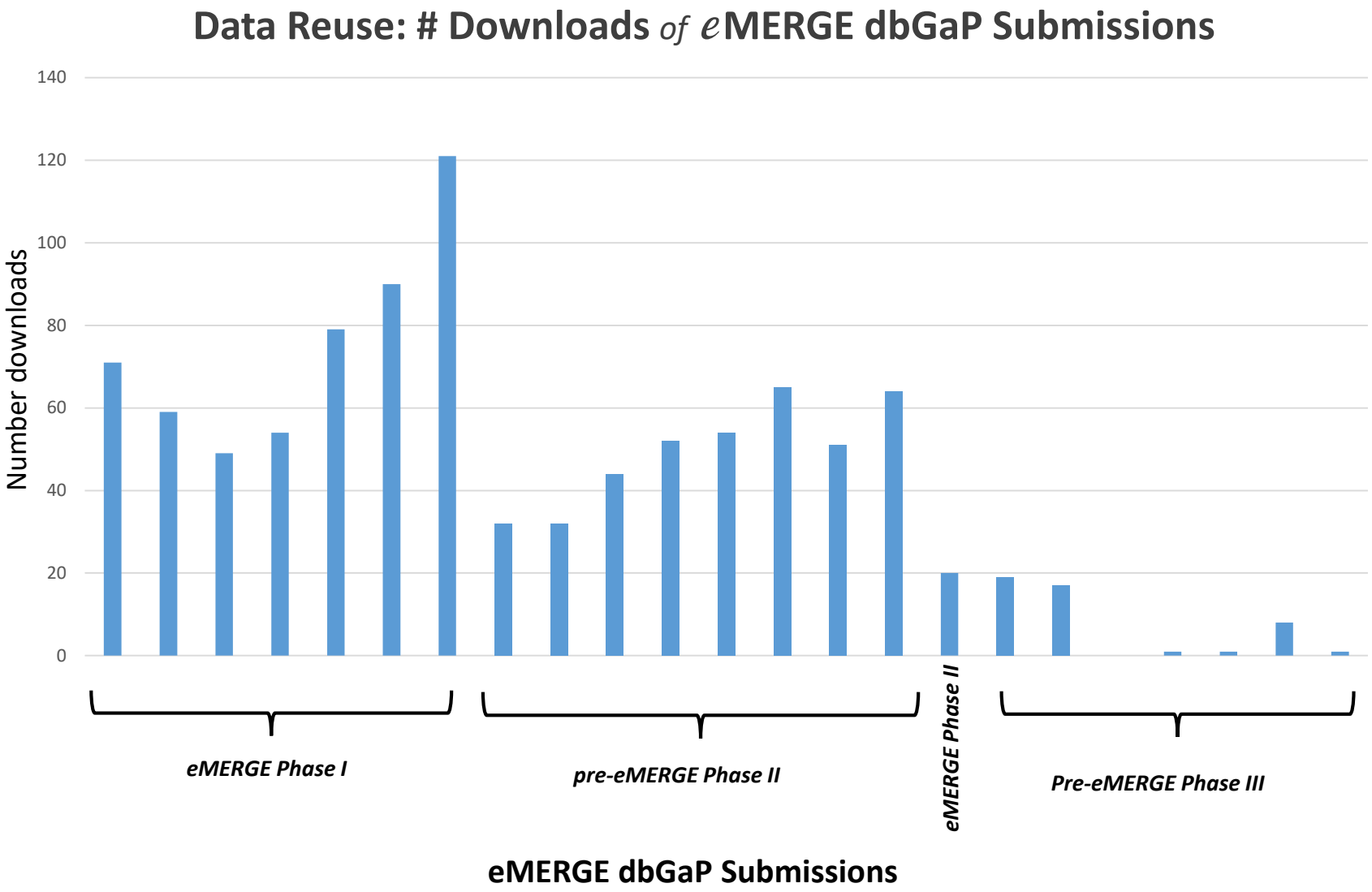
- All ~9000 BAMs re-aligned to the same genome reference hs37d5.fa (plus decoy)
  - There were multiple references used by the five sequencing centers
- VCF of PGRNseq multisample re-created (March, 2017) and provided to the Network for analysis through Aspera and DNANexus
  - Variant-only sites released in 8429 total samples included, ~5% of the original BAMs lacked paired mates, requiring the removal of 583 samples (have since been repaired by sequencing center and will attempt to add)
- Plan to add additional annotation to SPHINX to provide more bioinformatic data to users
  - SeattleSeq, Annovar, SnpEff, etc.
- Search capabilities like UCSC Genome Browser
  - e.g. search by chr#:base_pair_location
- Will also provide an additional summary of frequency in the Asian ancestry cohort in addition African and European ancestries
- Drs. Stanaway, Gordon, and Crosslin will lead Network-wide project describing the ~9000 participant data set; A manuscript concept sheet has been circulated
- Report released to the Network: link



