VANDERBILT UNIVERSITY
MEDICAL CENTER

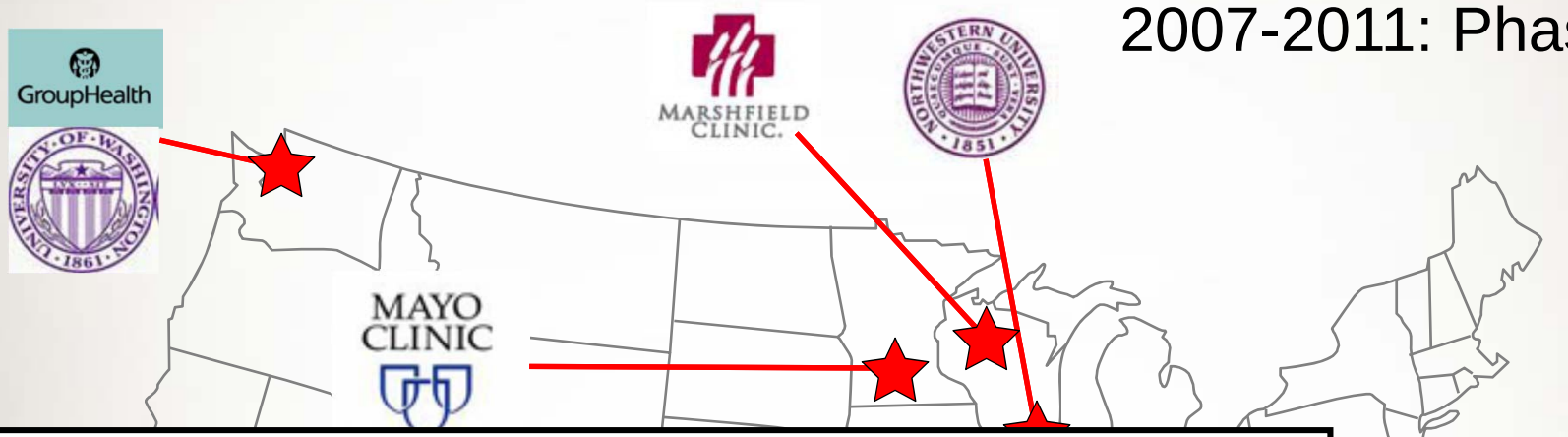
The electronic MEdical Records and GEnomics (eMERGE) network

The eMERGE Network

electronic Medical Records & Genomics

A consortium of biorepositories linked to electronic medical records data for conducting genomic studies

2007-2011: Phase I



eMERGE-I goal: to assess utility of DNA collections integrated with electronic medical records (EMRs) as resources for genome science

