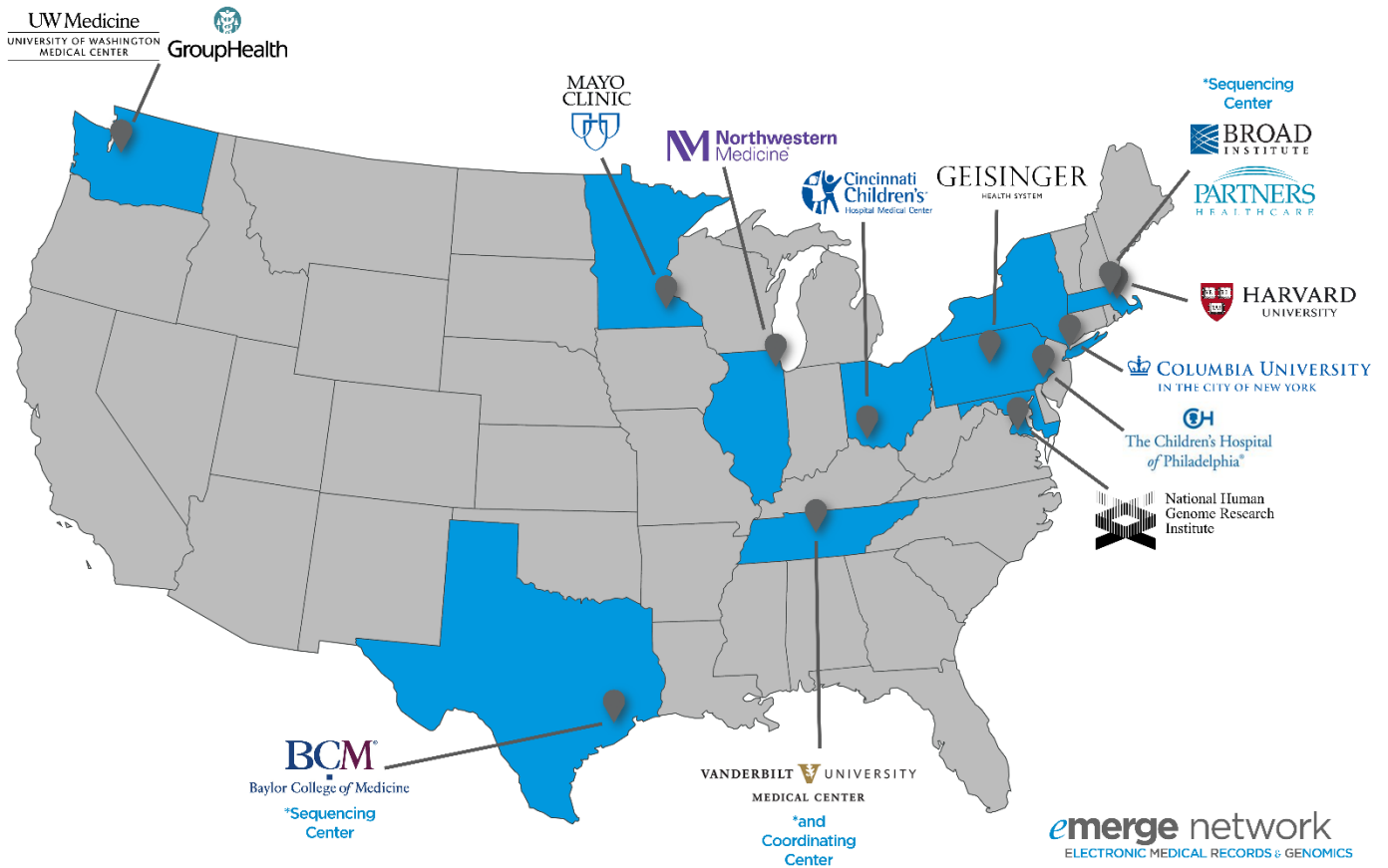


# External Scientific Panel

background materials

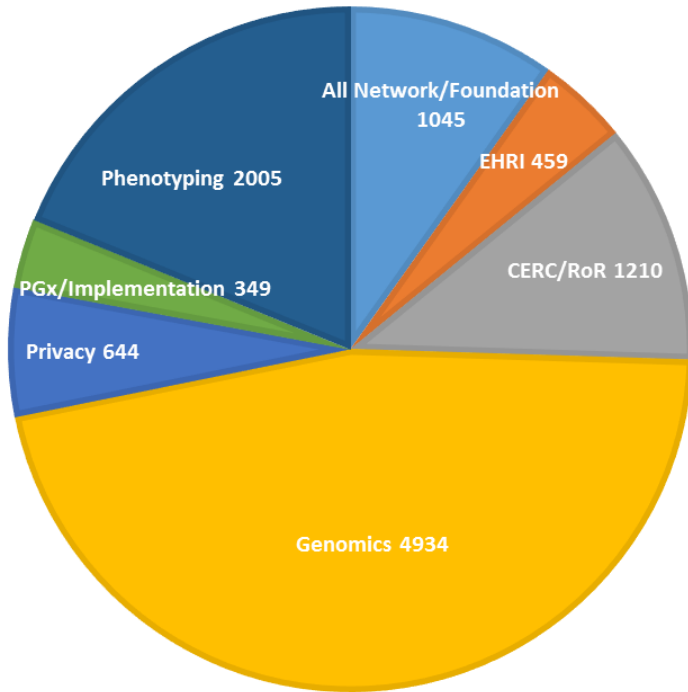


# TABLE *of* CONTENTS

Phase II Citation Analysis.....	<i>Pg. 2</i>
Phase II Publications.....	<i>Pg. 3</i>
Overview of eMERGE Tools.....	<i>Pg. 21</i>
Site Summaries.....	<i>Pg. 22</i>
Minutes from ESP Meeting, May 2015.....	<i>Pg. 32</i>
Minutes from ESP Meeting, May 2015, Executive Session.....	<i>Pg. 33</i>

# PHASE II CITATION ANALYSIS

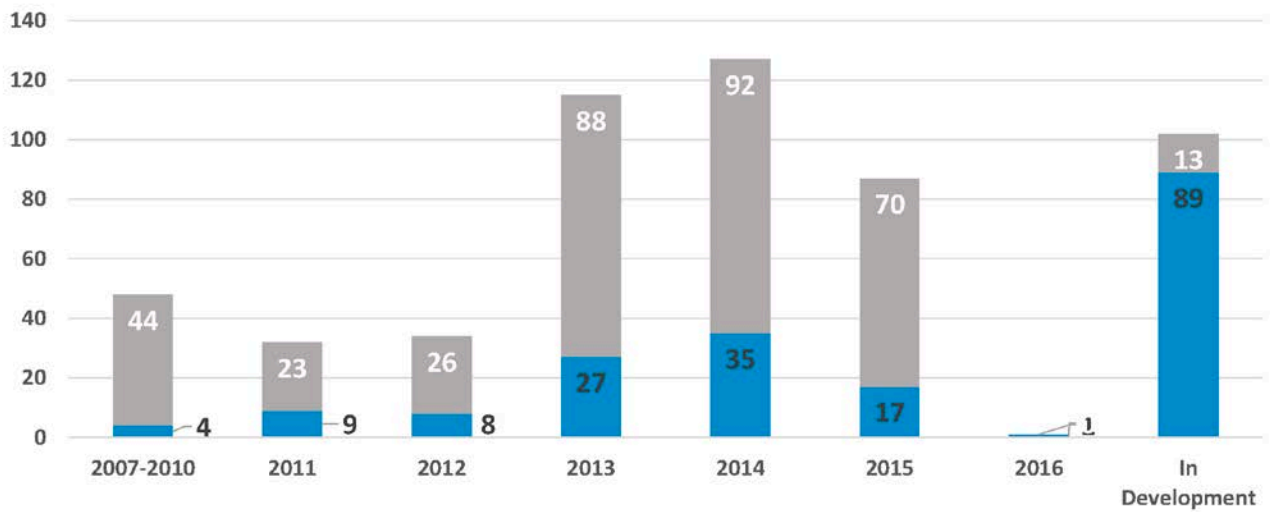
CITATIONS OF EMERGE PUBLICATIONS BY CATEGORY



## Cumulative Citation Count

2007 - January 2016: **10,646**

Number of published projects through January 2016



■ Site Projects (356)

■ Network Projects (190)

■ 546 Total Projects

# PHASE II BIBLIOGRAPHY & REMAINING IN-PROCESS PROJECTS

[ Digital Reference Library Available [Here](#) ]

## eMERGE Phase II Publications from June 2011 – January 2016

Digital Reference Library Available [Here](#)

