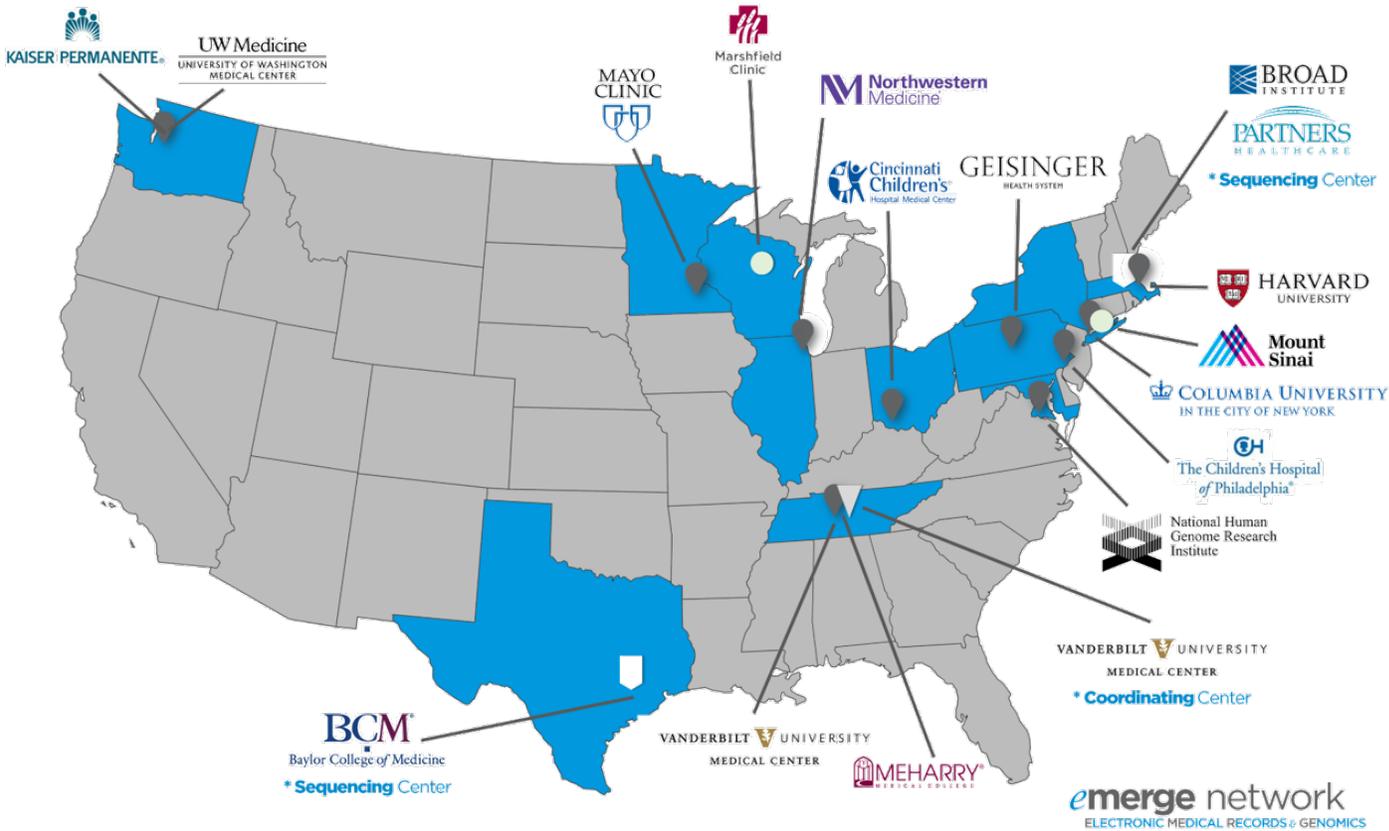


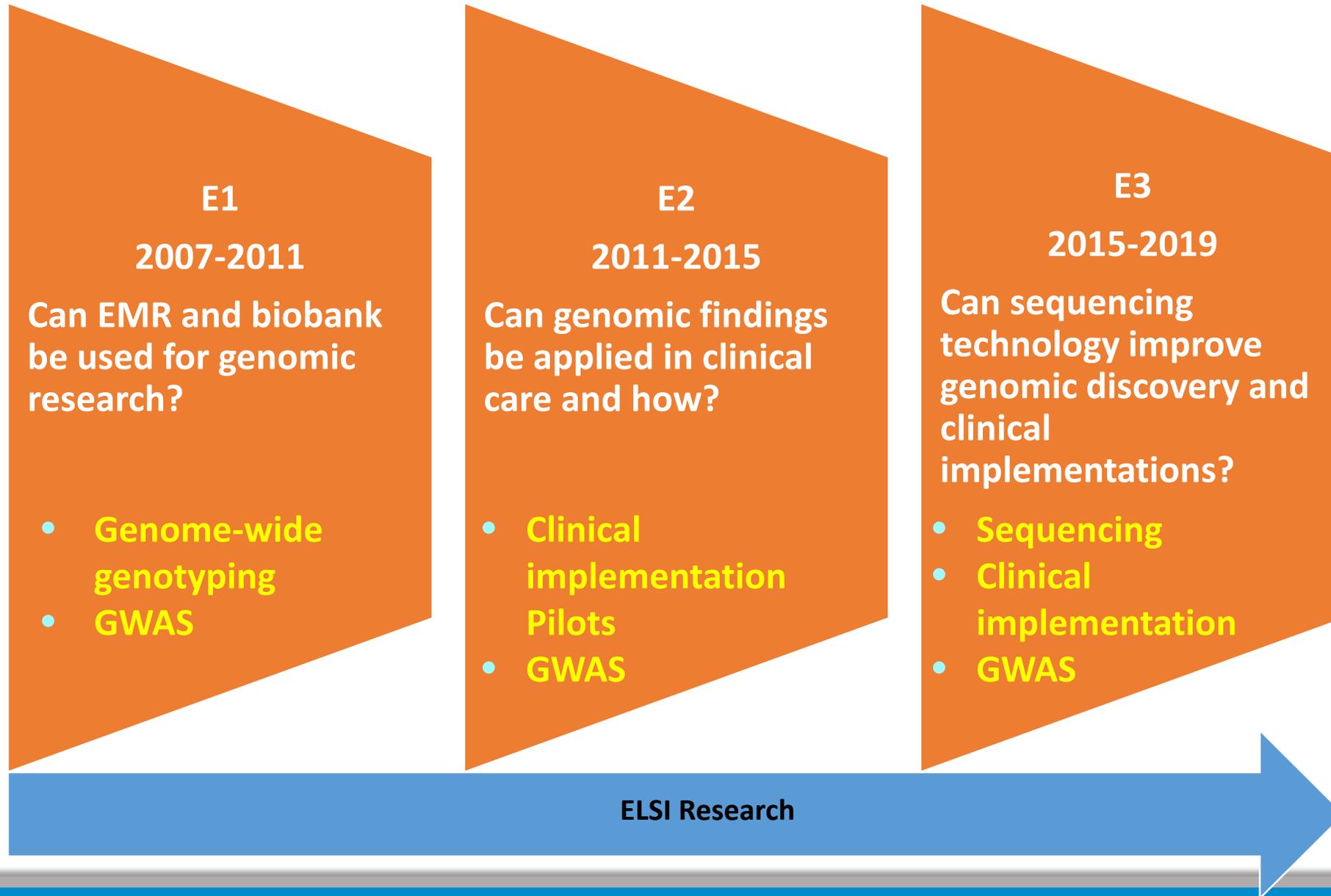
# The eMERGE Network



eMERGE WORKGROUPS	
Clinical Annotation	EHR Integration
Genomics	Outcomes
PGx	Phenotyping
Return of Results/ELSI	
eMERGE SUPPLEMENTS	
Geocoding	Health Care Provider Survey
Phenotyping – OMOP Model	
eMERGE SUBGROUPS	
Familial Implications of ROR	HLA
Infobutton	ROR Legal Considerations
Participant Survey	Phenotype Variables

**eMERGE PHASE III: SEPTEMBER 2015 – MAY 2019**

# eMERGE



# eMERGE III: What do we do?

## SPECIFIC AIMS *of the* eMERGE Network

- 1 Sequence and assess clinically relevant genes presumed to affect gene function in about 25,000 individuals
- 2 Assess the phenotypic implications of these variants
- 3 Integrate genetic variants into EMRs for clinical care
- 4 Create community resources

# Impact: 110k genomic dataset

- Data on over 110,000 participants and informatics tools with which to harness the data
- eMERGE Record Counter
  - Drag and drop demographics, phenotypes, ICD codes to obtain preliminary cohort counts
- SPHINX (Sequence and Phenotype Integration Exchange)<sup>+</sup>
  - Search catalog by genes, drugs, rsID and pathways
  - For each gene view: SNVs, pathways, drug interactions
  - For each variant view: SNPid, category, frequencies
  - European, African, & Asian ancestry allele data