GENOMIC DATA: Imputation

- Direction from eMERGE Leadership and the eMERGE Genomics Working Group Co-Chairs
  - Issues with prior imputed data prompted the imputation of all eMERGE array data, including the legacy data, against the HRC reference using the Michigan Imputation Server.
- HRC reference contains 39,235,157 SNPs, no indels, and will provide access to rare variation (low as 0.1%)
- 86,151 samples pre-QC, 81 total chip arrays
  - eMERGE I & II = 55,029 samples; eMERGE III = 31,122 samples
- The newly imputed merged set is expected to be released in April to the Network after QC
- Minimac calls and phases the genotypes directly: 0|0, 1|0, 0|1, 1|1 and gives genotype posterior probabilities in the VCF

IMPACT: dbGaP & Website Analytics



eMERGE Website

Average usage past 6 months

- 61.5% new visitors
- 1902 sessions/month
- 1206 users/month
- Views from 73 countries

PheKB Website

Average usage past 6 months

- 55.4% new visitors
- 966 sessions/month
- 553 users/month
- Views from 57 countries

dbGaP Update: eMERGE Merged II set now accessible via the dbGap site (Cohort size: 55,029)

*e*MERGE WORKGROUP PROGRESS

Clinical Annotation

Co-Chairs: Gail Jarvik (KPW/UW) & Heidi Rehm (Partners/Broad)

EHR Integration

Co-Chairs: Sandy Aronson (Harvard) & Casey Overby (Geisinger/JHU)

Genomics

Co-Chairs: Megan Roy-Puckelwartz (NU), Patrick Sleiman (CHOP) & David Crosslin (KPW/UW)

Outcomes

Co-Chairs: Hakon Hakonarson (CHOP), Josh Peterson (Vanderbilt/CC), & Marc Williams (Geisinger)

PGx

Co-Chairs: Laura Rasmussen-Torvik (Northwestern) & Cindy Prows (CCMHC)

Phenotyping

Co-Chairs: George Hripcsak (Columbia) & Peggy Peissig (Marshfield)