### In Progress Phase III Network Manuscripts

1. Approaches to Returning Clinically Actionable Results from Next Generation Sequencing Panel in a Healthy Population. Lead Investigator: Tracy McGregor (VU)
2. Knowledge driven rare variant PheWAS in eMERGE to identify regions associated with disease using collapsing based approach. Lead Investigator: Anna O Basile (Geisinger)
3. GWAS study on non-alcoholic fatty liver disease (NAFLD) in pediatric and adult population: comparison of size effect between adult and children using participants of the eMERGE Network. Lead Author: Bahram Namjou (CCHMC)
4. The identification and reporting of actionable incidental genetic variants from large scale clinical sequencing of drug response genes. Lead Investigator Quinn Wells (VU).
5. Feasibility of using geocoded US Census/ACS variables as proxy for socioeconomic status in genotype-phenotype interaction studies of T2DM and obesity. Lead Investigator: Kathryn Jackson (NU)

### Published/Accepted and Submitted Phase II Network Manuscripts

1. Kirby J, Speltz P, Basford M, Roden D, Haines J et al. *PheKB: A Catalog and Workflow for Creating Electronic Phenotype Algorithms for Transportability*. (Submitted).
2. Ritchie M, De R, Verma S, Holmes MV, Holzinger ER et al. *Gene-Gene interactions associated with BMI - Replication in the eMERGE network*. (In-Press).
3. Ritchie M, Verma S, Pendergrass S, Verma A, Kuivaniemi H et al. *Null Variant PheWAS*. (Accepted).
4. Jones G, Kuivaniemi H, Tromp G, Carey DJ, Smelser D et al. *Meta analysis of genome-wide association studies for abdominal aortic aneurysms*. (In Press).
5. Ritchie M, Lucas A, Verma S, Kim D, Crosslin D. et al. *Copy number variation burden analysis on a range of phenotypes in the eMERGE network*. (In Press).
6. Van Driest SL, Wells QS, Stallings S, et al. *Association of arrhythmia-related genetic variants with phenotypes documented in electronic medical records*. JAMA. 2016 Jan 5;315(1):47–57.
7. Brothers KB, Holm IA, Childerhose JE, Antommaria AHM, Bernhardt BA, et al. *When Participants in Genomic Research Grow Up: Contact and Consent at the Age of Majority*. J Pediatr. 2016 Jan;168:226–231.e1. **PMID: 26477867**
8. Xu J, Rasmussen LV, Shaw PL, Jiang G, Kiefer RC. *Review and evaluation of electronic health records-driven phenotype algorithm authoring tools for clinical and translational research*. J Am Med Inform Assoc. 2015 Nov;22(6):1251–1260. **PMID: 26224336**
9. Wang L-S, Naj AC, Graham RR, Crane PK, Kunkle BW, et al. *Rarity of the Alzheimer disease-protective APP A673T variant in the United States*. JAMA Neurol. 2015 Feb;72(2):209–216. **PMCID: PMC4324097**
10. Wan Z, Vorobeychik Y, Xia W, Clayton EW, Kantarcioglu M et al. *A game theoretic framework for analyzing re-identification risk*. PLoS ONE. 2015;10(3):e0120592. **PMCID: PMC4373733**
11. Van Driest SL, McGregor TL, Velez Edwards DR, Saville BR, Kitchner TE et al. *Genome-Wide Association Study of Serum Creatinine Levels during Vancomycin Therapy*. PLoS ONE. 2015;10(6):e0127791. **PMCID: PMC4452656**
12. Shirts BH, Salama JS, Aronson SJ, Chung WK, Gray SW et al. *CSER and eMERGE: current and potential state of the display of genetic information in the electronic health record*. J Am Med Inform Assoc. 2015 Nov;22(6):1231–1242. **PMID: 26142422**
13. Namjou B, Marsolo K, Lingren T, Ritchie MD, Verma SS et al. *A GWAS Study on Liver Function Test Using eMERGE Network Participants*. PLoS ONE. 2015;10(9):e0138677. **PMCID: PMC4586138**
14. Mosley JD, Shaffer CM, Van Driest SL, Weeke PE, Wells QS et al. *A genome-wide association study identifies variants in KCNIP4 associated with ACE inhibitor-induced cough*. Pharmacogenomics J. 2015 Jul 14; **PMID: 26169577**
15. Mo H, Thompson WK, Rasmussen LV, Pacheco JA, Jiang G et al. *Desiderata for computable representations of electronic health records-driven phenotype algorithms*. J Am Med Inform Assoc. 2015 Nov;22(6):1220–1230. **PMCID: PMC4639716**
16. Interleukin 1 Genetics Consortium. *Cardiometabolic effects of genetic upregulation of the interleukin 1 receptor antagonist: a Mendelian randomisation analysis*. Lancet Diabetes Endocrinol. 2015 Apr;3(4):243–253. **PMCID: PMC4648058**
17. Herr TM, Bielinski SJ, Bottinger E, Brautbar A, Brilliant M, et al. *A conceptual model for translating omic data into clinical action*. J Pathol Inform. 2015;6:46. **PMCID: PMC4584438**
18. Herr TM, Bielinski SJ, Bottinger E, Brautbar A, Brilliant M, et al. *Practical considerations in genomic decision support: The eMERGE*

- experience. *J Pathol Inform.* 2015;6:50. **PMCID: PMC4629307**
19. Hall MA, Verma SS, Wallace J, Lucas A, Berg RL, et al. *Biology-Driven Gene-Gene Interaction Analysis of Age-Related Cataract in the eMERGE Network.* *Genet Epidemiol.* 2015 Jul;39(5):376–384. **PMCID: PMC4550090**
  20. Goldstein BA, Knowles JW, Salfati E, Ioannidis JPA, Assimes TL. *Corrigendum: Simple, standardized incorporation of genetic risk into non-genetic risk prediction tools for complex traits: coronary heart disease as an example.* *Front Genet.* 2015;6:231. **PMCID: PMC4493401**
  21. Garrison NA, Sathe NA, Antommaria AHM, Holm IA, Sanderson SC, et al. *A systematic literature review of individuals' perspectives on broad consent and data sharing in the United States.* *Genet Med.* 2015 Nov 19; **PMID: 26583683**
  22. Gallego CJ, Burt A, Sundaresan AS, Ye Z, Shaw C, et al. *Penetrance of Hemochromatosis in HFE Genotypes Resulting in p.Cys282Tyr and p.[Cys282Tyr];[His63Asp] in the eMERGE Network.* *Am J Hum Genet.* 2015 Oct 1;97(4):512–520. **PMCID: PMC4596892**
  23. Crosslin DR, Carrell DS, Burt A, Kim DS, Underwood JG, et al. *Genetic variation in the HLA region is associated with susceptibility to herpes zoster.* *Genes Immun.* 2015 Feb;16(1):1–7. **PMCID: PMC4308645**
  24. Crosslin DR, Robertson PD, Carrell DS, Gordon AS, Hanna DS, et al. *Prospective participant selection and ranking to maximize actionable pharmacogenetic variants and discovery in the eMERGE Network.* *Genome Med.* 2015;7(1):67. **PMCID: PMC4517371**
  25. Borthwick KM, Smelser DT, Bock JA, Elmore JR, Ryer EJ, et al. *ePhenotyping for Abdominal Aortic Aneurysm in the Electronic Medical Records and Genomics (eMERGE) Network: Algorithm Development and Konstanz Information Miner Workflow.* *International Journal of Biomedical Data Mining* 2015 Jul 14;2015.
  26. Bielinski SJ, Pathak J, Carrell DS, Takahashi PY, Olson JE, et al. *A Robust e-Epidemiology Tool in Phenotyping Heart Failure with Differentiation for Preserved and Reduced Ejection Fraction: the Electronic Medical Records and Genomics (eMERGE) Network.* *J Cardiovasc Transl Res.* 2015 Nov;8(8):475–483. **PMCID: PMC4651838**
  27. Ye Z, Vasco DA, Carter T, Brilliant M, Schrodi SJ, et al. *Genome wide association study of SNP-, gene-, and pathway-based approaches to identify genes influencing susceptibility to Staphylococcus aureus infections.* *Staphylococcus aureus.* 2014 May 9;5:125. **PMCID: PMC4023021**
  28. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, et al. *Defining the role of common variation in the genomic and biological architecture of adult human height.* *Nat Genet.* 2014 Nov;46(11):1173–1186. **PMCID: PMC4250049**
  29. Verma SS, de Andrade M, Tromp G, Kuivaniemi H, Pugh E, et al. *Imputation and quality control steps for combining multiple genome-wide datasets.* *Front Genet.* 2014;5:370. **PMCID: PMC4263197**
  30. Sun X, Lu Q, Mukherjee S, Crane P, Elston RC, et al. *Analysis pipeline for the epistasis search – statistical versus biological filtering.* *Front Genet.* 2014 Apr 30;5:106. **PMCID: PMC4012196**
  31. Sleiman P, Bradfield J, Mentch F, Almoguera B, Connolly J, et al. *Assessing the functional consequence of loss of function variants using electronic medical record and large-scale genomics consortium efforts.* *Front Genet.* 2014 Apr 29;5:105. **PMCID: PMC4010747**
  32. Schrodi SJ, Mukherjee S, Shan Y, Tromp G, Sninsky JJ, et al. *Genetic-based prediction of disease traits: prediction is very difficult, especially about the future(†).* *Front Genet.* 2014;5:162. **PMID: 24917882**
  33. Ritchie MD, Verma SS, Hall MA, Goodloe RJ, Berg RL, et al. *Electronic medical records and genomics (eMERGE) network exploration in cataract: several new potential susceptibility loci.* *Mol Vis.* 2014;20:1281–1295. **PMCID: PMC4168835**
  34. Rasmussen-Torvik LJ, Stallings SC, Gordon AS, Almoguera B, Basford MA, et al. *Design and Anticipated Outcomes of the eMERGE-PGx Project: A Multi-Center Pilot for Pre-Emptive Pharmacogenomics in Electronic Health Record Systems.* *Clin Pharmacol Ther.* 2014 Jun 24; **PMCID: PMC4169732**
  35. Rasmussen LV, Thompson WK, Pacheco JA, Kho AN, Carrell DS, et al. *Design patterns for the development of electronic health record-driven phenotype extraction algorithms.* *J Biomed Inform.* 2014 Jun 21; **PMCID: PMC4194216**
  36. Patel ZH, Kottyan LC, Lazaro S, Williams MS, Ledbetter DH, et al. *The struggle to find reliable results in exome sequencing data: filtering out Mendelian errors.* *Front Genet.* 2014 Feb 12;5:16. **PMCID: PMC3921572**
  37. Parihar A, Wood GC, Chu X, Jin Q, Argyropoulos G, et al. *Extension of GWAS results for lipid-related phenotypes to extreme obesity using electronic health record (EHR) data and the MetaboChip.* *Front Genet.* 2014;5:222. **PMCID: PMC4123014**
  38. 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;2014:944–953. **PMCID: PMC4419923**
  39. Ng MCY, Shriner D, Chen BH, Li J, Chen W-M, et al. *Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes.* *PLoS Genet.* 2014 Aug;10(8):e1004517. **PMCID: PMC4125087**
  40. Namjou B, Marsolo K, Carroll RJ, Denny JC, Ritchie MD, et al. *Phenome-wide association study (PheWAS) in EMR-linked pediatric cohorts, genetically links PLCL1 to speech language development and IL5-IL13 to Eosinophilic Esophagitis.* *Front Genet.* 2014;5:401. **PMCID: PMC4235428**
  41. Nadkarni GN, Gottesman O, Linneman JG, Chase H, Berg RL, et al. *Development and validation of an electronic phenotyping algorithm for chronic kidney disease.* *AMIA Annu Symp Proc.* 2014;2014:907–916. **PMCID: PMC4419875**
  42. Muthalagu A, Pacheco JA, Aufox S, Peissig PL, Fuehrer JT, et al. *A Rigorous Algorithm To Detect And Clean Inaccurate Adult Height Records Within EHR Systems.* *Applied Clinical Informatics.* 2014 Feb 19;5(1):118–126. **PMCID: PMC3974252**