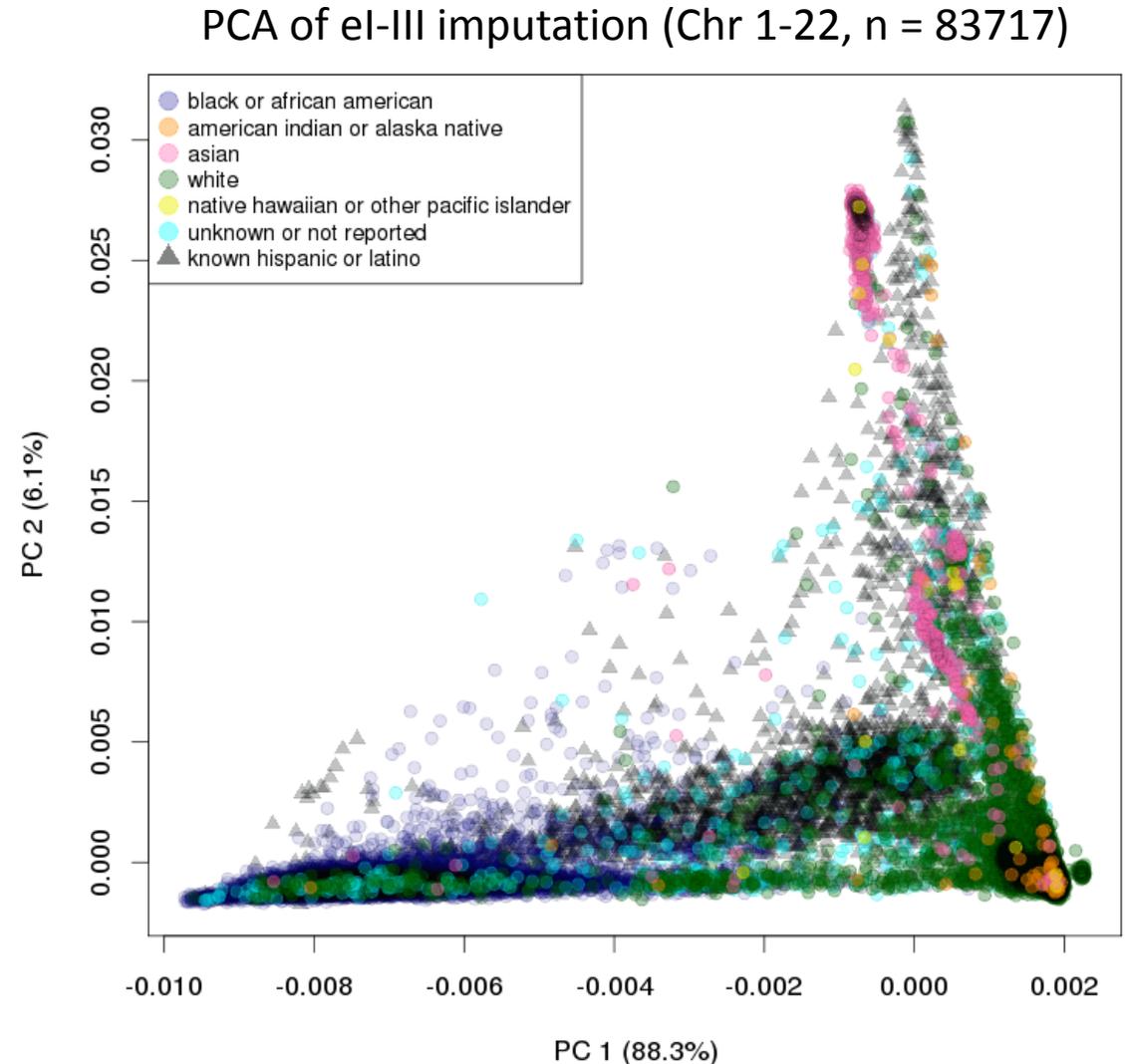
Set Name	Platform	Count
eI-eIII Merged*	GWAS	83,717
Exome chip	Exome	12,330
Whole exome	Sequencing	3,745
PGx	Sequencing	9,010
Whole genome	Sequencing	1,800
eMERGEseq	Sequencing	25,000
<b>Total Current</b>		<b>111,078</b>
<b>Total Expected</b>		<b>136,078</b>

\*eI-II: 55,029 samples; and eIII: 28,688 samples

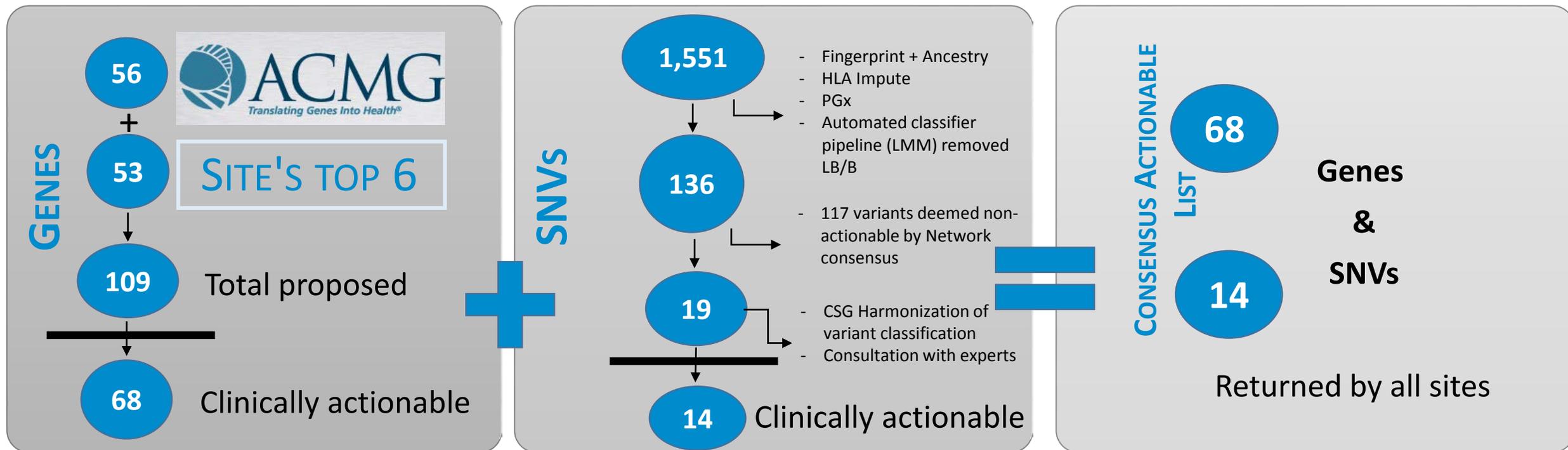
<sup>+</sup>Rasmussen-Torvik LJ, Stallings SC, Gordon AS, Almoguera B, Basford MA, et al. *Design and Anticipated Outcomes of the eMERGE-PGx Project: A Multicenter Pilot for Preemptive Pharmacogenomics in Electronic Health Record Systems*. *Clinical Pharmacology & Therapeutics*. 2014 Oct; 96(4):482-9. PMID: 24960519 PMCID: PMC4169732

# Deliverable: Imputation and merging of eI-III GWAS data

- Imputation of all eMERGE array data against the HRC reference using the Michigan Imputation Server
- HRC reference contains 39,235,157 SNPs, no indels, provides access to rare variation (low as 0.1%)
- 83,717 individuals in data set released to network
  - Principle components analysis (PCA) examined ancestry



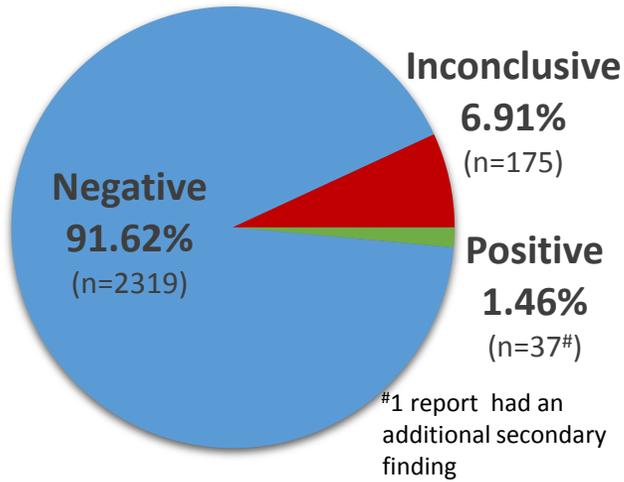
# Deliverable: Development of an eMERGEseq Platform



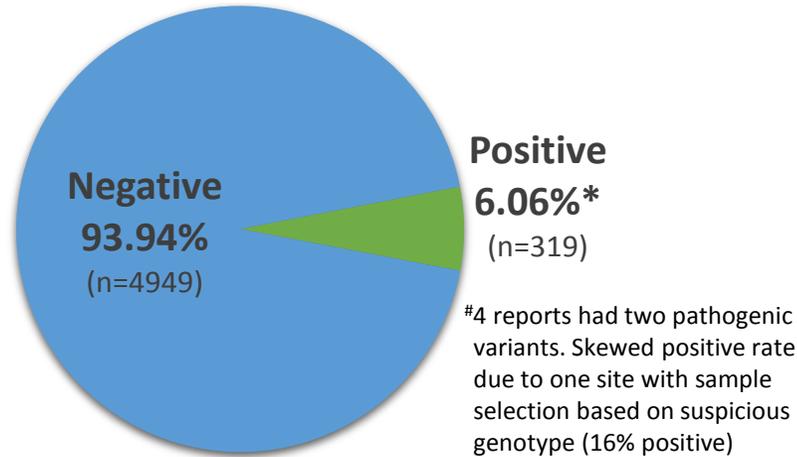
- Clinical reports are generated on the “Consensus Actionable List” and any specific genes or SNVs requested by individual sites
- *To date:* 14,077 samples sequenced and 3,716 reports issued