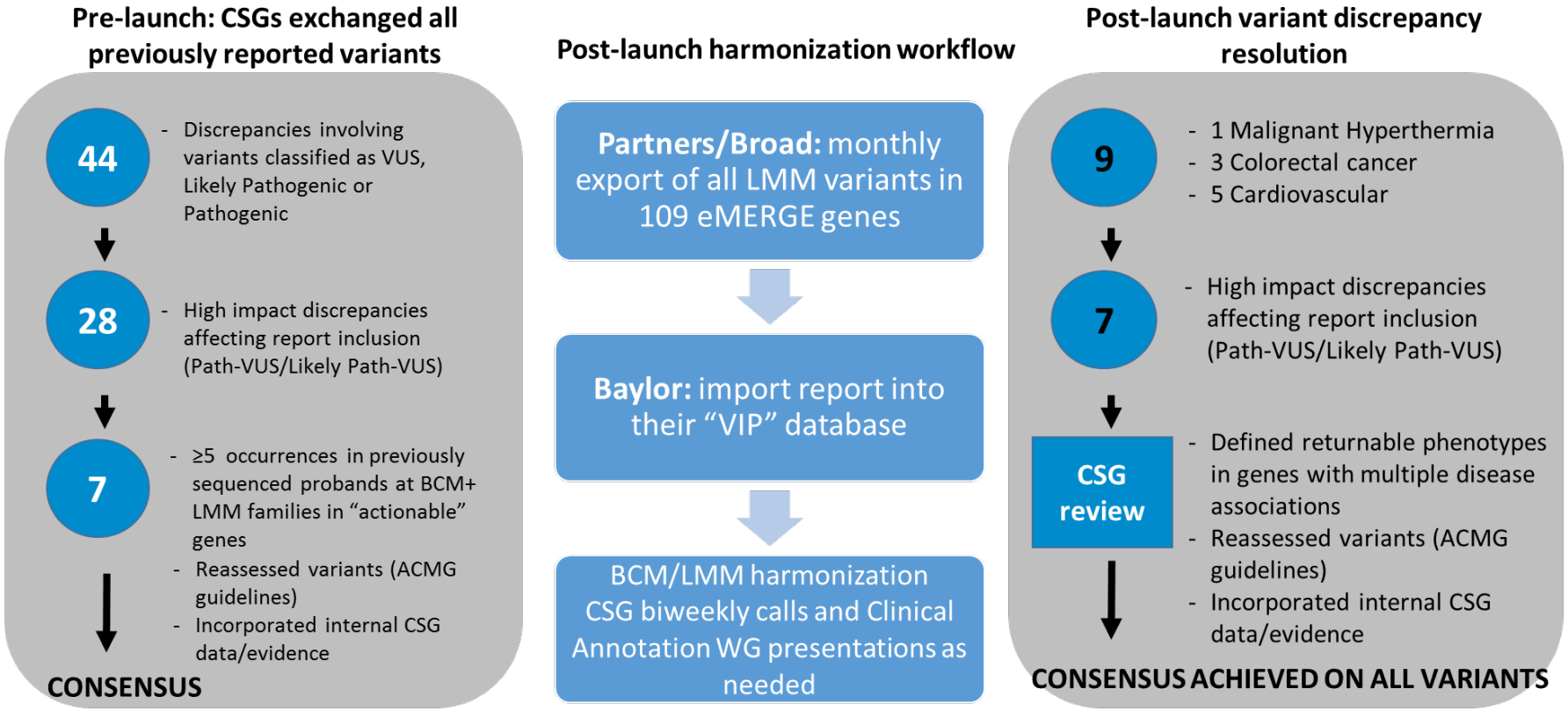
RoR/ELSI

Co-Chairs: Ingrid Holm (BCH) & Iftikhar Kullo (Mayo)

*e*MERGE CLINICAL ANNOTATION WORKGROUP: Status & Accomplishments

Co-Chairs: Gail Jarvik (KPW/UW) & Heidi Rehm (Partners/Broad)

- ✓ Completed clinical validity and actionability review of 109 proposed genes and 1551 SNVs for inclusion on the eMERGE Sequencing Panel using ClinGen approaches: Consensus Actionable List = 68 genes and 14 SNVs
- ✓ Developed a model for data sharing for both variant and de-identified case-level data using ClinVar and the GeneInsight Network, including recent approval for eMERGE to join the VariantWire network supporting clinical lab data sharing
- ✓ Developed an approach to ensure that variant interpretation is consistent across the CSGs and the network including a pre-launch harmonization effort between the CSGs and a post-launch process for ensuring ongoing consistency (see figure)



eMERGE CLINICAL ANNOTATION WORKGROUP: Future Efforts

Co-Chairs: Gail Jarvik (KPW/UW) & Heidi Rehm (Partners/Broad)

- The CSGs, in conjunction with our WG, are actively developing a standardized approach for the return of pharmacogenomics data from the eMERGE participants to each site. This involves working with the PGx WG to ensure adherence to standards and compatibility with network site needs.

- **Publications in progress or planned:**

    1. Applying a structured curation process to determine a minimal consensus set of actionable genes and variants for universal use on diverse biobank cohorts (concept sheet approved)

    2. Incidental and secondary findings (IFs) in 9,000 eMERGE participants (concept sheet approved)

    3. Variant harmonization efforts and interpretation results from the eMERGE Gene Panel (in planning)

eMERGE EHR INTEGRATION WORKGROUP: Status & Accomplishments

Co-Chairs: Sandy Aronson (Harvard) & Casey Overby (Geisinger/JHU)

Engineering

- Provided feedback on requirements for XML format
- Established cross transfer of de-identified case data between GeneInsight and DNAnexus repositories
- Achieved 1st transfers of identified clinical genetic lab reports in XML format
- Set-up a GitHub site for sharing XML file parsers
- Established a process to track milestones on monthly EHRI WG calls
- DocUBuild infrastructure to lower barriers to using infobuttons

Science

- "Empowering Genomic Medicine by Establishing Critical Data Flows: The eMERGE Example" paper in preparation
- Data collection underway for a program evaluation of genomic infobutton initiatives

Community

- Held a joint CSER/eMERGE EHR WG meeting and began planning for a collaboration on a Lynch Syndrome DIGITizE clinical decision support implementation guide
- Collaboration with ClinGen EHR WG on evaluation of genomic infobutton initiatives
- DocUBuild demonstrations to IGNITE CIIG, NHGRI Inter-Society Coordinating Committee, and GA4GH eHealth

eMERGE EHR INTEGRATION WORKGROUP: Future Efforts

Co-Chairs: Sandy Aronson (Harvard) & Casey Overby (Geisinger/JHU)

Engineering

- Enable local use of results from central sequencing centers
  - Share scripts to translate results received from central sequencing centers into local store formats on GitHub
  - Share experiences developing clinical decision support functionality based on eMERGE data

Science

- Longitudinal study of barriers to implementation
- Survey of IT capabilities and mechanisms for decision support delivery & reporting

Community

- Recent submissions to present at AMIA 2017 (under review)
- Plans to follow-up with survey respondents expressing an interest more information about genomic infobutton initiatives
- Plans to share results from genomic infobutton program evaluation with liaison group
- Collaboration with CSER to assess costs associated with implementing pharmacogenomic clinical decision support

eMERGE GENOMICS WORKGROUP: Status & Accomplishments

Co-Chairs: Megan Roy-Puckelwartz (NU), Patrick Sleiman (CHOP) & David Crosslin (KPW/UW)