43. Mitchell BD, Fornage M, McArdle PF, Cheng Y-C, Pulit S, et al. *Using previously genotyped controls in genome-wide association studies (GWAS): application to the Stroke Genetics Network (SiGN)*. *Front Genet.* 2014 Apr 29;5:95. **PMCID: PMC4010766**
44. Malinowski JR, Denny JC, Bielinski SJ, Basford MA, Bradford Y, et al. *Genetic variants associated with serum thyroid stimulating hormone (TSH) levels in European Americans and African Americans from the eMERGE Network*. *PLoS ONE.* 2014;9(12):e111301. **PMCID: PMC4249871**
45. Kullo IJ, Shameer K, Jouni H, Lesnick TG, Pathak J, et al. *The ATXN2-SH2B3 locus is associated with peripheral arterial disease: an electronic medical record-based genome-wide association study*. *Front Genet.* 2014;5:166. **PMCID: PMC4070196**
46. Kullo IJ, Haddad R 'ad, Prows CAM, Holm I, Sanderson SC, et al. *Return of Genomic Results in the Genomic Medicine Projects of the eMERGE Network*. *Front Genet.* 2014 Mar 26;5:50. **PMCID: PMC3972474**
47. Jeff JM, Brown-Gentry K, Goodloe R, Ritchie MD, Denny JC, et al. *Replication of SCN5A Associations with Electrocardio-graphic Traits in African Americans from Clinical and Epidemiologic Studies*. *Evol Comput Mach Learn Data Min Bioinform.* 2014;2014:939–951. **PMCID: PMC4290789**
48. Jeff JM, Armstrong LL, Ritchie MD, Denny JC, Kho AN, et al. *Admixture mapping and subsequent fine-mapping suggests a biologically relevant and novel association on chromosome 11 for type 2 diabetes in african americans*. *PLoS ONE.* 2014 Mar 3;9(3):e86931. **PMCID: PMC3940426**
49. Jarvik GP, Amendola LM, Berg JS, Brothers K, Clayton EW, et al. *Return of Genomic Results to Research Participants: The Floor, the Ceiling, and the Choices In Between*. *Am J Hum Genet.* 2014 May 7; **PMCID: PMC4121476**
50. Crosslin DR, Tromp G, Burt A, Kim DS, Verma SS, et al. *Controlling for population structure and genotyping platform bias in the eMERGE multi-institutional biobank linked to electronic health records*. *Front Genet.* 2014;5:352. **PMCID: PMC4220165**
51. Cronin RM, Field JR, Bradford Y, Shaffer CM, Carroll RJ, et al. *Phenome-wide association studies demonstrating pleiotropy of genetic variants within FTO with and without adjustment for body mass index*. *Front Genet.* 2014;5:250. **PMCID: PMC4134007**
52. Crawford DC, Crosslin DR, Tromp G, Kullo IJ, Kuivaniemi H, et al. *eMERGEing progress in genomics—the first seven years*. *Front Genet.* 2014;5:184. **PMCID: PMC4060012**
53. Connolly JJ, Glessner JT, Almoguera B, Crosslin DR, Jarvik GP, et al. *Copy number variation analysis in the context of electronic medical records and large-scale genomics consortium efforts*. *Front Genet.* 2014 Mar 18;5:51. **PMCID: PMC3957100**
54. Brothers KB, Lynch JA, Aufox SA, Connolly JJ, Gelb BD, et al. *Practical guidance on informed consent for pediatric participants in a biorepository*. *Mayo Clin Proc.* 2014 Nov;89(11):1471–1480. **PMCID: PMC4254313**
55. Arking DE, Pulit SL, Crotti L, van der Harst P, Munroe PB, et al. *Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization*. *Nat Genet.* 2014 Aug;46(8):826–836. **PMCID: PMC4124521**
56. Almoguera B, Vazquez L, Connolly JJ, Bradfield J, Sleiman P, et al. *Imputation of TPMT defective alleles for the identification of patients with high-risk phenotypes*. *TPMT.* 2014 May 12;5:96. **PMCID: PMC4026736**
57. Wei W-Q, Leibson CL, Ransom JE, Kho AN, Chute CG. *The absence of longitudinal data limits the accuracy of high-throughput clinical phenotyping for identifying type 2 diabetes mellitus subjects*. *Int J Med Inform.* 2013 Apr;82(4):239–247. **PMCID: PMC3478423**
58. Ury AG. *Storing and interpreting genomic information in widely deployed electronic health record systems*. *Genet Med.* 2013 Oct;15(10):779–785. **PMID: 23949573**
59. Tarczy-Hornoch P, Amendola L, Aronson SJ, Garraway L, Gray S, et al. *A survey of informatics approaches to whole-exome and whole-genome clinical reporting in the electronic health record*. *Genet Med.* 2013 Oct;15(10):824–832. **PMCID: PMC3951437**
60. Starren J, Williams MS, Bottinger EP. *Crossing the Omic Chasm: A Time for Omic Ancillary Systems*. *JAMA.* 2013 Mar 14;1–2. **PMCID: PMC3857698**
61. Schick UM, McDavid A, Crane PK, Weston N, Ehrlich K, et al. *Confirmation of the Reported Association of Clonal Chromosomal Mosaicism with an Increased Risk of Incident Hematologic Cancer*. *PLoS ONE.* 2013 Mar 22;8(3):e59823. **PMID: 23533652**
62. Ritchie MD, Denny JC, Zuvich RL, Crawford DC, Schildcrout JS, et al. *Genome- and Phenome-Wide Analysis of Cardiac Conduction Identifies Markers of Arrhythmia Risk*. *Circulation.* 2013 Apr 2;127(13):1377-85. **PMCID: PMC3713791**
63. Peterson JF, Bowton E, Field JR, Beller M, Mitchell J, et al. *Electronic health record design and implementation for pharmacogenomics: a local perspective*. *Genet Med.* 2013 Oct;15(10):833–841. **PMCID: PMC3925979**
64. Pathak J, Kho AN, Denny JC. *Electronic health records-driven phenotyping: challenges, recent advances, and perspectives*. *J Am Med Inform Assoc.* 2013 Dec;20(e2):e206–211. **PMCID: PMC3861925**
65. Overby CL, Pathak J, Gottesman O, Haerian K, Perotte A, et al. *A collaborative approach to developing an electronic health record phenotyping algorithm for drug-induced liver injury*. *J Am Med Inform Assoc.* 2013 Dec;20(e2):e243–252. **PMCID: PMC3861914**
66. Overby CL, Kohane I, Kannry JL, Williams MS, Starren J, et al. *Opportunities for genomic clinical decision support interventions*. *Genet Med.* 2013 Oct;15(10):817–823. **PMCID: PMC3858176**
67. Newton KM, Peissig PL, Kho AN, Bielinski SJ, Berg RL, et al. *Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network*. *J Am Med Inform Assoc.* 2013 Jun;20(e1):e147-54. **PMCID: PMC3715338**
68. Namjou B, Keddache M, Marsolo K, Wagner M, Lingren T, et al. *EMR-linked GWAS study: investigation of variation landscape of loci for body mass index in children*. *Front Genet.* 2013 Dec 3;4:268. **PMCID: PMC3847941**