# Partners-Broad Interpretation and Reporting: Review of 5268 cases

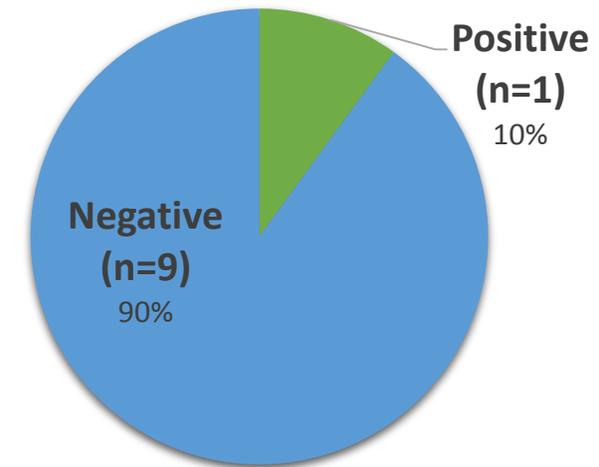
## Indication-based returnable results (n=2531)



## Non indication-based consensus returnable results\*\* (n=5268)



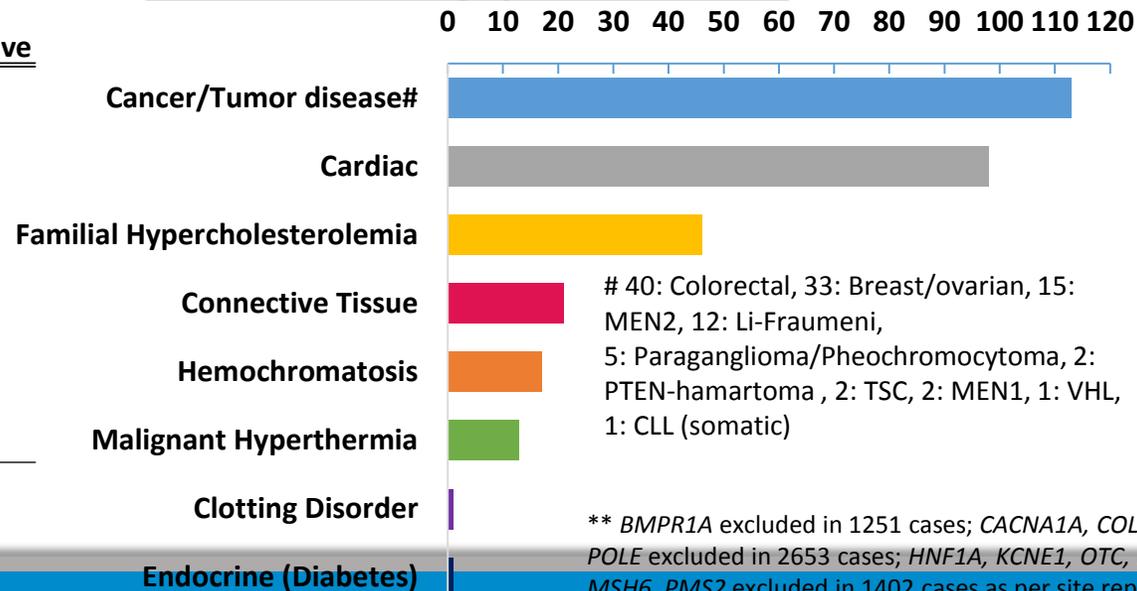
## Non indication-based site-specific returnable results (n=10)



Indication	Total	Positive	Negative	Inconclusive
Colorectal cancer/Polyps	1165	24*	966	175
Ehlers-Danlos Syndrome	66	1	65	n/a
Abnormality of pain sensation	545	0	545	n/a
Pediatric migraine	443	0	443	n/a
Hyperlipidemia	258	12	246	n/a
Autistic Behavior	54	0	54	n/a
<b>TOTAL</b>	<b>2531</b>	<b>13</b>	<b>2319</b>	<b>175</b>

\* 1 report had an additional secondary finding

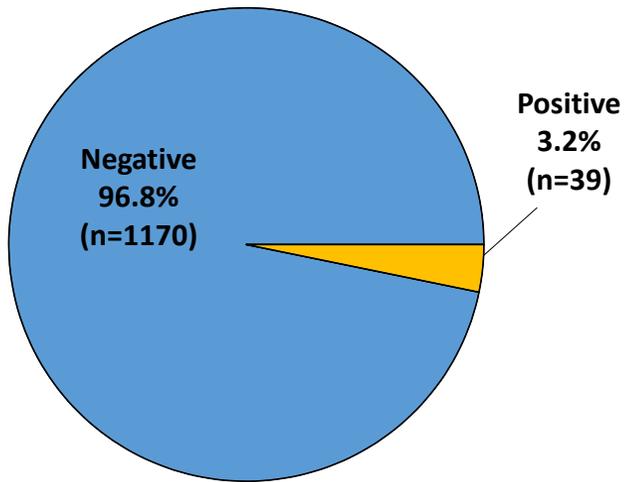
### Returnable findings per disease area



Path/Likely Path Variants in CCHMC-adolescent cohort specific genes	Total
<i>CHEK2</i>	1/10

# Interpretation & reporting: Baylor

## Indication based returnable results



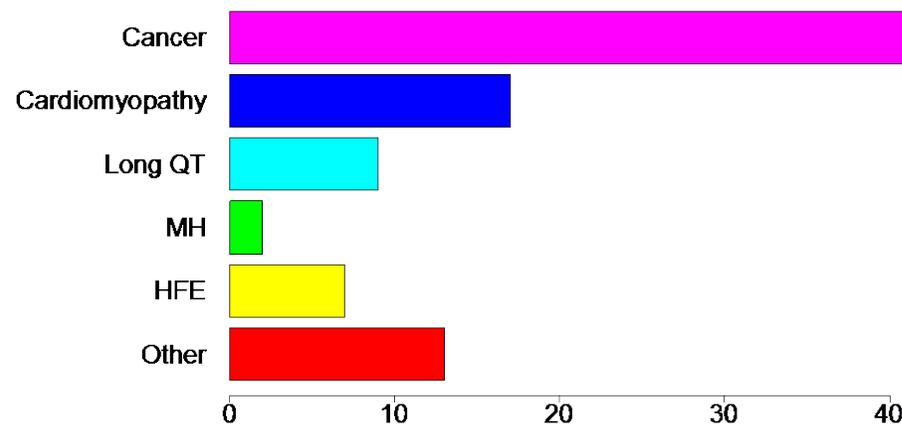
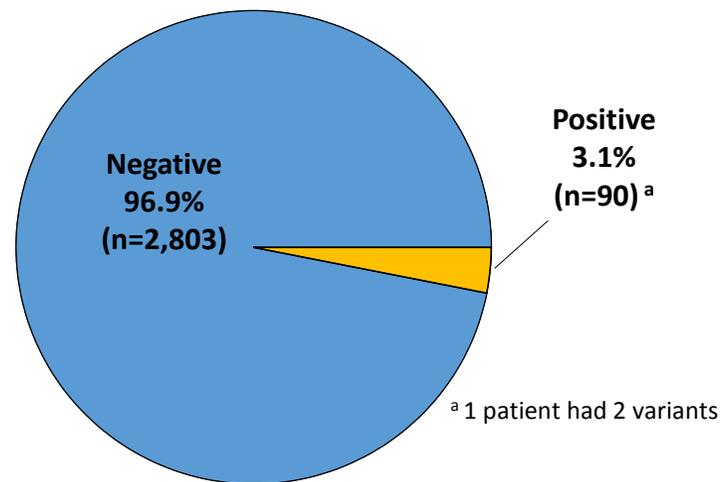
Indications	Total	Positive	Negative
Cardiomyopathy	1	1	0
Cardiac Arrythmia	31	0	31
Hyperlipidemia <sup>a, b</sup>	808	22	786
Colorectal Cancer	595	3	592
Breast/Ovarian Cancer <sup>c</sup>	72	16	56

<sup>a</sup> 298 patients had colorectal cancer and hyperlipidemia

<sup>b</sup> Hyperlipidemia includes FH, hypertriglyceridemia, hyperlipidemia and coronary artery disease indications.