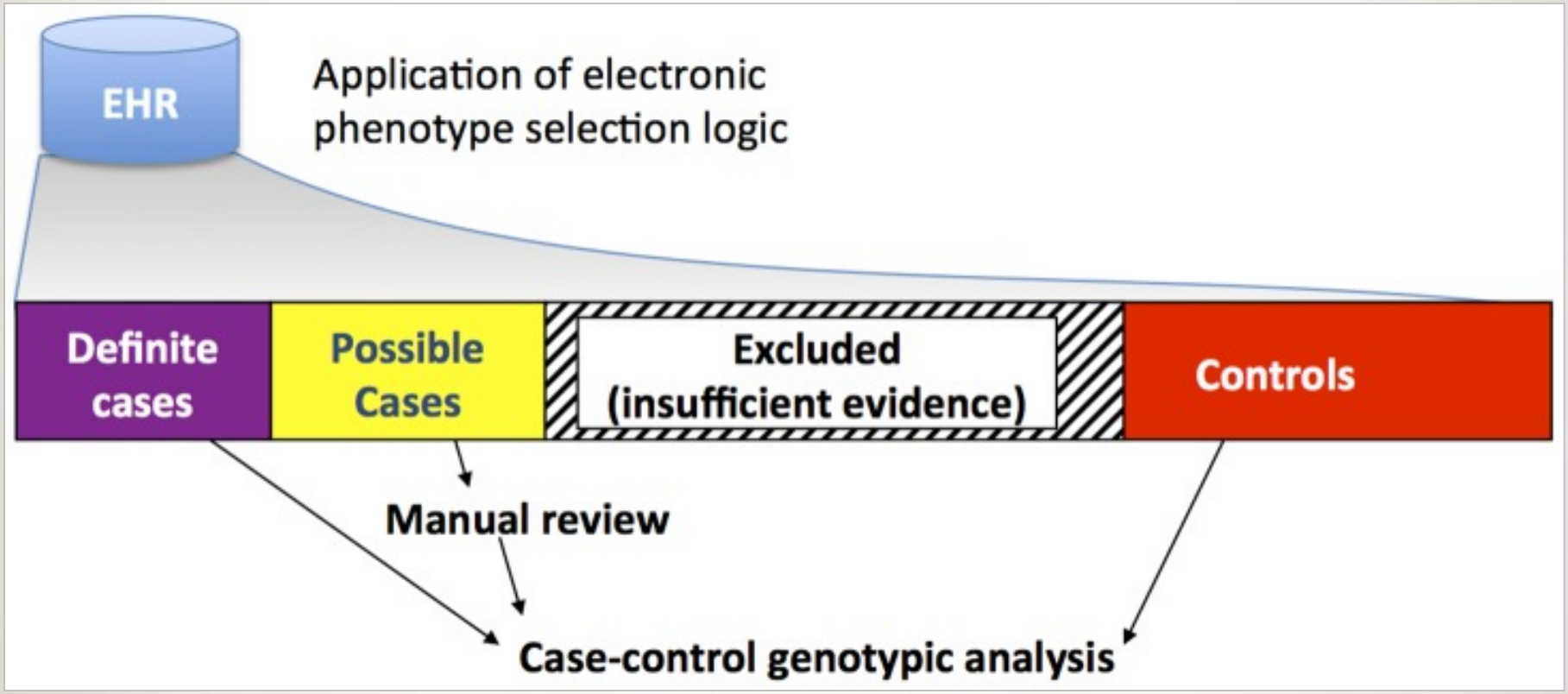
- Each site identified a phenotype of interest in ~3,000 subjects and conducted a genome-wide association study (GWAS)
- To what extent can identifiers be stripped from EMRs and research utility retained?
- Assess consent for genomic technologies & data sharing

Coordinating Center

VANDERBILT UNIVERSITY

CIDR Center for Inherited Disease Research

A common general approach to study design in BioVU and eMERGE



Phenotype definitions are portable across EMRs

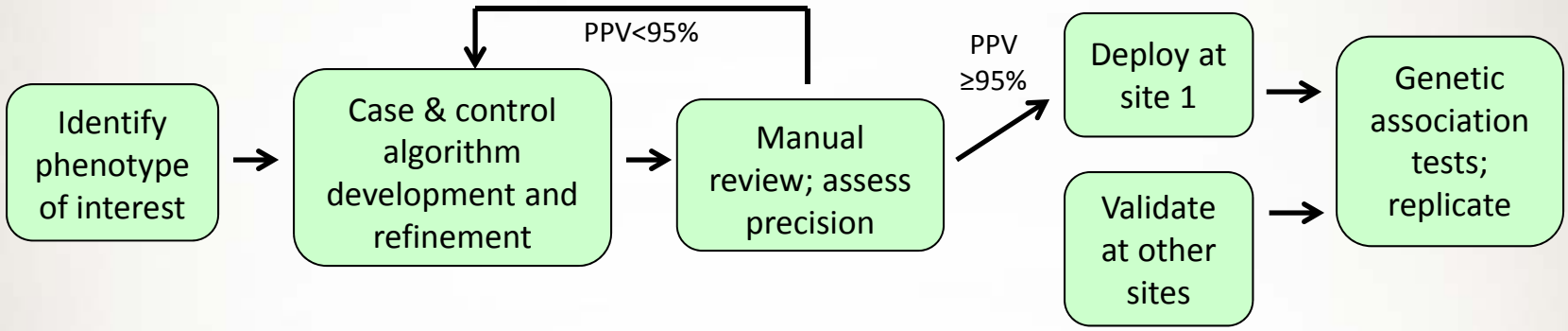


Table 1. Evaluation of Primary Hypothyroidism Algorithm at the Five eMERGE Sites

			Primary Hypothyroidism			
Site	**Primary Phenotype**	**Total Genotyped Subjects**	**Cases**	**Controls**	**Case PPV (%)**	**Control PPV (%)**
Group Health	dementia	2532	397	1,160	98	100
Marshfield	cataracts	4113	514	1,187	91	100
Mayo Clinic	peripheral arterial disease	3043	233	1,884	82	96
Northwestern	type 2 diabetes	1217	92	470	98	100
Vanderbilt	normal cardiac conduction	2712	81	352	98	100
All sites		13,617	1317	5053	92.4[a]	98.5[a]