69. Mosley JD, Van Driest SL, Larkin EK, Weeke PE, Witte JS, et al. *Mechanistic Phenotypes: An Aggregative Phenotyping Strategy to Identify Disease Mechanisms Using GWAS Data*. PLoS ONE. 2013 Dec 12;8(12):e81503. **PMCID: PMC3861317**
70. McDavid A, Crane PK, Newton KM, Crosslin DR, McCormick W, et al. *Enhancing the Power of Genetic Association Studies through the Use of Silver Standard Cases Derived from Electronic Medical Records*. PLoS ONE. 2013 Jun 10;8(6):e63481. **PMCID: PMC3677889**
71. Marsolo K, Spooner SA. *Clinical genomics in the world of the electronic health record*. Genet Med. 2013 Oct;15(10):786–791. **PMCID: PMC4096770**
72. Kho AN, Rasmussen LV, Connolly JJ, Peissig PL, Starren J, et al. *Practical challenges in integrating genomic data into the electronic health record*. Genet Med. 2013 Oct;15(10):772–778. **PMCID: PMC4201621**
73. Kannry JM, Williams MS. *Integration of genomics into the electronic health record: mapping terra incognita*. Genet Med. 2013 Oct;15(10):757–760. **PMCID: PMC4157459**
74. Kannry J, Williams MS. *The undiscovered country: the future of integrating genomic information into the EHR*. Genet Med. 2013 Oct;15(10):842–845. **PMCID: PMC4259267**
75. Jeff JM, Ritchie MD, Denny JC, Kho AN, Ramirez AH, et al. *Generalization of Variants Identified by Genome-Wide Association Studies for Electrocardiographic Traits in African Americans*. Ann Hum Genet. 2013 Jul;77(4):321-32. **PMCID: PMC3743946**
76. Hazin R, Brothers KB, Malin BA, Koenig BA, Sanderson SC, et al. *Ethical, legal, and social implications of incorporating genomic information into electronic health records*. Genet Med. 2013 Oct;15(10):810–816. **PMID: 2403043**
77. Hartzler A, McCarty CA, Rasmussen LV, Williams MS, Brilliant M, et al. *Stakeholder engagement: a key component of integrating genomic information into electronic health records*. Genet Med. 2013 Oct;15(10):792-801. **PMCID: PMC3909653**
78. Gottesman O, Kuivaniemi H, Tromp G, Faucett WA, Li R, et al. *The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future*. Genet Med. 2013 Oct;15(10):761-71. **PMCID: PMC3795928**
79. Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, 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. **PMCID: PMC3969265**
80. Crosslin DR, McDavid A, Weston N, Zheng X, Hart E, et al. *Genetic variation associated with circulating monocyte count in the eMERGE Network*. Hum Mol Genet. 2013 May 15;22(10):2119-27. **PMCID: PMC3633369**
81. Chute CG, Ullman-Cullere M, Wood GM, Lin SM, He M, et al. *Some experiences and opportunities for big data in translational research*. Genet Med. 2013 Oct;15(10):802–809. **PMCID: PMC3906918**
82. Chute CG, Kohane IS. *Genomic medicine, health information technology, and patient care*. JAMA. 2013 Apr 10;309(14):1467–1468. **PMCID: PMC3959893**
83. Wei W-Q, Leibson CL, Ransom JE, Kho AN, Caraballo PJ, et al. *Impact of data fragmentation across healthcare centers on the accuracy of a high-throughput clinical phenotyping algorithm for specifying subjects with type 2 diabetes mellitus*. J Am Med Inform Assoc. 2012 Mar 2;19(2):219–224. **PMCID: PMC3277630**
84. Rasmussen-Torvik LJ, Pacheco JA, Wilke RA, Thompson WK, Ritchie MD, et al. *High Density GWAS for LDL Cholesterol in African Americans Using Electronic Medical Records Reveals a Strong Protective Variant in APOE*. Clinical and Translational Science. 2012;5(5):394–399. **PMCID: PMC3521536**
85. Peissig PL, Rasmussen LV, Berg RL, Linneman JG, McCarty CA, et al. *Importance of multi-modal approaches to effectively identify cataract cases from electronic health records*. J Am Med Inform Assoc. 2012 Mar 1;19(2):225–234. **PMCID: PMC3277618**
86. Kullo IJ, Jarvik GP, Manolio TA, Williams MS, Roden DM. *Leveraging the electronic health record to implement genomic medicine*. Genet Med. 2012 Sep 27; **PMCID: PMC4206937**
87. Kho AN, Hayes MG, Rasmussen-Torvik L, Pacheco JA, Thompson WK, et al. *Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study*. J Am Med Inform Assoc. 2012 Mar 1;19(2):212–218. **PMCID: PMC3277617**
88. Fullerton SM, Wolf WA, Brothers KB, Clayton EW, Crawford DC, et al. *Return of individual research results from genome-wide association studies: experience of the Electronic Medical Records and Genomics (eMERGE) Network*. Genet Med. 2012 Apr;14(4):424–431. **PMCID: PMC3723451**
89. Ding K, Shameer K, Jouni H, Masys DR, Jarvik GP, et al. *Genetic Loci implicated in erythroid differentiation and cell cycle regulation are associated with red blood cell traits*. Mayo Clin Proc. 2012 May;87(5):461–474. **PMCID: PMC3538470**
90. Crosslin DR, McDavid A, Weston N, Nelson SC, Zheng X, et al. *Genetic variants associated with the white blood cell count in 13,923 subjects in the eMERGE Network*. Hum Genet. 2012 Apr;131(4):639–652. **PMCID: PMC3640990**
91. Zuvich RL, Armstrong LL, Bielinski SJ, Bradford Y, Carlson CS, et al. *Pitfalls of merging GWAS data: lessons learned in the eMERGE network and quality control procedures to maintain high data quality*. Genet Epidemiol. 2011 Dec;35(8):887–898. **PMCID: PMC3592376**
92. Pathak J, Wang J, Kashyap S, Basford M, Li R, et al. *Mapping clinical phenotype data elements to standardized metadata repositories and controlled terminologies: the eMERGE Network experience*. J Am Med Inform Assoc. 2011 Jul;18(4):376–386. **PMCID: PMC3128396**

93. McGuire AL, Basford M, Dressler LG, Fullerton SM, Koenig BA, et al. *Ethical and practical challenges of sharing data from genome-wide association studies: The eMERGE Consortium experience*. *Genome Res*. 2011 Jul;21(7):1001–1007. **PMCID: PMC3129243**
94. Kullo IJ, Ding K, Shameer K, McCarty CA, Jarvik GP, et al. *Complement Receptor 1 Gene Variants Are Associated with Erythrocyte Sedimentation Rate*. *Am J Hum Genet*. 2011 Jul 15;89(1):131–138. **PMCID: PMC3135803**