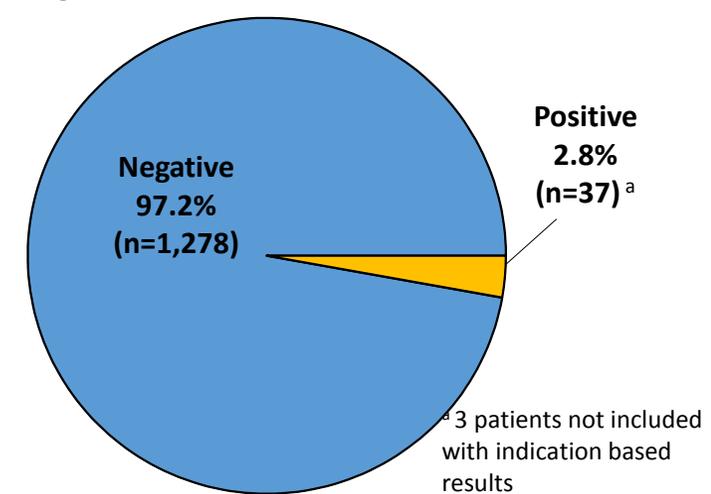
<sup>c</sup> All returned genes belong to the 68 consensus except for CHEK2 in a breast cancer patient

## Non indication based consensus returnable results



Others include MEFV, HNF1A, CACNA1A, OTC, LDLR

## Non indication based site-specific returnable results



Path and Lpath variants in NU and Vanderbilt specific returned	Total
<i>CHEK2</i>	24
<i>ATM</i>	7
<i>SERPINA1</i>	3
<i>MC4R</i>	3
<i>F11, FLG, KCNE2 (x1)</i>	3

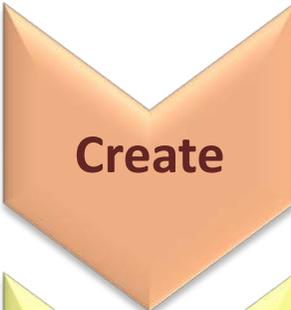
# Impact: Electronic phenotyping & PheKB

- PheKB (Phenotype KnowledgeBase)
  - Collaborative environment to building and validating electronic algorithms
  - Computational algorithm library
    - 37 finalized, public phenotypes
- Demonstrated feasibility of use in Genomic Medicine
- Tools and process allowed for computational and algorithm development cross collaboration around the world

Kirby JC, Speltz P, Rasmussen LV, Basford M, Gottesman O, et al. *PheKB: a catalog and workflow for creating electronic phenotype algorithms for transportability*. J Am Med Inform Assoc. 2016 Nov;23(6):1046-1052. doi: 10.1093/jamia/ocv202. PMID: 27026615  
PMCID: PMC5070514.

# Phenotype Development Workflow

## Tool Support



- Phenotype algorithm and data dictionary are in development
  - ❖ Share algorithm with project team
  - ❖ Standardize Phenotype Development
  - ❖ Standardize data collection



Logos for KNIME, eMERGE RecordCounter, eleMAP, and PheKB.



- Algorithm and Data Dictionary in review by validation site(s)
  - ❖ Share algorithm with validation team
  - ❖ Validate algorithm
  - ❖ Validate Data Dictionary



Logos for Dictionary/Dataset Validation and PheKB.



- Share and implement algorithm and data dictionary for multi-site data collection
  - ❖ Validate Dataset against Data dictionary



Logos for Dictionary/Dataset Validation, eMERGE RecordCounter, and PheKB.



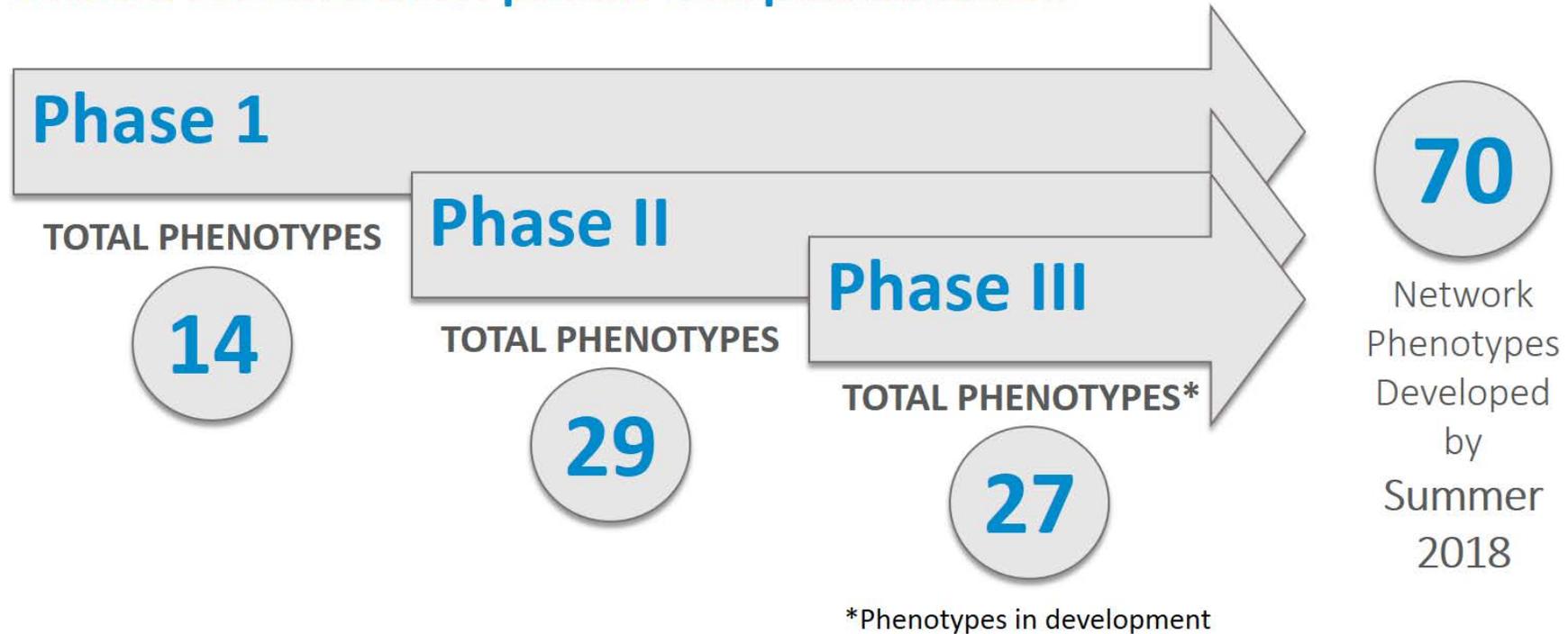
- Phenotype published and Algorithm is sharable to public



Logo for PheKB.