The following projects were implemented or facilitated by the Genomics Workgroup

DNAnexus

- Basic infrastructure including access/permissions has been determined (implementation ongoing)
- Pipeline for data analysis determined
- Necessary tools identified (a number of tools are already available on DNAnexus)
  - For tools not yet available, Genomics WG is facilitating tool development
- In-person training being scheduled
  - Additional web-based training will be offered as needed

DataSet Availability

- PGRNSeq
  - New multisample call complete and available
  - Report and PCA analysis complete on these data
- HRC imputation complete for eMERGE I, II, & III data
  - Merged imputation data and QC will be available to the Network late-April.
- Phenotype Data
  - CC gathering/collating phenotype data for imputed data set and additional legacy data
    - These data will include basic demographic information, available to all members
- All data are available on UW's Aspera servers

The WG is also discussing updates of SPHINX, including the addition of demographic information by variants

eMERGE GENOMICS WORKGROUP: Future Efforts

Co-Chairs: Megan Roy-Puckelwartz (NU), Patrick Sleiman (CHOP) & David Crosslin (KPW/UW)

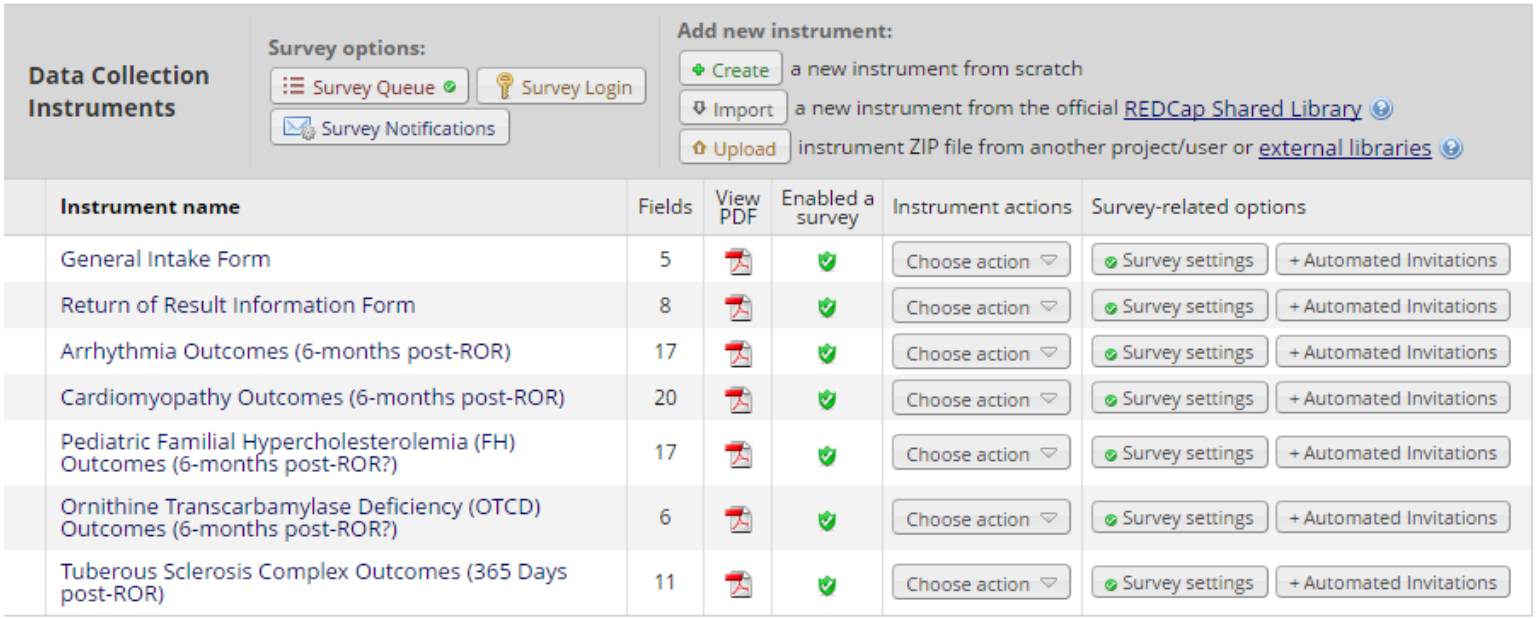
- The Genomics Workgroup will propose 3 Network-wide genetic analysis manuscripts for the PGRNseq (n~9,000), HRC-imputed (n~86,000), and eMERGE-Seq (n~25,000), respectively.
- The Genomics Workgroup will leverage the above-mentioned datasets and EHR-derived phenotypes to collaborate with other consortia, such as CSER, GIANT, and TOPMed.
- The Genomics Workgroup will continue to provide guidance to the eMERGE CC regarding genetic data activities.
- The Genomics Workgroup will continued to provide guidance to the DNAnexus Group regarding genetic and phenotype data organization, and analysis tools.