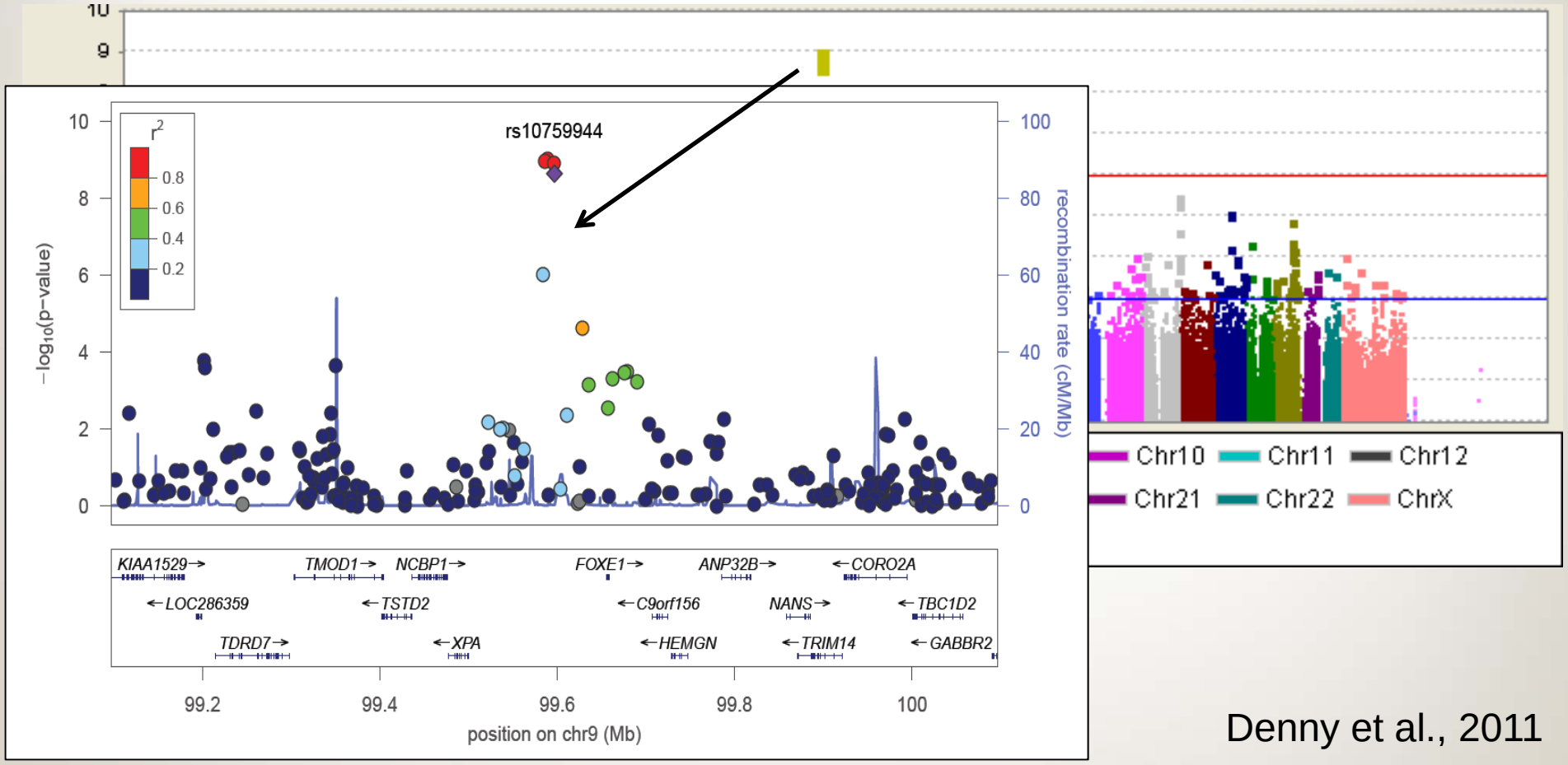
Genotype counts represent all subjects who were found by the hypothyroidism algorithms at each site and who were genotyped. Counts are limited to those classified as "white" in the electronic medical record of each site. PPV = positive predictive value.