# Phenotypes

## PHENOTYPES: Development & Implementation

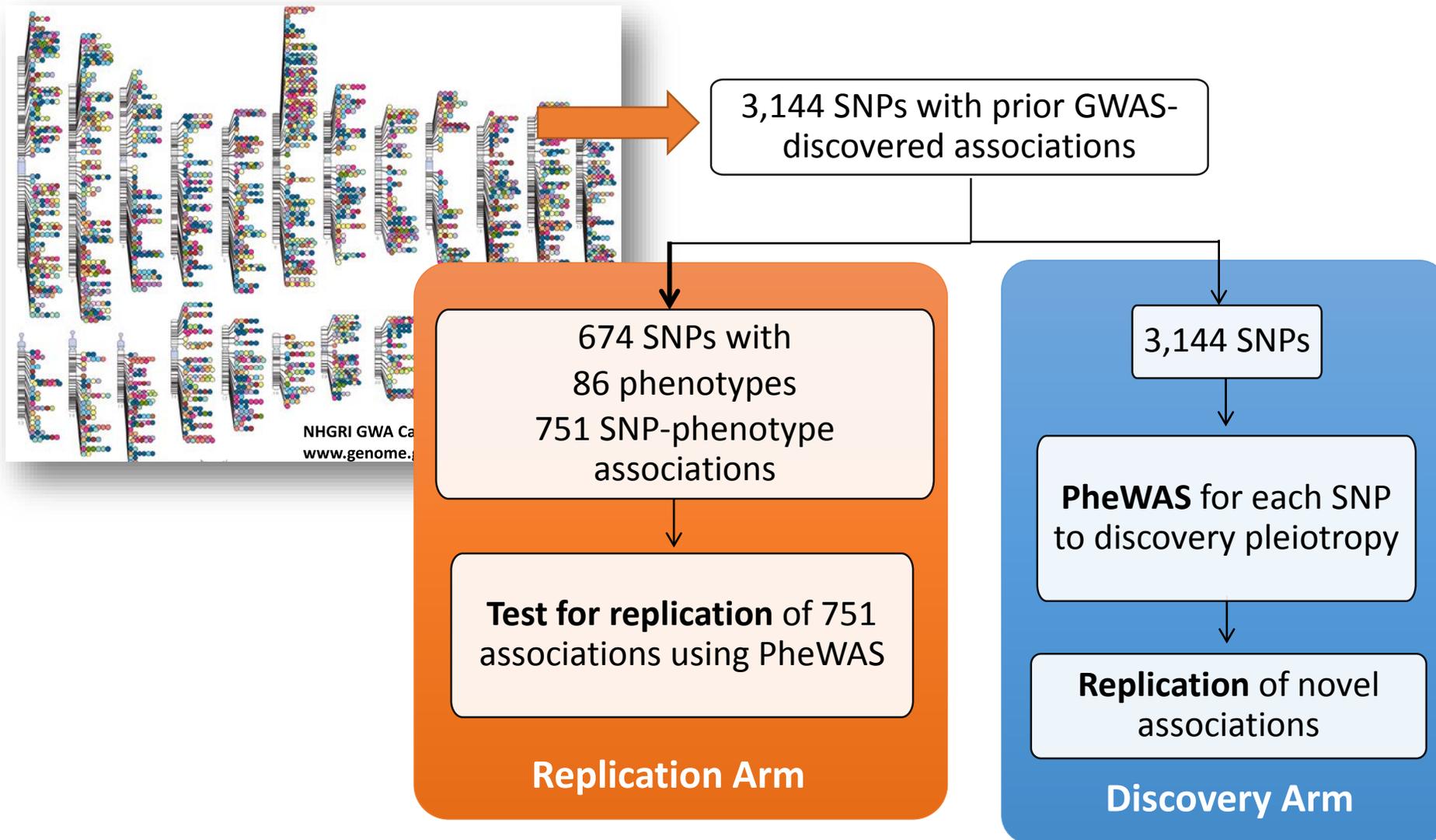


## Impact: eMERGE PheWAS

- Developed methods for large scale genotype/phenotype analyses and implemented them across an entire collaborative Network
- Phenome-wide association studies (PheWAS)
  - 3144 SNPS present in NHGRI catalog (2012) in 13,835 individuals across 5 sites.
    - 1358 phenotypes analyzed for each SNP
  - Addition of Neanderthal PheWAS catalogue
  - Creation of Phecode mappings from ICD codes

Denny JC, Bastarache L, Ritchie MD et al. *Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data*. Nat Biotechnol. 2013 Dec;31(12):1102-10. PMID: 24270849 PMCID: PMC3969265

# PheWAS of “all” NHGRI GWAS Catalog SNPs



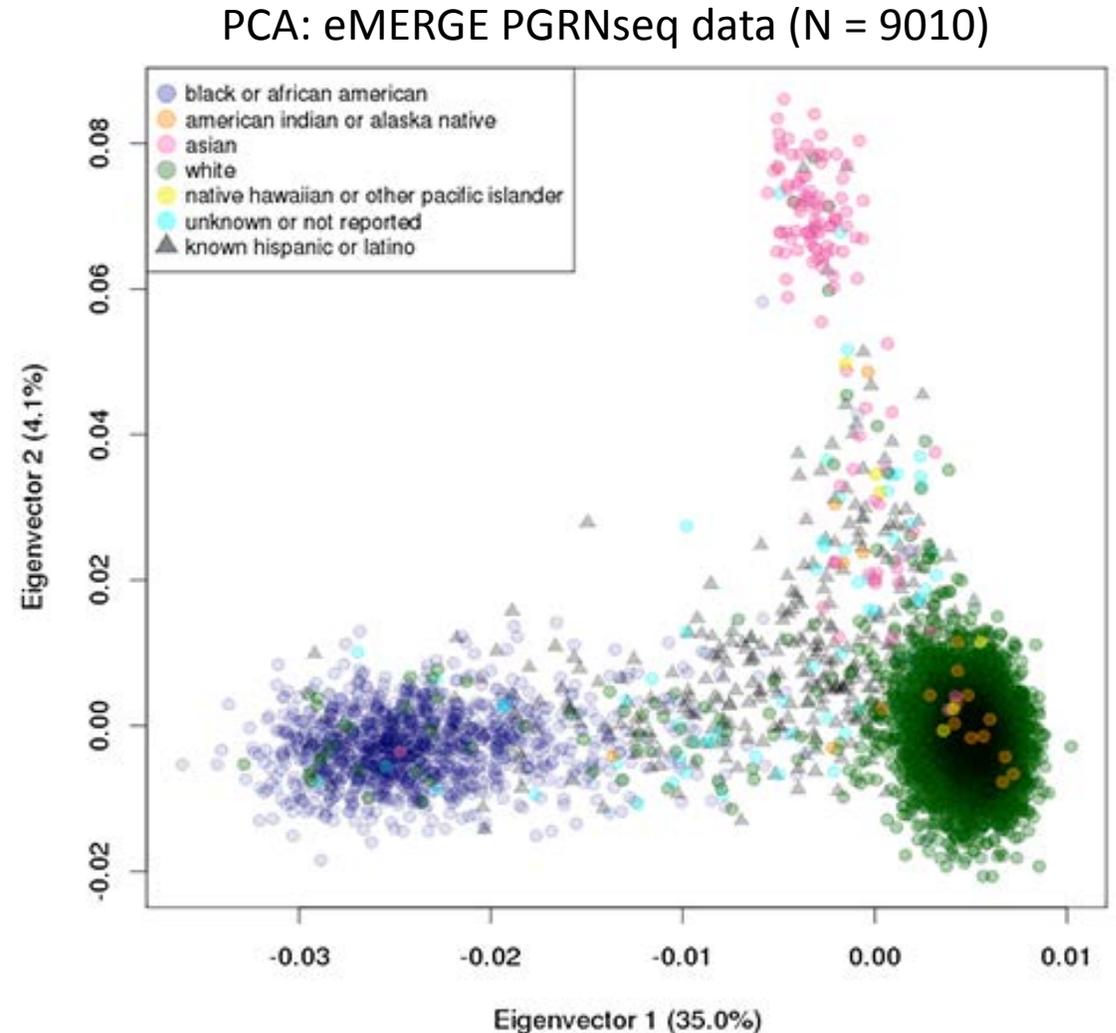
## Impact: eMERGE Pharmacogenomics (PGx)

- Multi-site test of the concept that genetic sequence information can be coupled to electronic medical records (EMRs) for use in healthcare
- Genetic sequencing on a 9010 participant data set
  - Sequencing and phenotype data available on SPHINX
- 82 pharmacogenetic genes investigated
- Many more opportunities for research on these data
  - PGx SNVs on the eMERGE-Seq panel
- Sites continue to collect utilization and outcomes data

Bush WS, Crosslin DR, Owusu-Obeng A, Wallace J, Almoguera B. et al. *Genetic variation among 82 pharmacogenes: The PGRNseq data from the eMERGE network. Clin Pharmacol Ther. 2016 Aug;100(2):160-9. doi: 10.1002/cpt.350. PMID: 26857349 PMCID: PMC5010878.*

# Deliverable: PGRNseq multi-sample calling

- Original PGRNseq aligned to multiple references used by the original five sequencing centers
- All 9010 BAMs re-aligned to the same genome reference hs37d5.fa
- 9010 individuals in data set provided to the network for analysis
- Principle components analysis (PCA) examined ancestry



# Impact: Return of genomic data via EMR

- Infrastructure and tools, in particular decision support tools, to enable genomic medicine
- InfoButton\*
  - Explored use of infobuttons as a decision support tool to provide context specific links within the electronic health record (EHR) to relevant genomic medicine content
  - Assessed the coverage of content topics among information resources developed
- CDS\_KB (Clinical Decision Support KnowledgeBase)
  - Partnership with IGNITE network
  - Goal is to catalog and share CDS implementation artifacts and design considerations for genomic medicine programs from a broad community of institutions

