

Table S2. eMERGE Phase II Phenotypes

Phenotype	Site	
	Primary	Secondary
Abdominal aortic aneurysm (AAA)	Geisinger	Mayo, Vanderbilt
buGWAS: <i>Clostridium difficile</i>	Group Health	Northwestern, Vanderbilt
Cardiorespiratory fitness (CRF)	Mayo	Geisinger, Vanderbilt
Diabetes/Hypertension-associated Chronic Kidney Disease (CKD)	Mount Sinai	Marshfield, Northwestern
Lower GI (non-syndromic polyps)	Northwestern	Vanderbilt
Ocular Hypertension and response to medication	Marshfield	Geisinger, Group Health
ACEI cough	Vanderbilt	Northwestern, Group Health
Lipids	CHOP	
Childhood Obesity	CCHMC & BCH	
Extreme Obesity (BMI)	Geisinger	Marshfield
Dose response for lipid lowering agents	Vanderbilt	Mayo, Mount Sinai
Glaucoma	Marshfield	Geisinger, Group Health
Lower GI (diverticulosis and -itis)	Northwestern	Marshfield, Vanderbilt
Drug Induced Liver Injury (DILI)	Mount Sinai	Marshfield, Mayo
Venous Thromboembolism (VTE)	Mayo	Vanderbilt
buGWAS: onychomycosis (<i>Tinea unguium</i>)	Group Health	Marshfield
Autism Spectrum Disorder	CCHMC & BCH	
Childhood Asthma	CHOP	
buGWAS: shingles (<i>Varicella zoster</i>)	Group Health	Vanderbilt
Heart Failure	Mayo	Group Health
Rapid renal decline in Diabetes/Hypertension-associated Chronic Kidney Disease (CKD)	Mount Sinai	Marshfield, Mayo
Community acquired MRSA	Northwestern	Marshfield, Geisinger
AMD or Dry eye & response to medication	Marshfield	Vanderbilt, Northwestern
Dose Response to lipid lowering agents	Vanderbilt	Mayo, Mount Sinai
Complications of Obesity -- resolution of diabetes after ROUX-EN-Y	Geisinger	Marshfield
Upper GI (Peptic ulcer disease)	Vanderbilt	Northwestern

The phenotyping workgroup of the eMERGE network has as its goal the creation, validation, and execution of phenotype algorithms across eMERGE. Algorithms are adapted and executed at each site to identify cases and controls for each phenotype of interest. These phenotype algorithms typically consist of combinations of billing codes, laboratory and test results, medication exposures, and natural language processing. The algorithms typically take 6-12 months to develop through a process of design and iterative validation. Validation across multiple sites is key to develop a transportable algorithm which can then be used in all eMERGE sites and outside the Network.