## In Process Phase II Network Manuscripts

1. Association of rare and common variants with insulin resistance: results from the eMERGE Network. Lead Investigator Daniel Kim (UW)
2. Investigation of CETP SNPs with LDL-C, BMI and risk of T2D. Lead Investigator: Brendan Keating (UPenn)
3. Quantitative and discreet trait analysis across eMERGE-I phenotypes. Lead Investigator: Joseph Glessner (CHOP)
4. Association of Y chromosomal variants with leukocyte count and cardiovascular disease: results from the eMERGE Network. Lead Investigator: Yatong Li (UW)
5. Multi-site IRB review – experience of the eMERGE Network. Lead Investigator: Jen McCormick (Mayo)
6. A phenome-wide association study to discover pleiotropic effects of PCSK9. Lead Investigator: Maya Saforova (Mayo)
7. An investigation of somatic mutations in PGx-eMERGE dataset. Lead Investigator Kenneth Kaufman (CCHMC)
8. The eMERGE Network: The practice of patient education in the return of genomic medicine results. Lead Investigator: Cassandra Perry (BCH)
9. Assessing opinions about healthcare provider-targeted MyResults.org content. Lead Investigator: Casey Overby (NU)
10. Genome wide association studies and gene based interaction study for Ocular Hypertension. Lead Investigator: Shefali Verma (Penn State)
11. Agreement between research-grade sequencing and CLIA validation genotyping in eMERGE-PGx. Lead Investigator: Laura R Torvik (NU)
12. Replication of another sites' GWAS findings for Squamous Cell Carcinoma (SCC) and Actinic Keratosis (AK) findings using existing PheWAS codes from the "Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data" paper. Lead Investigator: Josh Denny (VU)
13. The identification of adverse events in the eMERGE PGx cohort using the electronic health record, and assessing association with genetic variation in the 84 pharmacogenes. Lead Investigator: David Crosslin (UW)
14. An investigation into the genetics of Intractable Epilepsy in the pediatric population. Lead Investigator: Berta Castillo (CHOP)
15. Patients' views on consent and data sharing in biobank research: A large multisite experimental survey in the US. Lead Investigator: Saskia Sanderson (MSSM)
16. The eMERGE Network: Healthcare provider education to support genomic medicine in practice. Lead Investigator: Carolyn R. Vitek (Mayo)
17. A targeted sampling scheme utilizing both EHR and census information. Lead Investigator: Nate Mercaldo (VU)
18. Feasibility of using geocoded US Census/ACS variables as proxy for socioeconomic status in genotype-phenotype interaction studies of T2DM and obesity. Lead Investigator: Kathryn Jackson (NU)
19. Cognitive Interviews associated with developing a national survey on consent across a national network of genomic medicine sites. Lead Investigator: Melanie Myers (CCHMC)
20. Developing a national survey on consent across a national network of genomic medicine sites. Lead Investigator: Ingrid Holm & Maureen Smith
21. A Literature Review of U.S. Individuals' Perspectives on Privacy, Trust, and Perceived Risks and Benefits in Biobanking. Lead Investigator: Nanibaa' Garrison (VU)
22. A Review of U.S. Individuals' Perspectives on Governance and Consent in Biobanking. Lead Investigator: Nanibaa' Garrison
23. A Highly Accurate Electronic Algorithm for the Classification of Asthma Severity in Children. Lead Investigator: Erik Hysinger (CHOP)
24. Epistatic gene-based interaction analyses for glaucoma in eMERGE network and NEIGHBOR consortium. Lead Investigator: Shefali Verma (PSU)
25. Exploring the genetic architecture of Age-Related Macular Degeneration (AMD) in the eMERGE network. Lead Investigator: Molly Hall (PSU)
26. Genome-wide Association Study of Gastroesophageal Reflux Disease (GERD) in Adult and Pediatric Populations. Lead Investigator: Patrick Sleiman (CHOP)
27. Genome-wide Association Study of Atopic Dermatitis in Adult and Pediatric Populations. Lead Investigator: Berta Almoguera (CHOP)
28. Genome-wide Association Study of Asthma in Adult and Pediatric Populations. Lead Investigator: Berta Almoguera (CHOP)



29. Genome-wide Association Study of Attention Deficit Hyperactivity Disorder (ADHD). Lead Investigator: John Connolly (CHOP)
30. Investigation of PCSK9 SNPs with LDL-C, BMI and risk of T2D. Lead Investigator: Brendan Keating (CHOP)
31. Phenotype transportability across Electronic Health Records. Lead Investigator: Joshua Denny (VU)
32. The COGENT consortium meta-analysis of blood pressure African ancestry cohorts. Lead Investigator: Dinga Velez Edwards (VU)
33. Genetic variation among 84 pharmacogenes from the PGRNSeq in the eMERGE Network. Lead Investigators: Will Bush (Case) & David Crosslin (GH/UW)
34. PheWAS analysis of a functional variant in CDHR3. Lead Investigator: Michael March (CHOP)
35. PheWAS analysis of homozygous deletions in GWAS data. Lead Investigator: Michael March (CHOP)
36. Using PheWAS to assess disease comorbidity and potential pleiotropy of genetic risk scores for rheumatoid arthritis. Lead Investigator: Robert Carroll (VU)
37. Association of rare and common variants in LDLR, HMGCR, NAT2, ABCA1, and APOA1 with plasma lipid levels: results from 9000 participants of the eMERGE Network. Lead Investigators: Daniel Kim (Michigan), Erin Austin (Mayo)
38. Exploring genetics and outcomes associated with acute kidney injury (AKI) using electronic health records and genomics. Lead Investigator: Girish N Nadkarni (Mt. Sinai)
39. PGRNseq and GWAS predictors of Methylphenidate (MPH) response. Lead Investigator: Tanya Froelich (CCHMC)
40. Discovery, Replication and Clinical Associations of Pathway-Based Trans-eQTL. Lead Investigator: Laura Wiley (VU)
41. Multiscale Analysis Of Influenza Host-Pathogen Interactions: Fluomics. Lead Investigator: Ellie Sang Sukerman (Northwestern)
42. Rare RYR1, CACNA1S variant annotation, exposure history, observed phenotypes in cases and controls. Lead Investigator: Senthilkumar Sadhasivam (CCHMC)
43. Variant Calling and Annotation for 82 known pharmacogenes (tentative). Lead Investigator: Will Bush (CC)
44. Characterizing the individual and shared genetic components of pheWAS phenotypes. Lead Investigator: Jonathan Mosley (VU)
45. Examining gene variants in eMERGE samples for association with uterine fibroids. Lead Investigator: Todd Edwards (VU)
46. A Phenome-wide Survey of the Phenotypic Effects of Neanderthal Admixture. Lead Investigator: Tony Capra (VU)
47. MVtest: a method to flexibly model the genetic determinants of trait variability. Lead Investigator: Todd Edwards (VU)
48. Autism Spectrum Disorders: Electronic Health Record Mining and Comorbidity Clustering. Lead Investigator: Todd Lingren (CCHMC)
49. Pediatric Providers are Poor at Identifying Severe Obesity in Young Children at Two Tertiary Pediatric Medical Centers. Lead Investigator: Cassandra Brady (CCHMC)
50. Developing an Algorithm to Detect Early Childhood Obesity Two Tertiary Pediatric Medical Centers. Lead Investigator: Todd Lingren, Vidhu Thaker (BCH)
51. Association of APOL1 G1/G2 risk alleles with metabolic and cardiovascular traits. Lead Investigator: Girish Nadkarni (Mt. Sinai) & Miriam Udler (BCH)
52. Development of a dynamic XML event-driven ophthalmologic data capture framework. Lead Investigator: Peggy Peissig (Marshfield)
53. Burden of structural variation and PheWAS. Lead Investigator: David Crosslin (GroupHealth)
54. Knowledge driven search for gene-gene interactions associated with hypothyroidism in the eMERGE network. Lead Investigator: Molly Hall (MC/EIRH/PSU)
55. Discovery and replication of genetic interactions for quantitative lipid traits. Lead Investigator: Marylyn Ritchie (MC/EIRH/PSU & CC)
56. Genetic Risk Scores for Complex Diseases in the eMERGE Network: Characterization and Predictive Abilities in Clinical Settings. Lead Investigator: Logan Dumitrescu (VU)
57. Phenotypes Seen in Cohorts with Rare Variants in Six PGRN-Seq (VIP) Genes also Identified by the ACMG as Priority Genes for Reporting Incidental Findings. Lead Investigator: Josh Denny (VU)
58. Chromosomal Anomalies that Affect Levels of White Blood Count (WBC) and its Differential. Lead Investigator: David Crosslin (UW)
59. Extracting the Quality of Prostate Cancer Care from Electronic Healthcare Records. Lead Investigator: Tina Hernandez Bousard
60. Practical Approaches to the Omic Chasm. Lead Investigator: Justin Starren (NU)
61. Developing Consents for Returning Pharmacogenomics Results: The eMERGE Experience. Lead Investigator: Maureen Smith (NU)
62. Portable Applications for Implementing Multi-Site Clinical NLP Algorithms. Lead Investigator: David Carrell (GroupHealth)
63. Evaluation of a Differentially Private Top-k SNP Publication Strategy. Lead Investigator: Mehmet Kuzu (External Collaborator); Brad Malin (VU)
64. Effective Use of Electronic Health Records to Identify Venous Thromboembolism: Results from the eMERGE Network. Lead Investigator: Jyoti Pathak (Mayo)
65. GWAS of Infection or Colonization with Community Associated Methicillin-Resistant Staphylococcus Aureus (CA-MRSA). Lead Investigator: Abel Kho (NU)
66. Genetic Risk Factors for Development of Diverticulitis. Lead Investigator: Abel Kho (NU)