eMERGE OUTCOMES WORKGROUP: Status & Accomplishments

Co-Chairs: Hakon Hakonarson (CHOP), Josh Peterson (Vanderbilt/CC), & Marc Williams (Geisinger)

Developing Data Collection Instruments and Cross-Site Outcome Data Infrastructure with REDCap

https://emerge.mc.vanderbilt.edu/workgroups/workgroup-outcomes/



Data Collection Instruments

Survey options: Survey Queue | Survey Login | Survey Notifications

Add new instrument:
- Create a new instrument from scratch
- Import a new instrument from the official REDCap Shared Library
- Upload instrument ZIP file from another project/user or external libraries

Instrument name	Fields	View PDF	Enabled a survey	Instrument actions	Survey-related options
General Intake Form	5			Choose action	Survey settings + Automated Invitations
Return of Result Information Form	8			Choose action	Survey settings + Automated Invitations
Arrhythmia Outcomes (6-months post-ROR)	17			Choose action	Survey settings + Automated Invitations
Cardiomyopathy Outcomes (6-months post-ROR)	20			Choose action	Survey settings + Automated Invitations
Pediatric Familial Hypercholesterolemia (FH) Outcomes (6-months post-ROR?)	17			Choose action	Survey settings + Automated Invitations
Ornithine Transcarbamylase Deficiency (OTCD) Outcomes (6-months post-ROR?)	6			Choose action	Survey settings + Automated Invitations
Tuberous Sclerosis Complex Outcomes (365 Days post-ROR)	11			Choose action	Survey settings + Automated Invitations