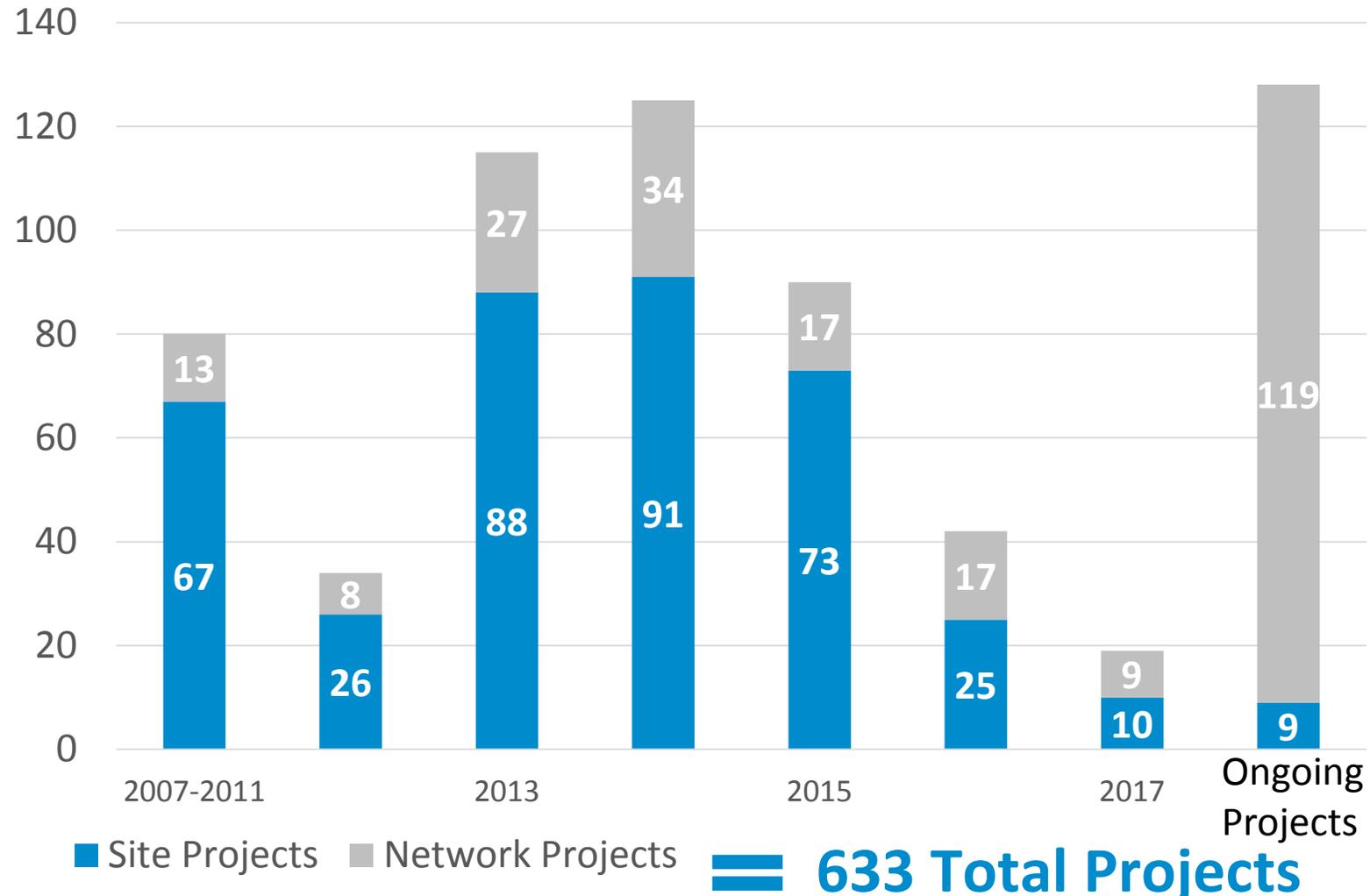
\*(Overby CL, Rasmussen LV, Hartzler A, Connolly JJ, Peterson JF, et al. *A Template for Authoring and Adapting Genomic Medicine Content in the eMERGE Infobutton Project*. AMIA Annu Symp Proc. 2014 Nov 14;2014:944-53. PMID: 25954402 PMCID: PMC4419923.)

## Impact: Network wide analyses 'DNAnexus'

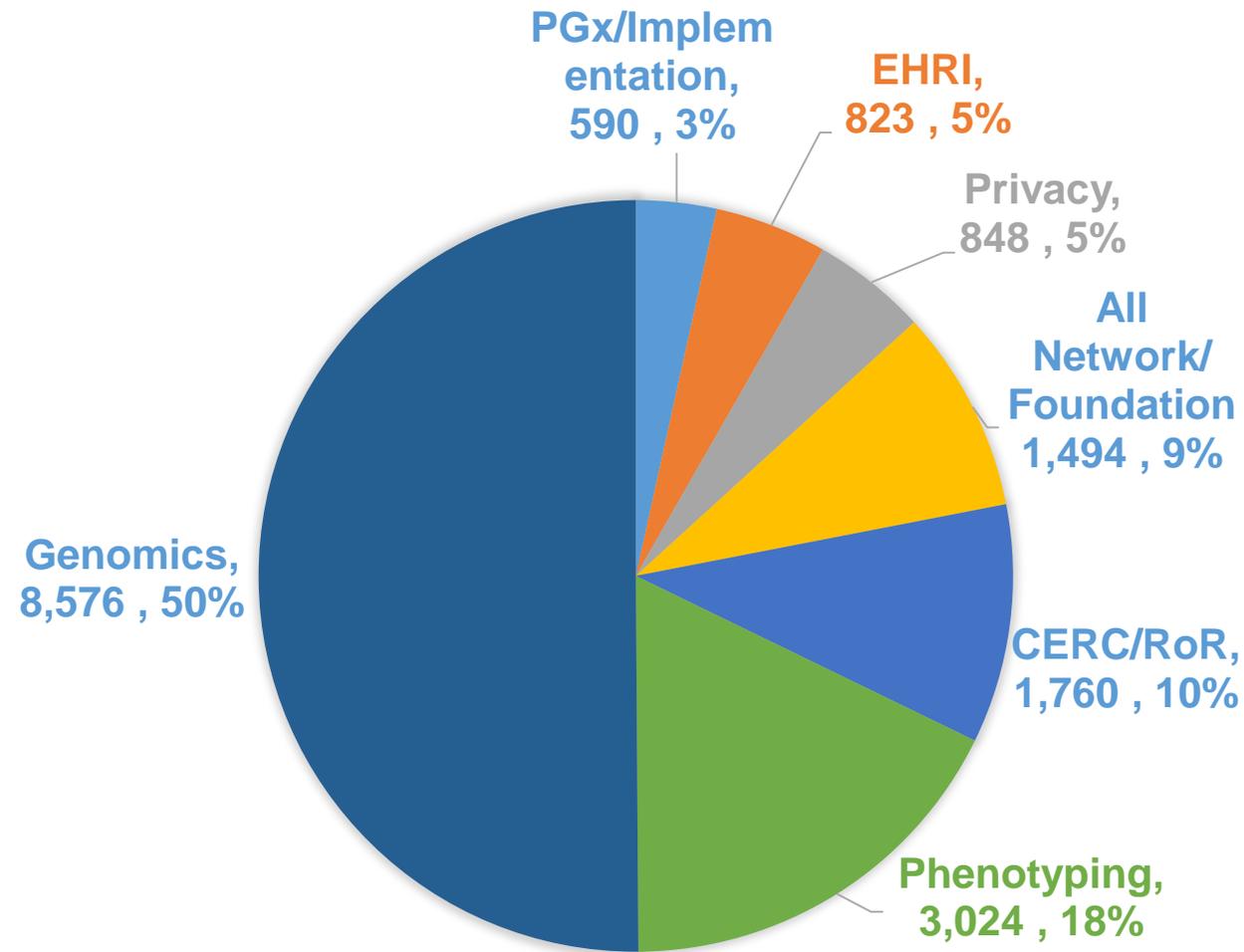
- Utilization for sharing and managing genetic data in a cloud-based system
- Network seminar series demonstrating utility of the analysis pipeline and development of apps
  - Large scale analyses possible for all investigators, regardless of local computing power
- DNAnexus houses Network wide genetic datasets
  - GWAS
    - Including a subset of geocoded samples
  - PGRNseq
  - eMERGEseq

# Impact: eMERGE Publications 2007-2017

## Number of Published Projects Through August 2017



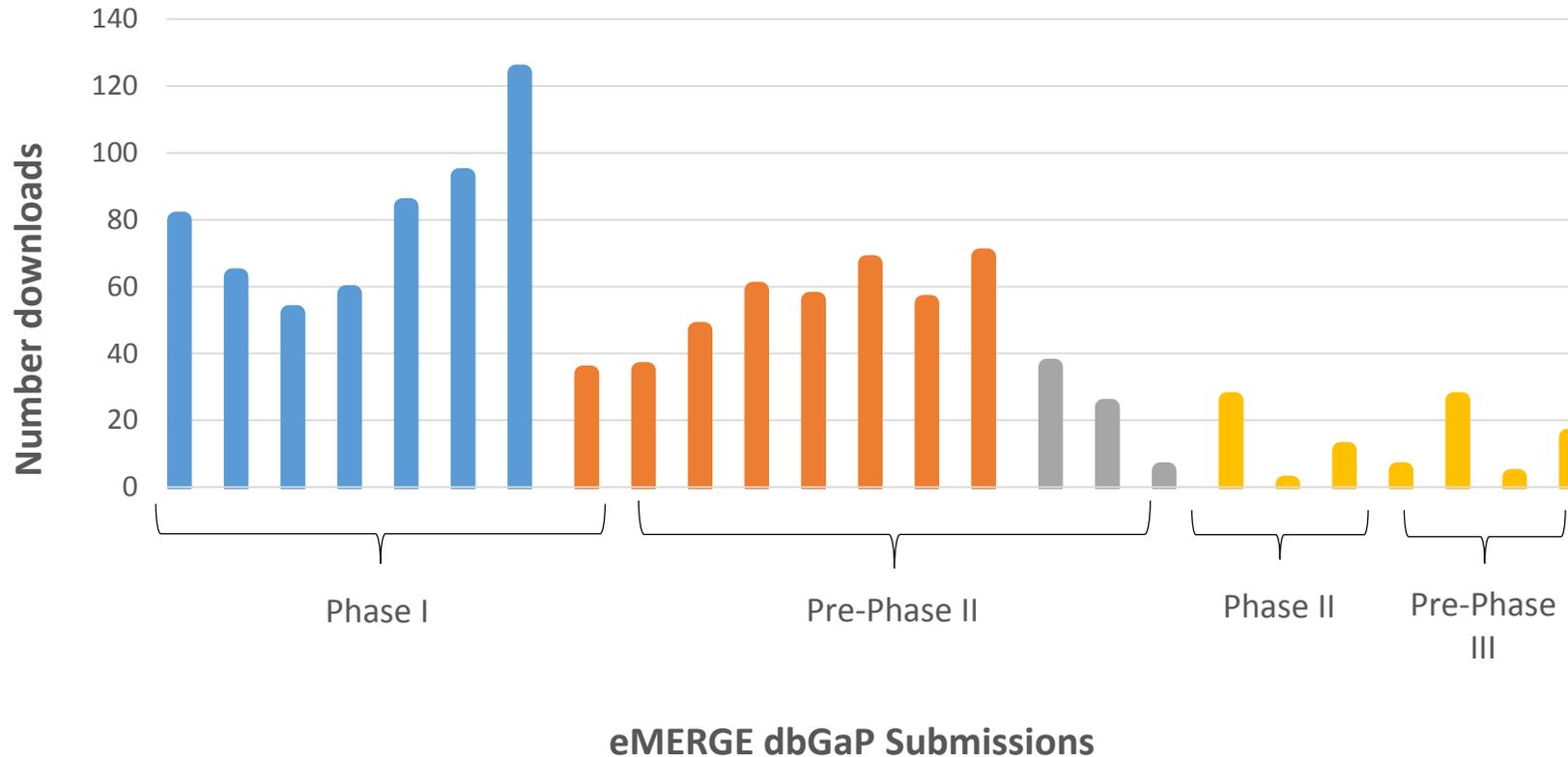
# Citations of eMERGE Publications by Category