67. Diverticulosis. Lead Investigator: Will Thompson (NU)
68. The Geographic Distribution of Colon Polyps. Lead Investigator: Will Thompson (NU)
69. Colon Polyps. Lead Investigator: Abel Kho (NU)
70. Genetic Variants Associated with Response to Heart Failure Treatment: The Electronic Medical Records and Genomics (eMERGE) Network. Lead Investigator: Sue Bielinski (Mayo)
71. Genome Wide Association of Risk of Heart Failure: The Electronic Medical Records and Genomics (eMERGE) Network. Lead Investigator: Sue Bielinski (Mayo)
72. Genetic Variation that Predicts Susceptibility to Clostridium Difficile. Lead Investigator: Josh Denny (VU)
73. Genome-wide study of resistant hypertension using existing genomic data and electronic medical records. Lead Investigator: Logan Dumitrescu (VU).
74. GWAS of Venous Thromboembolism (VTE) among White Americans. Lead Investigator: John Heit (Mayo)
75. GWAS of Venous Thromboembolism (VTE) among African-Americans. Lead Investigator: John Heit (Mayo)
76. Genome-Wide Association Study of Abdominal Aortic Aneurysms with Electronic Medical Record Phenotyping. Lead Investigators: Helena Kuivaniemi and Gerard Tromp (Geisinger)

## Site-Specific Manuscripts Published During Phase II

### BCH/CCHMC

1. Sadhasivam S, Zhang X, Chidambaran V, Mavi J, Pilipenko V, et al. *Novel associations between FAAH genetic variants and postoperative central opioid-related adverse effects*. Pharmacogenomics J. 2015 Oct;15(5):436–442. **PMCID: PMC4492912**
2. Ni Y, Wright J, Perentesis J, Lingren T, Deleger L, et al. *Increasing the efficiency of trial-patient matching: automated clinical trial eligibility pre-screening for pediatric oncology patients*. BMC Med Inform Decis Mak. 2015;15:28. **PMCID: PMC4407835**
3. Li Q, Spooner SA, Kaiser M, Lingren N, Robbins J, et al. *An end-to-end hybrid algorithm for automated medication discrepancy detection*. BMC Med Inform Decis Mak. 2015;15:37. **PMCID: PMC4427951**
4. Li Q, Kirkendall ES, Hall ES, Ni Y, Lingren T, et al. *Automated detection of medication administration errors in neonatal intensive care*. J Biomed Inform. 2015 Jul 17; **PMID: 26190267**
5. Zhai H, Brady P, Li Q, Lingren T, Ni Y, et al. *Developing and evaluating a machine learning based algorithm to predict the need of pediatric intensive care unit transfer for newly hospitalized children*. Resuscitation. 2014 Aug;85(8):1065–1071. **PMCID: PMC4087062**
6. Prows CA, Zhang X, Huth MM, Zhang K, Saldaña SN, et al. *Codeine-related adverse drug reactions in children following tonsillectomy: a prospective study*. Laryngoscope. 2014 May;124(5):1242–1250. **PMID: 24122716**
7. Namjou B, Ni Y, Harley ITW, Chepelev I, Cobb B, et al. *The effect of inversion at 8p23 on BLK association with lupus in Caucasian population*. PLoS ONE. 2014;9(12):e115614. **PMCID: PMC4278715**
8. Morris DL, Fernando MMA, Taylor KE, Chung SA, Nititham J, et al. *MHC associations with clinical and autoantibody manifestations in European SLE*. Genes Immun. 2014 Apr;15(4):210–217. **PMCID: PMC4102853**
9. Li Q, Melton K, Lingren T, Kirkendall ES, Hall E, et al. *Phenotyping for patient safety: algorithm development for electronic health record based automated adverse event and medical error detection in neonatal intensive care*. Journal of the American Medical Informatics Association. 2014 Sep 1;21(5):776–784. **PMCID: PMC4147599**
10. Kohane IS. *An Autism Case History to Review the Systematic Analysis of Large-Scale Data to Refine the Diagnosis and Treatment of Neuropsychiatric Disorders*. Biol Psychiatry. 2014 Jun 12; **PMCID: PMC4260993**
11. Kaufman KM, Linghu B, Szustakowski JD, Husami A, Yang F, et al. *Whole-Exome Sequencing Reveals Overlap Between Macrophage Activation Syndrome in Systemic Juvenile Idiopathic Arthritis and Familial Hemophagocytic Lymphohistiocytosis*. Arthritis & Rheumatology. 2014 Dec 1;66(12):3486–3495. **PMCID: PMC4321811**
12. Holm IA. *Clinical Management of Pediatric Genomic Testing*. Curr Genet Med Rep. 2014 Aug 9;2(4):212–215. **PMCID: PMC4226515**
13. Harley JB, Zoller EE. *Editorial: What Caused All These Troubles, Anyway? Epstein-Barr Virus in Sjögren’s Syndrome Reevaluated*. Arthritis & Rheumatology (Hoboken, NJ). 2014 Sep;66(9):2328–2330. **PMCID: PMC4203322**
14. Zhai H, Lingren T, Deleger L, Li Q, Kaiser M, et al. *Web 2.0-based crowdsourcing for high-quality gold standard development in clinical natural language processing*. J Med Internet Res. 2013 Apr 2;15(4):e73. **PMCID: PMC3636329**
15. Mangale D, Kariuki SN, Chrabot BS, Kumabe M, Kelly et al. *Familial aggregation of high tumor necrosis factor alpha levels in systemic lupus erythematosus*. Clin Dev Immunol. 2013 Sep 25;2013:267430. **PMCID: PMC3800640**
16. Li Q, Zhai H, Deleger L, Lingren T, Kaiser M, et al. *A sequence labeling approach to link medications and their attributes in clinical notes and clinical trial announcements for information extraction*. J Am Med Inform Assoc. 2013 Oct;20(5):915–921. **PMCID:**

## PMC3756265

17. Li Q, Deleger L, Lingren T, Zhai H, Kaiser M, et al. *Mining FDA drug labels for medical conditions*. BMC Med Inform Decis Mak. 2013 Apr 24;13:53. **PMCID: PMC3646673**
18. Deleger L, Brodzinski H, Zhai H, Li Q, Lingren T, et al. *Developing and evaluating an automated appendicitis risk stratification algorithm for pediatric patients in the emergency department*. J Am Med Inform Assoc. 2013 Dec;20(e2):e212–220. **PMCID: PMC3861926**