Map Universe of Possible Outcomes for Returned Genes

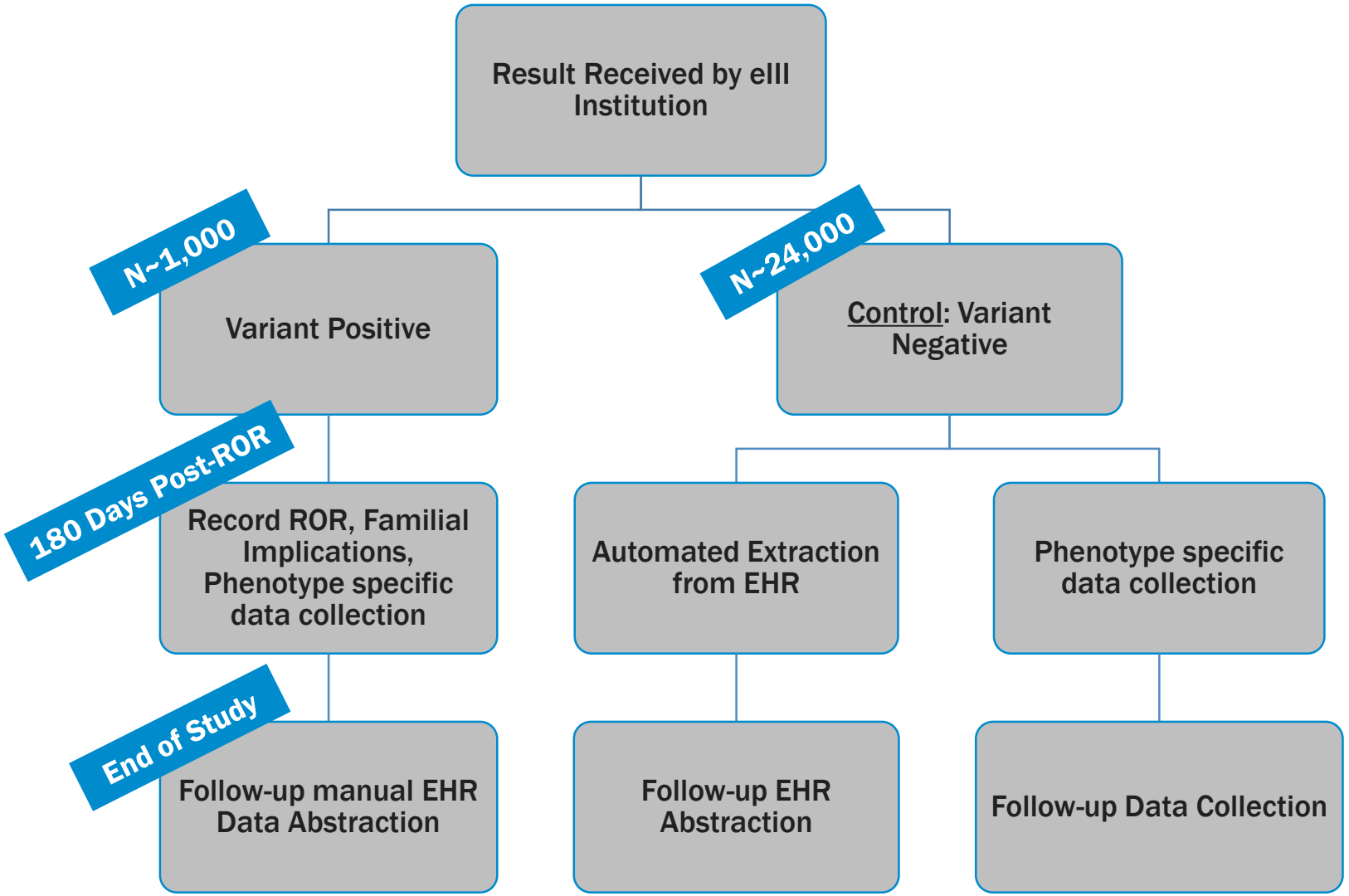
Prioritize Gene(s)-Outcomes Pairs

Define Specific Outcomes Projects

*e*MERGE OUTCOMES WORKGROUP: Future Efforts

Co-Chairs: Hakon Hakonarson (CHOP), Josh Peterson (Vanderbilt/CC), & Marc Williams (Geisinger)

Developing Workflow *for* Outcome Assessment



*e*MERGE PGx WORKGROUP: Status & Accomplishments

Co-Chairs: Laura Rasmussen-Torvik (Northwestern) & Cindy Prows (CCHMC)

- Monitoring PGx project progress
  - 8 Non-phenotyping papers in progress; 5 papers published

    https://docs.google.com/spreadsheets/d/1ewHXIm2NcaHOG85Ysfdc_GI39BSP9A08bI9gCzohRww/edit#gid=2059913263
  - 7 PGx-specific phenotypes and 6 combination (eII or eIII + PGx) phenotypes included on prioritization grid

    https://docs.google.com/spreadsheets/d/108r7-OOsCMp9pIHCXafzI4yhPA_iNCI2JYWfJ8rZe00/edit#gid=1638173722
- Grid created to monitor ongoing PGx clinical decision support activity

  https://docs.google.com/spreadsheets/d/1ewHXIm2NcaHOG85Ysfdc_GI39BSP9A08bI9gCzohRww/edit#gid=0
- Understand eI/II/III GWAS and PGx overlap

Site	PGx / eI-II-III Overlap
Marshfield	114
GHC/UW	177
Mayo	34
Northwestern	87
Geisinger	39
Mt Sinai	162
CCHMC	7
CHOP	515
Total	**1135**

eMERGE PGx WORKGROUP: Future Efforts

Co-Chairs: Laura Rasmussen-Torvik (Northwestern) & Cindy Prows (CCHMC)

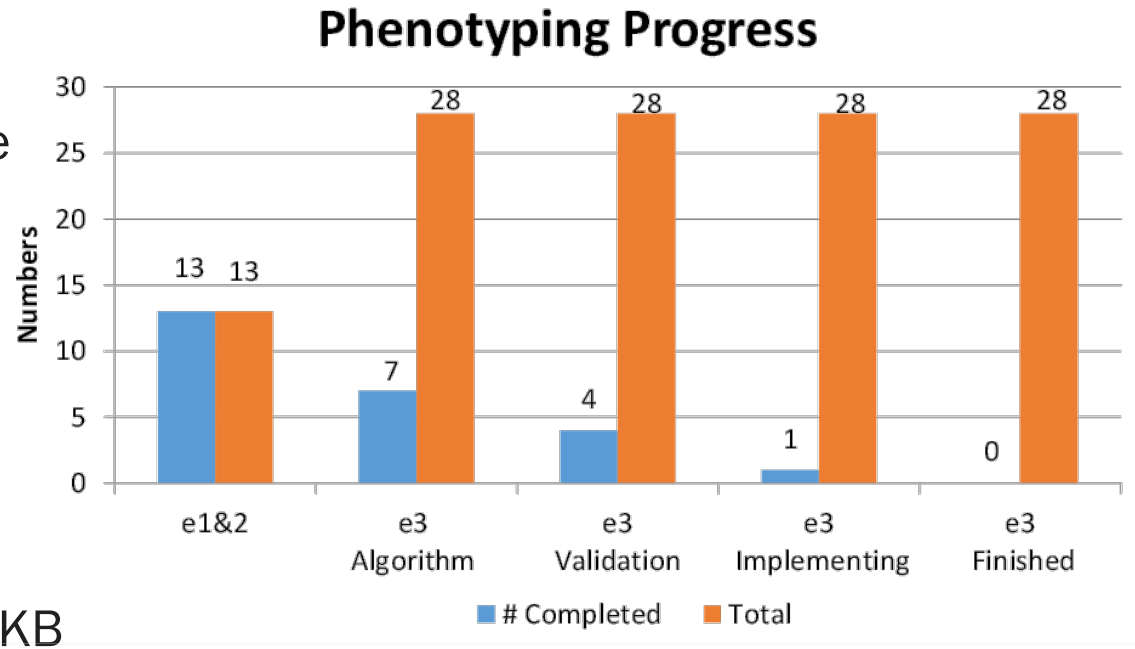
- Determine e3 PGx return of results processes
  - Baylor & LMM reports
    - .pdf and XML
  - Site specific return plans
    - Process for determining snvs to return
    - Plans for CDS?
    - Considerations for sites studying return of "negative results"
  - Overlap with existing projects
    - e3 / eMERGE-PGx
- Develop analysis plans with new PGx multisample call dataset
  - create eMERGE-approved rare variant analysis pathway on DNA Nexus and associated methods paper
- Examine cross-site PGx implementation process
- Compare *allele annotation tools for pharmacogenomic regions resistant to interpretation when using short-read sequencing technology (e.g. CYP2D6-CYP2D7; HLA)