[a] Average weighted for number of samples contributed to the total.

Denny et al., 2011

An eMERGE-wide phenotype analyzed with no extra genotyping: hypothyroidism

European Americans (1,306 cases and 5,013 controls)



Denny et al., 2011

Firefox | What is the Phenotype KnowledgeBase? ... | phekb.org | Bing

PheKB

a knowledgebase for discovering phenotypes from electronic medical records

Login | Register

Home | Phenotypes | Implementations | Groups | Institutions

What is the Phenotype KnowledgeBase?

The reuse of data from electronic medical records (EMRs) and other clinical data systems holds tremendous promise for improving the efficiency and effectiveness of health research. Clinical data in the EMR is a potential source of rich longitudinal data for research, and the recent government efforts to promote the use of EMRs in the clinical setting may further promote the use of such systems in the US healthcare system. As the use of EMRs expands, the demand for usable data from these systems for research has also expanded.

One such effort by the Electronic Medical Records and Genomics Network (eMERGE) has investigated whether data captured through routine clinical care using EMRs can identify disease phenotypes with sufficient positive and negative predictive values for use in genome-wide association studies (GWAS). Most EMRs captured key information (diagnoses, medications, laboratory tests) used to define phenotypes in a structured format; in addition, natural language processing has also been shown to improve case identification rates.*

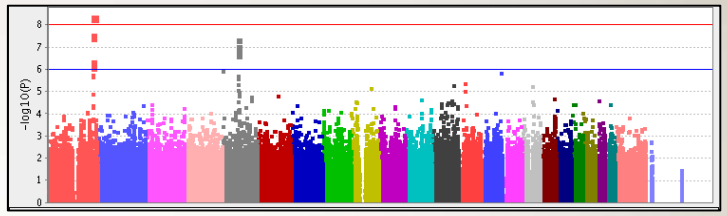
PheKB is an outgrowth of that validation effort and provides a collaborative environment of building and

MEDICAL CENTER

Most Recent Phenotypes

- Clopidogrel Poor Metabolizers
- Atrial Fibrillation - Demonstration Project
- Rhematoid Arthritis - Demonstration Project
- Multiple Sclerosis - Demonstration Project
- Crohn's Disease - Demonstration Project

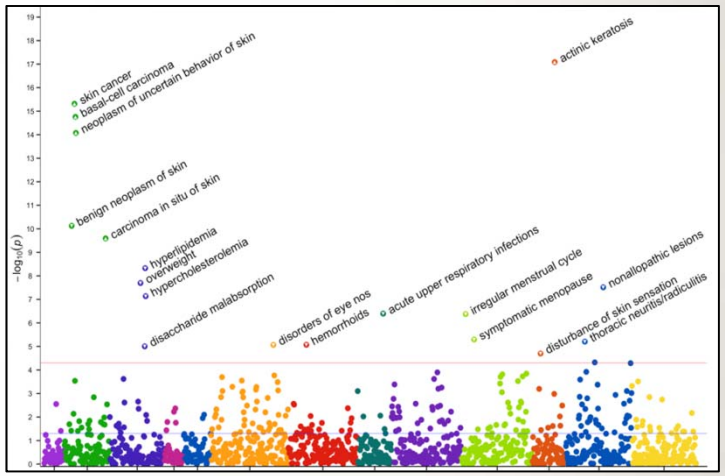
GWAS: Target phenotype → association P value