## CHOP

1. Sun F, Oristaglio J, Levy SE, Hakonarson H, Sullivan N, et al. *Genetic Testing for Developmental Disabilities, Intellectual Disability, and Autism Spectrum Disorder* [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2015 [cited 2016 Jan 8]. **PMID: 26158183**
2. Sgariglia F, Pedrini E, Bradfield JP, Bhatti TR, D'Adamo P, et al. *The type 2 diabetes associated rs7903146 T allele within TCF7L2 is significantly under-represented in Hereditary Multiple Exostoses: Insights into pathogenesis*. Bone. 2015 Mar 1;72:123–127. **PMCID: PMC4300120**
3. Roy SM, Chesi A, Mentch F, Xiao R, Chiavacci R, et al. *Body Mass Index (BMI) Trajectories in Infancy Differ by Population Ancestry and May Presage Disparities in Early Childhood Obesity*. J Clin Endocrinol Metab. 2015 Jan 30;jc20144028. **PMCID: PMC4399305**
4. Rodriguez S, Gaunt TR, Guo Y, Zheng J, Barnes MR, et al. *Lipids, obesity and gallbladder disease in women: insights from genetic studies using the cardiovascular gene-centric 50K SNP array*. Eur J Hum Genet. 2015 Apr 29; **PMID: 25920552**
5. Nead KT, Li A, Wehner MR, Neupane B, Gustafsson S, et al. *Contribution of common non-synonymous variants in PCSK1 to body mass index variation and risk of obesity: a systematic review and meta-analysis with evidence from up to 331 175 individuals*. Hum Mol Genet. 2015 Jun 15;24(12):3582–3594. **PMCID: PMC4498155**
6. Merikangas KR, Calkins ME, Burstein M, He J-P, Chiavacci R, et al. *Comorbidity of physical and mental disorders in the neurodevelopmental genomics cohort study*. Pediatrics. 2015 Apr;135(4):e927–938. **PMCID: PMC4379458**
7. Menezes MJ, Guo Y, Zhang J, Riley LG, Cooper ST, et al. *Mutation in mitochondrial ribosomal protein S7 (MRPS7) causes congenital sensorineural deafness, progressive hepatic and renal failure and lactic acidemia*. Hum Mol Genet. 2015 Jan 2; **PMID: 25556185**
8. Mancini C, Nassani S, Guo Y, Chen Y, Giorgio E, et al. *Adult-onset autosomal recessive ataxia associated with neuronal ceroid lipofuscinosis type 5 gene (CLN5) mutations*. J Neurol. 2015 Jan;262(1):173–178. **PMID: 25359263**
9. Maier R, Moser G, Chen G-B, Ripke S, Cross-Disorder et al. *Joint analysis of psychiatric disorders increases accuracy of risk prediction for schizophrenia, bipolar disorder, and major depressive disorder*. Am J Hum Genet. 2015 Feb 5;96(2):283–294. **PMCID: PMC4320268**
10. Maggadottir SM, Li J, Glessner JT, Li YR, Wei Z, et al. *Rare variants at 16p11.2 are associated with common variable immunodeficiency*. J Allergy Clin Immunol. 2015 Jun;135(6):1569–1577. **PMCID: PMC4461447**
11. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, et al. *Genetic studies of body mass index yield new insights for obesity biology*. Nature. 2015 Feb 12;518(7538):197–206. **PMCID: PMC4382211**
12. Li YR, Li J, Zhao SD, Bradfield JP, Mentch FD, et al. *Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases*. Nat Med. 2015 Sep;21(9):1018–1027. **PMID: 26301688**
13. Li W-D, Jiao H, Wang K, Yang F, Grant SFA, et al. *Pathway-Based Genome-wide Association Studies Reveal That the Rac1 Pathway Is Associated with Plasma Adiponectin Levels*. Sci Rep. 2015;5:13422. **PMCID: PMC4642532**
14. Li J, Jørgensen SF, Maggadottir SM, Bakay M, Warnatz K, et al. *Association of CLEC16A with human common variable immunodeficiency disorder and role in murine B cells*. Nat Commun. 2015;6:6804. **PMCID: PMC4444044**
15. Li J, Fung I, Glessner JT, Pandey R, Wei Z, et al. *Copy Number Variations in CTNNA3 and RBFox1 Associate with Pediatric Food Allergy*. J Immunol. 2015 Aug 15;195(4):1599–1607. **PMID: 26188062**
16. Kruszka P, Li D, Harr MH, Wilson NR, Swarr D, et al. *Mutations in SPECC1L, encoding sperm antigen with calponin homology and coiled-coil domains 1-like, are found in some cases of autosomal dominant Opitz G/BBB syndrome*. J Med Genet. 2015 Feb;52(2):104–110. **PMCID: PMC4393015**
17. International Genomics of Alzheimer's Disease Consortium (IGAP). *Convergent genetic and expression data implicate immunity in Alzheimer's disease*. Alzheimers Dement. 2015 Jun;11(6):658–671. **PMCID: PMC4672734**
18. Guo Y, Menezes MJ, Menezes MP, Liang J, Li D, et al. *Delayed diagnosis of congenital myasthenia due to associated mitochondrial enzyme defect*. Neuromuscul Disord. 2015 Mar;25(3):257–261. **PMID: 25557462**
19. Connolly JJ, Glessner JT, Elia J, Hakonarson H. *ADHD & Pharmacotherapy: Past, Present and Future: A Review of the Changing Landscape of Drug Therapy for Attention Deficit Hyperactivity Disorder*. Ther Innov Regul Sci. 2015 Sep;49(5):632–642. **PMCID: PMC4564067**
20. Chang X, Li J, Guo Y, Wei Z, Mentch FD, et al. *Genome-wide association study of serum minerals levels in children of different ethnic background*. PLoS ONE. 2015;10(4):e0123499. **PMCID: PMC4401557**
21. Brehm JM, Ramratnam SK, Tse SM, Croteau-Chonka DC, Pino-Yanes M, et al. *Stress and Bronchodilator Response in Children with Asthma*. Am J Respir Crit Care Med. 2015 Apr 28; **PMCID: PMC4511425**

22. Zhang Y, Ghosh S, Hakonarson H. *Dynamic Bayesian testing of sets of variants in complex diseases*. Genetics. 2014 Nov;198(3):867–878. **PMCID: PMC4224176**
23. Zamzow J, Culnan E, Spiers M, Calkins M, Satterthwaite T, et al. *B-37The Relationship between Body Mass Index and Executive Function from Late Childhood through Adolescence*. Arch Clin Neuropsychol. 2014 Sep;29(6):550. **PMID: 25176787**
24. White PS, Xie HM, Werner P, Glessner J, Latney B, et al. *Analysis of chromosomal structural variation in patients with congenital left-sided cardiac lesions*. Birth Defects Res Part A Clin Mol Teratol. 2014 Jul 26; **PMID: 25066379**
25. van der Valk RJP, Kreiner-Møller E, Kooijman MN, Guxens M, Stergiakouli E, et al. *A novel common variant in DCST2 is associated with length in early life and height in adulthood*. Hum Mol Genet. 2014 Oct 3; **PMCID: PMC4447786**
26. Swerdlow DI, Preiss D, Kuchenbaecker KB, Holmes MV, Engmann JEL, et al. *HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials*. Lancet. 2014 Sep 24; **PMCID: PMC4322187**
27. St Pourcain B, Skuse DH, Mandy WP, Wang K, Hakonarson H, et al. *Variability in the common genetic architecture of social-communication spectrum phenotypes during childhood and adolescence*. Mol Autism. 2014 Feb 24;5:18. **PMCID: PMC3940728**
28. St Pourcain B, Haworth CMA, Davis OSP, Wang K, Timpson NJ, et al. *Heritability and genome-wide analyses of problematic peer relationships during childhood and adolescence*. Hum Genet. 2014 Dec 17; **PMCID: PMC4424375**
29. Soleimanpour SA, Gupta A, Bakay M, Ferrari AM, Groff DN, et al. *The diabetes susceptibility gene Clec16a regulates mitophagy*. Cell. 2014 Jun 19;157(7):1577–1590. **PMCID: PMC4184276**
30. Sleiman PMA, Wang M-L, Cianferoni A, Aceves S, Gonsalves N, et al. *GWAS identifies four novel eosinophilic esophagitis loci*. Nat Commun [Internet]. 2014 Nov 19 [cited 2015 Mar 27];5. **PMCID: PMC4238044**
31. Scott D, Cook-Sather JL. *Modulatory effects of TAOK3 variants on morphine requirement in acute postoperative pain: An early genome wide association study contribution to the field of pediatric pain*. Pain. 2014;155(11). **PMID: 25456570**
32. Robinson EB, Kirby A, Ruparel K, Yang J, McGrath L, et al. *The genetic architecture of pediatric cognitive abilities in the Philadelphia Neurodevelopmental Cohort*. Mol Psychiatry. 2014 Jul 15; **PMCID: PMC4294962**
33. Purkey MT, Li J, Mentch F, Grant SFA, Desrosiers M, et al. *Genetic Variation in Genes Encoding Airway Epithelial Potassium Channels Is Associated with Chronic Rhinosinusitis in a Pediatric Population*. PLoS One [Internet]. 2014 Mar 3 [cited 2014 Sep 11];9(3). **PMCID: PMC3940609**
34. Prokudin I, Li D, He S, Guo Y, Goodwin L, et al. *Value of whole exome sequencing for syndromic retinal dystrophy diagnosis in young patients*. Clin Experiment Ophthalmol. 2014 Jul 24; **PMID: 25060287**
35. Pizzino A, Pierson TM, Guo Y, Helman G, Fortini S, et al. *TUBB4A de novo mutations cause isolated hypomyelination*. Neurology. 2014 Sep 2;83(10):898–902. **PMCID: PMC4153852**
36. Picard M, Zhang J, Hancock S, Derbeneva O, Golhar R, et al. *Progressive increase in mtDNA 3243A>G heteroplasmy causes abrupt transcriptional reprogramming*. Proc Natl Acad Sci USA. 2014 Sep 23;111(38):E4033–4042. **PMCID: PMC4183335**
37. Pellegrino R, Kavakli IH, Goel N, Cardinale CJ, Dinges DF, et al. *A novel BHLHE41 variant is associated with short sleep and resistance to sleep deprivation in humans*. Sleep. 2014;37(8):1327–1336. **PMCID: PMC4096202**
38. Naj AC, Jun G, Reitz C, Kunkle BW, Perry W, et al. *Effects of multiple genetic Loci on age at onset in late-onset Alzheimer disease: a genome-wide association study*. JAMA Neurol. 2014 Nov 1;71(11):1394–1404. **PMCID: PMC4314944**
39. Moore TM, Reise SP, Gur RE, Hakonarson H, Gur RC. *Psychometric Properties of the Penn Computerized Neurocognitive Battery*. Neuropsychology. 2014 Sep 1; **PMCID: PMC4345134**
40. Mitchell LE, Agopian AJ, Bhalla A, Glessner JT, Kim CE, et al. *Genome-wide association study of maternal and inherited effects on left-sided cardiac malformations*. Hum Mol Genet. 2014 Aug 18; **PMCID: PMC4326325**
41. Mestdagh P, Hartmann N, Baeriswyl L, Andreasen D, Bernard N, et al. *Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study*. Nat Methods. 2014 Aug;11(8):809–815. **PMID: 24973947**
42. Matsunami N, Hensel CH, Baird L, Stevens J, Otterud B, et al. *Identification of rare DNA sequence variants in high-risk autism families and their prevalence in a large case/control population*. Mol Autism. 2014 Jan 27;5(1):5. **PMCID: PMC4098669**
43. Holmes MV, Dale CE, Zuccolo L, Silverwood RJ, Guo Y, et al. *Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data*. BMJ. 2014;349:g4164. **PMCID: PMC4091648**
44. Hadley D, Wu Z-L, Kao C, Kini A, Mohamed-Hadley A, et al. *The impact of the metabotropic glutamate receptor and other gene family interaction networks on autism*. Nat Commun. 2014;5:4074. **PMCID: PMC4059929**
45. Guo Y, Kartawinata M, Li J, Pickett HA, Teo J, et al. *Inherited bone marrow failure associated with germline mutation of ACD, the gene encoding telomere protein TPP1*. Blood. 2014 Oct 30;124(18):2767–2774. **PMCID: PMC4215308**
46. Glessner JT, Bick AG, Ito K, Homsy JG, Rodriguez-Murillo L, et al. *Increased frequency of de novo copy number variants in congenital heart disease by integrative analysis of single nucleotide polymorphism array and exome sequence data*. Circ Res. 2014 Oct 24;115(10):884–896. **PMCID: PMC4209190**
47. Escott-Price V, Bellenguez C, Wang L-S, Choi S-H, Harold D, et al. *Gene-wide analysis detects two new susceptibility genes for Alzheimer's disease*. PLoS ONE. 2014;9(6):e94661. **PMCID: PMC4055488**
48. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, et al. *Synaptic, transcriptional and chromatin genes disrupted in autism*. Nature. 2014 Nov 13;515(7526):209–215. **PMCID: PMC4402723**