Cumulative Citation Counts: 17,115 (2007-March 2017)

# Impact: dbGaP & Website analytics 2007-2017

Data Reuse: # Downloads of eMERGE dbGaP Submissions as of August 2017



> **1100** external downloads as of August 2017

**emerge network**  
 ELECTRONIC MEDICAL RECORDS & GENOMICS

## eMERGE Website

**Average usage past 6 months**

- 63.1% new visitors
- 1596 sessions/month
- 1043 users/month
- Views from 96 countries

**PheKB**  
 a knowledgebase for discovering phenotypes from electronic medical records

## PheKB Website

**Average usage past 6 months**

- 56.2% new visitors
- 1171 sessions/month
- 540 users/month
- Views from 76 countries

# eMERGE Tools

**PheKB**

A **knowledgebase** for discovering phenotypes from electronic medical records

**MyResults.org**

An **informational tool** for educating patients about genetic test results

**SPHINX**

A **data exploration tool** for genetics-related drug response hypothesis generation

**Infobutton Project**  
template

*e*merge **Model Consent Language**

**PheWAS**  
catalog

## Additional Tools

**GENOTYPING** tools

**PHENOTYPING** tools

**CDS** tools

Natural Language Processing (NLP) Tools

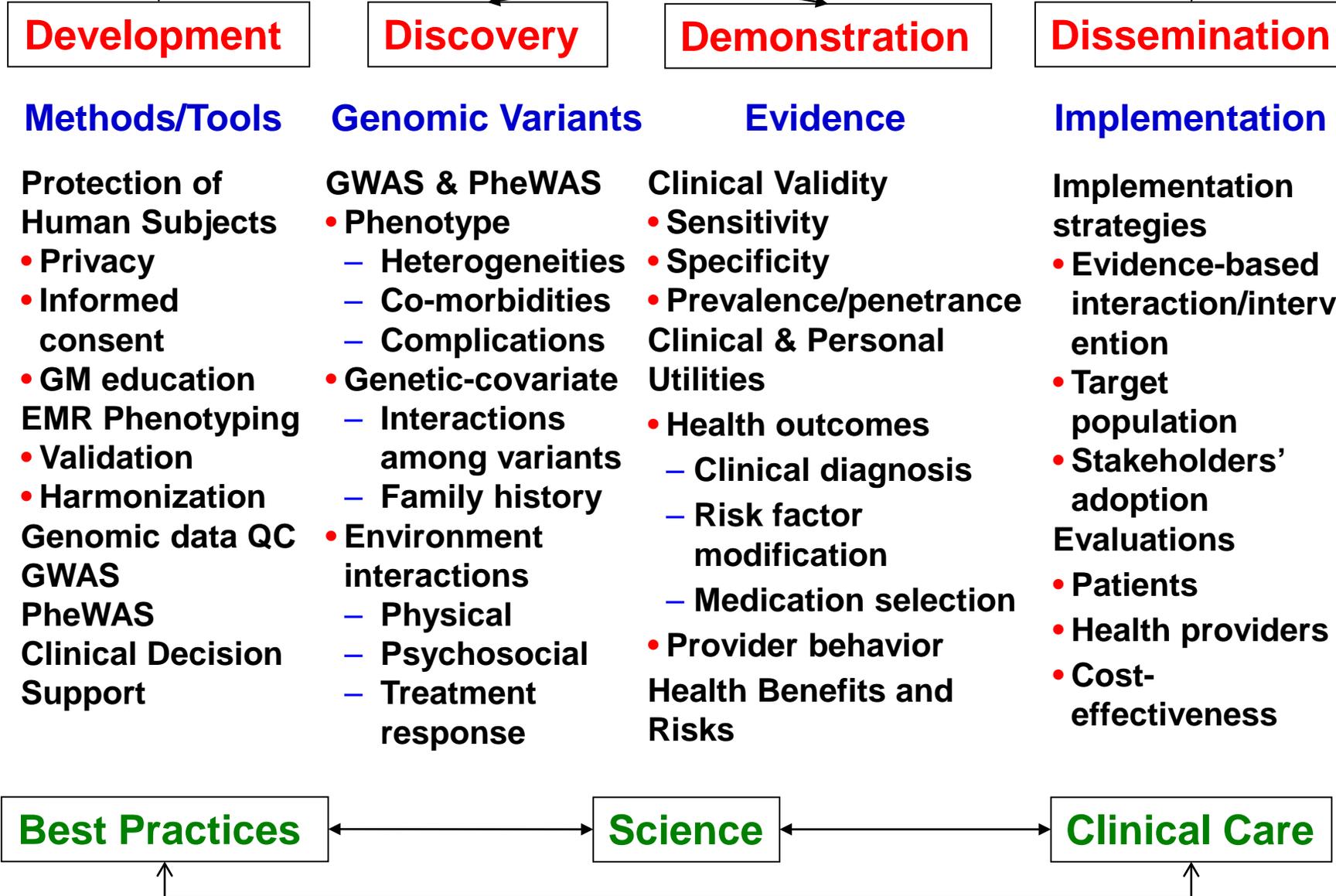
# eMERGE III: Future Deliverables

- dbGaP submissions
  - GWAS eI-III imputed set (*ready to submit*)
  - Interim (Fall 2017) and final eMERGEseq data
- Return of clinical results and EHR integration at all sites
  - Establishment of IT support for return of results processes based on data delivered through the network
  - Analysis of solutions, challenges and lessons learned
    - Manuscripts and methods documentation of Network-wide efforts
    - Sharing with CDSKB and standard bodies as appropriate
- Outcomes analysis for effect of return of results on patients and providers across sites
  - Compare differences in health outcomes and provider behaviors for return of negative and positive results
  - EHR and survey based methods for examining patient impacts and changes in care or awareness by providers
- Creation and deployment of 27 phenotypes, 8 deployed to date
  - 25 eI-III imputed GWAS
  - 13 PGRNseq
  - 24 eMERGEseq

Questions??