*e*MERGE PHENOTYPING WORKGROUP: Status & Accomplishments

Co-Chairs: George Hripcsak (Columbia) & Peggy Peissig (Marshfield)

Phenotype Development & Implementation

- Reimplementation of 13 eI&II phenotypes – Complete
- eIII phenotypes
  - Schedule established
  - Algorithm completed – 25%
  - Validation completed – 14%
  - Implementing across network – 4%



Data Standardization Efforts

- PHeMA: phenotype authoring tool integrated into PheKB
- Cardio Core: narrative ECG and ECHO data into a coded repository
- Common data model: modular definition, OMOP sites, i2b2 middleware
- Consistent care definition: mature definition
- PheWAS codes: now include ICD10-CM and SNOMED

eMERGE PHENOTYPING WORKGROUP: Future Efforts

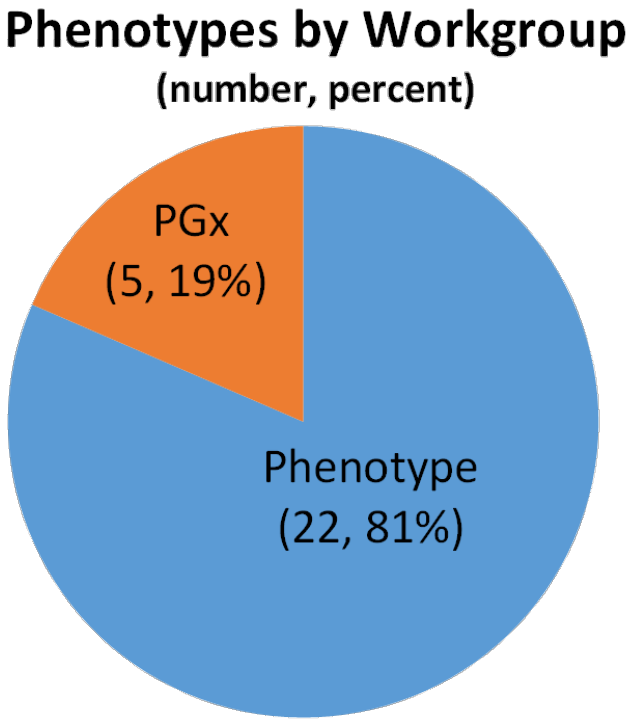
Co-Chairs: George Hripcsak (Columbia) & Peggy Peissig (Marshfield)

Phenotype Development & Implementation

- Stay on schedule
- Understand and meet needs of other workgroups
- Network data submissions
  - (dbGAP, SPHINX, eRecordCounter)

Data Standardization Efforts

- PHeMA: continue work
- Cardio Core: to be completed 2018
- NLP: sharing effort on parsing for breast cancer phenotype
- Common data model: carry out sharing experiment and publish
- Consistent care definition: validate and add to PheKB
- Develop a “common set” of data elements to use with all studies



eMERGE ROR/ELSI WORKGROUP: Status & Accomplishments

Co-Chairs: Ingrid Holm (BCH) & Iftikhar Kullo (Mayo)

- Project: Approaches to returning clinically actionable results at each site
  - Collected data and conduct interviews of investigators at all sites – Completed
- Project: ELSI impact of ROR on participants across the eMERGE sites: Develop data collection tools to implement across sites.
  - Participant survey subgroup – weekly calls
  - Domains: Baseline – decisional conflict; Post-disclosure 1 month and 6 month – Decisional regret, Privacy, Intent to/actual sharing with family, Impact of genetic findings
  - Baseline survey and 1 month post-disclosure survey questions to be included on most sites' surveys – Completed. 6 month post-disclosure survey question – In process
- Project: ELSI impact of disclosure of unsolicited genomic results on health care providers (HCPs) across eMERGE (see separate presentation)
- IRB Perspectives Project –experiences of eMERGE sites with respective IRBs related to approval of the return of unsolicited genetic results projects
  - Manuscript summarizing how IRBs at the various eMERGE sites reviewed each site's proposed return of results – in process
- Family history project - family communication supplement designed to understand how to contact family members
  - Geisinger Lead; Joint Outcomes/RoR project
- Joint meetings with the Outcomes and Clinical Annotation WGs to coordinate efforts

eMERGE ROR/ELSI WORKGROUP: Future Efforts

Co-Chairs: Ingrid Holm (BCH) & Iftikhar Kullo (Mayo)