chromosomal location

The phenome-wide association study

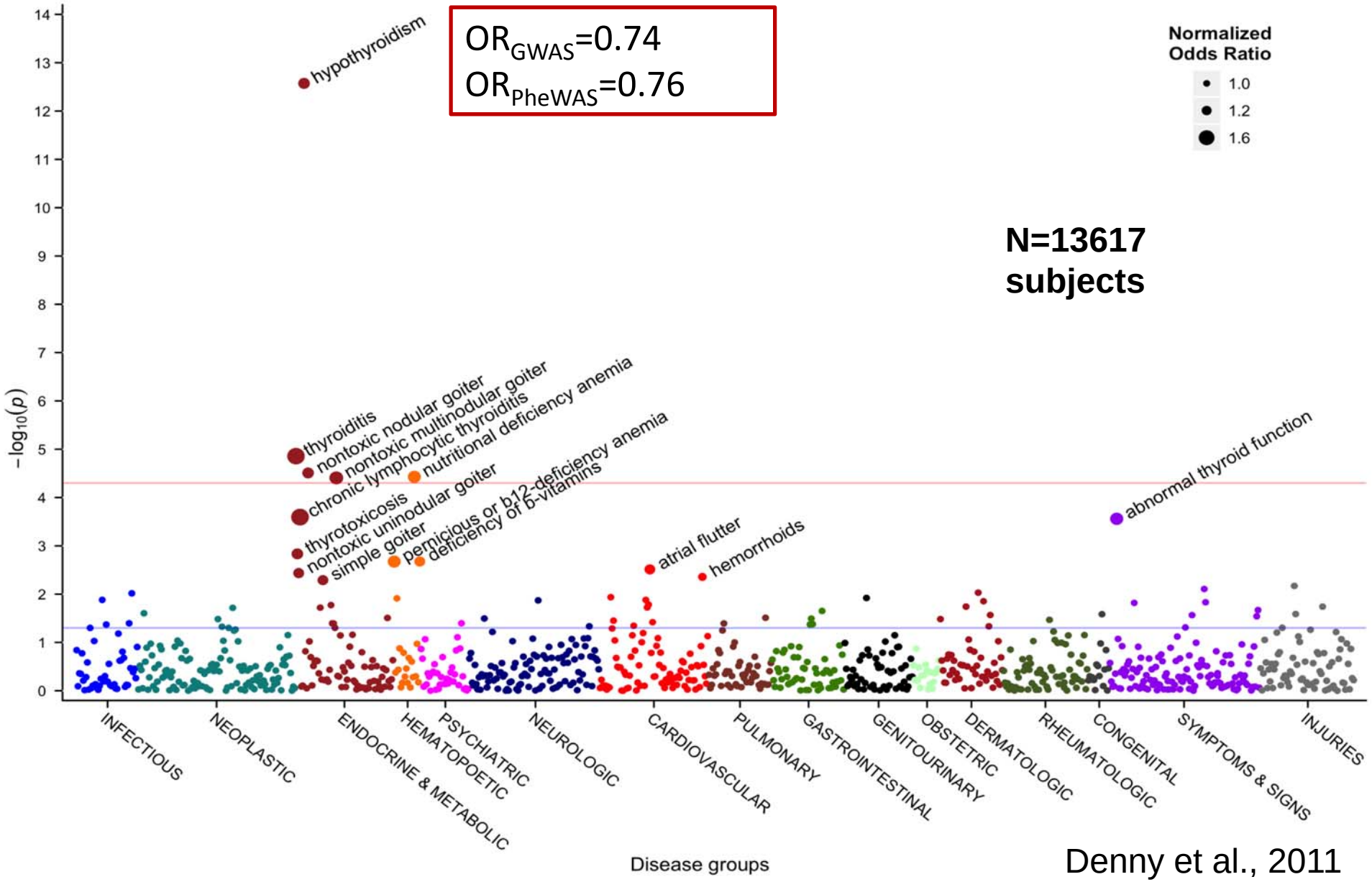
PheWAS (ΦWAS): Target genotype → association P value



diagnosis code

PheWAS requirement: A large cohort of patients with genotype data and many diagnoses

PheWAS for rs10759944 near *FOXE1*



Denny et al., 2011

GWAS catalog: 1751

PheWAS applications

- Replicating genotype-phenotype associations
- Discovering pleiotropic gene effects
- Disease subsetting
- Drug repurposing
- Expanding our understanding of gene regulation
- Engaging basic scientists

National Human Genome Research Institute

eMERGE Network

electronic medical records & genomics

2011-2015: Phase II

VANDERBILT UNIVERSITY MEDICAL CENTER

CIDR Center for Inherited Disease Research Johns Hopkins University

eMERGE Network

electronic medical records & genomics

2011-2015: Phase II
+ pediatric sites

eMERGE-II goals

- Expand the electronic phenotyping library and apply to genotyped samples
- Initiate implementation of actionable variants into the EMR
  - Site-specific projects
  - Cross network initiatives

ordinating Center

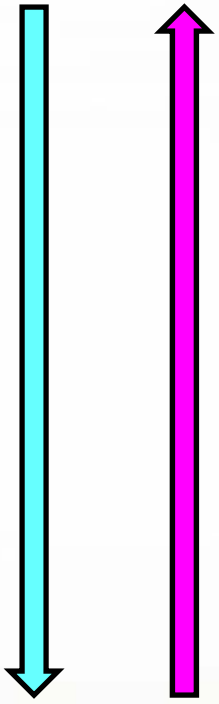
CIDR Center for Inherited Disease Research

eMERGE-PGRN Partnership

PGRN Pharmacogenomics Research Network

<u>PGx capabilities:</u>

- CPIC guidelines
- Resequencing platform for 84 Very Important Pharmacogenes
- CLIA & QC standards