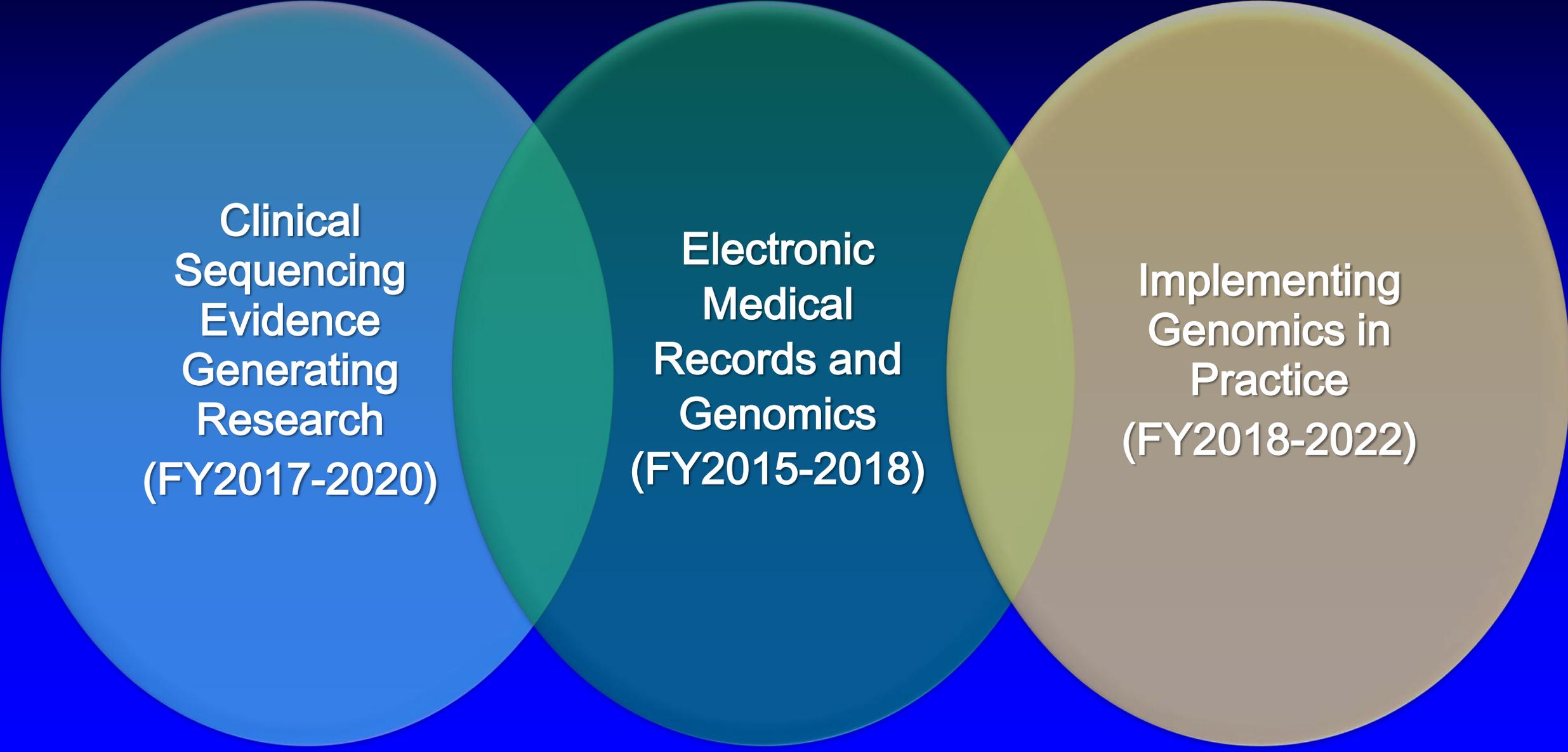
# eMERGE



# eMERGE Geocoding

Factors	Source	Resolution	National/ Local
Demographics	Coordinating Center/Site EDW	Patient Level	National
SES	Census/ACS	Block Group Level	National
Built Environment	RUCA (rural-urban-commuting-area-codes)	Tract Level	National
Traffic Volume	Google?		
Road Density	ArcGIS shapefiles	Block Group Level	National
Food Accessibility	Food Environment Atlas (USDA Economic Research Service)	County Level	National
Water Quality	NURE-HSSR database; Enviromapper?	Various	
Density of Parks	ArcGIS shapefiles	Block Group Level	National
Walkability	Walk Score Professional	Zip Code	National
Entropy Index	Census/ACS	Block Group Level	National
Crime			Local
Hospital Utilization	AHRF, HHS, HRSA	County Level	National

# CSER, eMERGE, and IGNITE



Clinical  
Sequencing  
Evidence  
Generating  
Research  
(FY2017-2020)

Electronic  
Medical  
Records and  
Genomics  
(FY2015-2018)

Implementing  
Genomics in  
Practice  
(FY2018-2022)

# Commonalities and Complementarity of CSER and eMERGE

## CSER (FY2017-2020)

- ~4,600 pts, 6 sites
- Community clinical scenarios
- Focus: clinical encounter
- Increased ethnic and socioeconomic diversity
- Evidence generation for clinical utility of genomic sequencing
- Real-world barriers to integrating genomic data for healthcare utilization

- EMR integration
- Clinical impact of RoR
- Data sharing concerns

## eMERGE (FY2015-2018)

- 25K pts, 9 sites
- Electronic phenotyping
- Focus: system-wide
- Health outcomes of rare variants in ~100 clinically relevant genes
- System-wide impact of reporting actionable variants
- Improved e-phenotyping
- Novel variant discovery
- Electronic CDS

# Commonalities and Complementarity of eMERGE and IGNITE

## eMERGE (FY2015-2018)

- 25K pts, 9 sites
- Electronic phenotyping
- Focus: system-wide
- Health outcomes of rare variants in ~100 clinically relevant genes
- System-wide impact of reporting actionable variants
- Improved e-phenotyping
- Novel variant discovery
- Electronic CDS

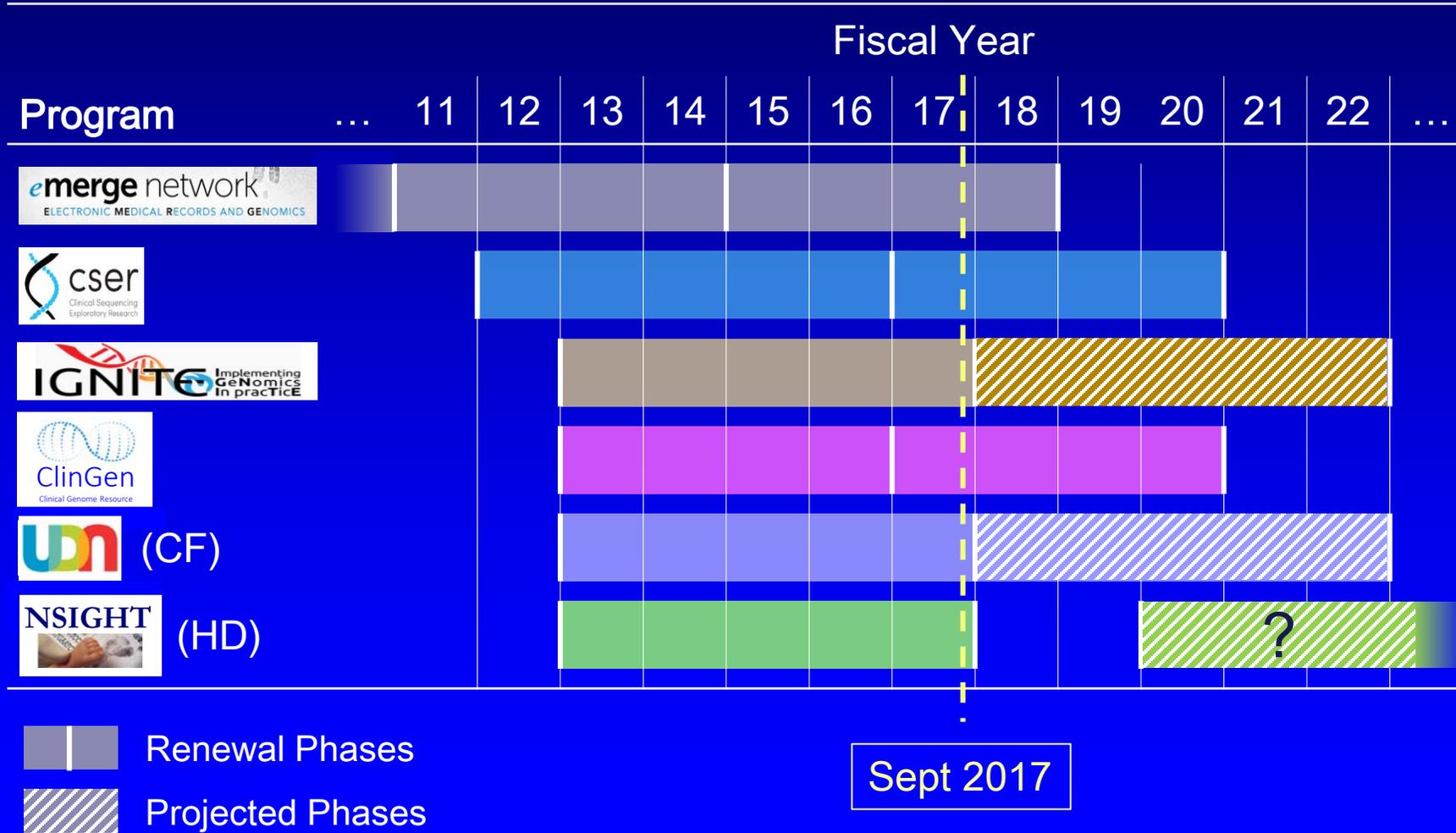
- EMR integration
- Cost-effectiveness
- Patient/clinician education

## IGNITE (FY2018-2022)

- ~15K pts, 4-6 sites
- Diverse, real-world clinical settings
- Focus: pragmatic trials
- Clinical utility of established genomic medicine interventions
- Increased ethnic and socioeconomic diversity
- Generalizable knowledge on use of trials in genomic medicine interventions

# Timeline of NHGRI Genomic Medicine Programs

You are here



# Resources



[www.gwas.org](http://www.gwas.org)

## Manuscripts *(to date)*

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



**dbGaP**

*(published to date)*

<https://emerge.mc.vanderbilt.edu/dbgap/>

## GWAS sequencing platforms *(eI-III)*

<https://emerge.mc.vanderbilt.edu/wp-content/uploads/2015/02/Platform-Information-eMERGE.docx>

# TOOLS



<https://phewascatalog.org/>



<https://phekb.org/>



<https://cdskb.org/>



<https://www.emergesphinx.org/>