- Describe approaches to returning clinically actionable results at each site
  - Manuscript outlining the ROR pipelines and common deviations from those pipelines at each site
- Develop and publish standards for ROR in eMERGE
  - Timeline: 1-2 years
- Complete comprehensive evaluation of impact of ROR on participants
  - Finalize questions for 6 month post-disclosure surveys – in process
  - Sites to implement surveys and include cross-site questions – in process
  - Analysis of data across sites to address hypotheses
- Best practice informed consent document for genome sequencing projects that involve ROR
- Cascade screening: explore best practices for contacting family members and methods to increase efficiency of cascade screening
- Coordinate joint projects with the Clinical Annotations and Outcomes WGs

*e*MERGE SUPPLEMENT PROJECTS

eMERGE GEOCODING SUPPLEMENT: Status & Accomplishments

Co-Chairs: Patrick Sleiman (CHOP) & Abel Kho (Northwestern)

- Two main challenges of this project: 1) Identifying the requisite environmental variables for extraction 2) Identify a platform that allows for centralized analysis of the data while respecting HIPAA rules surrounding the sharing of addresses (PHI)
- We have achieved network wide consensus on list of environmental variables to derive
- Prerequisites included: covariates for eMERGE phenotypes, both completed and in progress and availability of data with sufficient granularity across the US
- 15 environmental variables were selected. All variables have also been scored for importance by each site to generate a priority list. Socioeconomic status and hospital utilization scored most highly followed by a deprivation index and air quality
- We anticipate generating all variables on the list, however, to ensure the most desirable variables are generated we intend to begin the analysis at the top of the list and work down
- We have also identified a suitable platform to enable centralized data analysis using the same software, data tables and quality controls across the network
- DNAnexus provides a HIPAA compliant cloud based computation and storage solution on which we can run geocoding programs
- The Tibco Spotfire program has been successfully demoed on the platform

eMERGE GEOCODING SUPPLEMENT: Future Efforts

Co-Chairs: Patrick Sleiman (CHOP) & Abel Kho (Northwestern)

- Finalize preparation of DNAnexus as a platform includes, testing software, housing environmental data on the platform, setting up restricted access directories for each site to house addresses. All sites in the network are already DNAnexus customers with the necessary business agreements in place.

- Define and test a complete, end to end, analysis pipeline using representative test samples. Users will simply provide the addresses of their patients and the pipeline will generate the environmental exposures for those patients.

- Disseminate the SOP across the network.

- Deposit the demographic and exposure data generated by the geocoding in a central repository for use in genetic studies

eMERGE HEALTHCARE PROVIDER SUPPLEMENT: Status & Accomplishments

Chair: Ingrid Holm (BCH)

- Project: Impact of disclosure of unsolicited genomic results on health care providers (HCPs) across eMERGE 3
  - Ancillary Study Pilot Project funded by NHGRI ELSI Branch to develop survey
  - Plan is to obtain future funding to implement the survey across the eMERGE sites
  - Subgroup of the RoR-ELSI Workgroup: Vanderbilt clinical site (Clayton, Wiesner), Geisinger (Williams J), CCHMC (Myers), and Vanderbilt CC (Howell) with subcontracts to BCH (Holm), Louisville (Brothers), Colorado Children's Hospital (Ziniel)
- Literature review – completed
- Interviews of HCP across 4 institutions to inform survey
  - Adult and pediatric PCPs, oncologists, and cardiologists – 23 interviews completed
  - Coding – in progress
- Survey of HCP after receiving results on their patients
  - Draft informed by literature review and interviews – completed

eMERGE HEALTHCARE PROVIDER SUPPLEMENT: Future Efforts

Chair: Ingrid Holm (BCH)

- Interviews – complete coding; prepare and submit manuscript
- Surveys
  - Cognitive interviews
  - Finalize survey
- Submit RO1 for funding to start fall of 2017
  - Implement survey across eMERGE sites
  - Interview subset of HCPs 6 months after receiving results

MATERIALS *of* INTEREST

February 2017 Steering Committee Meeting Materials, Joint w/CSER

https://emerge.mc.vanderbilt.edu/february-2017-steering-committee-meeting/

October 2016 Steering Committee Meeting Materials

https://emerge.mc.vanderbilt.edu/3169-2/

Manuscripts *(to date)*

https://emerge.mc.vanderbilt.edu/publications/

Data Resources *(used to date)*

https://emerge.mc.vanderbilt.edu/tools/phenotype-data/

*e*MERGE TOOLS

eRC

https://biovu.vanderbilt.edu/EmergeRC/

PheKB

https://phekb.org/

CDSKB

https://cdskb.org/

SPHINX

https://www.emergesphinx.org/