<u>EMR-informatics capabilities</u>

- Privacy
- Electronic phenotyping
- Large populations
- Decision support

eMERGE Network
electronic medical records & genomics

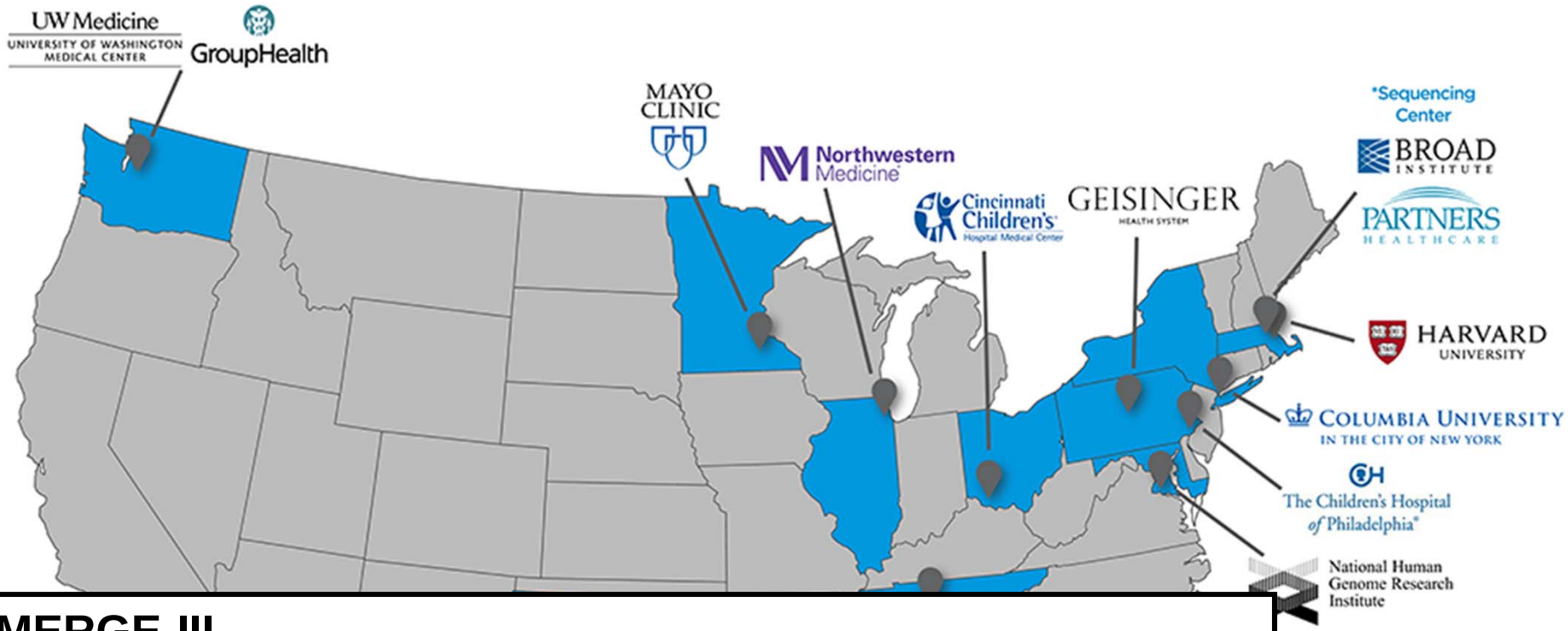
VANDER
MEDICAL CENTER

eMERGE-PGx Pharmacogene sequencing project

- identify patients (n=9,000)
- identify “actionable” variants
- sequence (82 key pharmacogenes)

implement actionable variants

create a repository of all variants



eMERGE-III

- Expand the electronic phenotyping library and apply to genotyped samples
- Targeted sequencing of 100 selected genes in 25,000 subjects:
  - Return “Known or Expected Pathogenic” variants
  - Create a repository of other variants: assess EMR phenotypes, ?penetrance

rge network
ONIC MEDICAL RECORDS & GENOMICS