49. Culnan E, Zamzow J, Spiers M, Calkins M, Satterthwaite T, et al. *B-36 Relationships between Body Mass Index and Social Cognition among 8-19 Year-Olds*. Arch Clin Neuropsychol. 2014 Sep;29(6):550. **PMID: 25176786**
50. Cook-Sather SD, Li J, Goebel TK, Sussman EM, Rehman MA, et al. *TAOK3, a novel genome-wide association study locus associated with morphine requirement and postoperative pain in a retrospective pediatric day surgery population*. Pain. 2014 Sep;155(9):1773–1783. **PMCID: PMC4157963**
51. Connolly JJ, Hakonarson H. *Etiology of autism spectrum disorder: a genomics perspective*. Curr Psychiatry Rep. 2014 Nov;16(11):501. **PMID: 25212713**
52. Calkins ME, Moore TM, Merikangas KR, Burstein M, Satterthwaite TD, et al. *The psychosis spectrum in a young U.S. community sample: findings from the Philadelphia Neurodevelopmental Cohort*. World Psychiatry. 2014 Oct;13(3):296–305. **PMCID: PMC4219071**
53. Banerjee A, Wang H-Y, Borgmann-Winter KE, MacDonald ML, Kaprielian H, et al. *Src kinase as a mediator of convergent molecular abnormalities leading to NMDAR hypoactivity in schizophrenia*. Mol Psychiatry. 2014 Oct 21; **PMID: 25330739**
54. Almoguera B, He S, Corton M, Jose PF-S, Blanco-Kelly F, et al. *Expanding the phenotype of PRPS1 syndromes in females: neuropathy, hearing loss and retinopathy*. Orphanet Journal of Rare Diseases. 2014 Dec 10;9(1):190. **PMCID: PMC4272780**
55. St Pourcain B, Whitehouse AJO, Ang WQ, Warrington NM, Glessner JT, et al. *Common variation contributes to the genetic architecture of social communication traits*. Mol Autism. 2013 Sep 18;4(1):34. **PMCID: PMC3853437**
56. Shi L, Zhang X, Golhar R, Otieno FG, He M, et al. *Whole-genome sequencing in an autism multiplex family*. Mol Autism. 2013 Apr 18;4(1):8. **PMCID: PMC3642023**
57. Ong BA, Li J, McDonough JM, Wei Z, Kim C, et al. *Gene network analysis in a pediatric cohort identifies novel lung function genes*. PLoS ONE. 2013 Sep 2;8(9):e72899. **PMCID: PMC3759429**
58. Mortensen LJ, Kreiner-Møller E, Hakonarson H, Bønnelykke K, Bisgaard H. *The PCDH1-gene and asthma in early childhood*. Eur Respir J. 2013 Aug 29; **PMID: 23988763**
59. Monda KL, Chen GK, Taylor KC, Palmer C, Edwards TL, et al. *A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry*. Nat Genet. 2013 Jun;45(6):690–696. **PMCID: PMC3694490**
60. Mitchell JA, Hakonarson H, Rebbeck TR, Grant SFA. *Obesity-susceptibility loci and the tails of the pediatric BMI distribution*. Obesity (Silver Spring) [Internet]. 2013 Feb 14. **PMCID: PMC3661695**
61. March ME, Sleiman PM, Hakonarson H. *Genetic polymorphisms and associated susceptibility to asthma*. Int J Gen Med. 2013 Apr 17;6:253–265. **PMCID: PMC3636804**
62. Keller M, Glessner J, Resnick E, Perez E, Chapel H, et al. *Burden of Copy Number Variation in Common Variable Immunodeficiency*. Clin Exp Immunol. 2013 Dec 13; **PMCID: PMC4089176**
63. Kasperaviciute D, Catarino CB, Matarin M, Leu C, Novy J, et al. *Epilepsy, hippocampal sclerosis and febrile seizures linked by common genetic variation around SCN1A*. Brain. 2013 Oct;136(Pt 10):3140–3150. **PMCID: PMC3784283**
64. Hamshere ML, Langley K, Martin J, Agha SS, Stergiakouli E, et al. *High loading of polygenic risk for ADHD in children with comorbid aggression*. Am J Psychiatry. 2013 Aug 1;170(8):909–916. **PMCID: PMC3935265**
65. Guo Y, Lanktree MB, Taylor KC, Hakonarson H, Lange LA, et al. *Gene-centric meta-analyses of 108 912 individuals confirm known body mass index loci and reveal three novel signals*. Hum Mol Genet. 2013 Jan 1;22(1):184–201. **PMCID: PMC3522401**
66. Glessner JT, Li J, Hakonarson H. *ParseCNV integrative copy number variation association software with quality tracking*. Nucleic Acids Res [Internet]. 2013 Jan 4. **PMCID: PMC3597648**
67. Cardinale CJ, Kelsen JR, Baldassano RN, Hakonarson H. *Impact of exome sequencing in inflammatory bowel disease*. World J Gastroenterol. 2013 Oct 28;19(40):6721–6729. **PMCID: PMC3812471**
68. Asselbergs FW, Guo Y, van Iperen EPA, Sivapalaratnam S, Tragante V, et al. *Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci*. Am J Hum Genet. 2012 Nov 2;91(5):823–838. **PMCID: PMC3487124**

## Geisinger

1. Fragoulakis V, Mitropoulou C, Williams M, Patrinos G. *Economic Evaluation in Genomic Medicine*.
2. Williams MS. *Perspectives on what is needed to implement genomic medicine*. Mol Genet Genomic Med. 2015 May;3(3):155–159. **PMCID: PMC4444156**
3. Smith-Packard B, Myers SM, Williams MS. *Girls with Seizures Due to the c.320A>G Variant in ALG13 Do Not Show Abnormal Glycosylation Pattern on Standard Testing*. JIMD Rep. 2015 Mar 3; **PMCID: PMC4486277**
4. Williams MS. *Genomic medicine implementation: Learning by example*. Am J Med Genet. 2014 Mar 1;166(1):8–14. **PMID: 24616394**
5. Williams JL, Faucett WA, Smith-Packard B, Wagner M, Williams MS. *An Assessment of Time Involved in Pre-test Case Review and Counseling for a Whole Genome Sequencing Clinical Research Program*. J Genet Couns. 2014 Feb 27; **PMCID: PMC4090811**
6. Wade JE, Ledbetter DH, Williams MS. *Implementation of genomic medicine in a health care delivery system: A value proposition?* Am J Med Genet. 2014 Mar 1;166(1):112–116. **PMID: 24619641**

7. Snyder SR, Mitropoulou C, Patrinos GP, Williams MS. *Economic Evaluation of Pharmacogenomics: A Value-Based Approach to Pragmatic Decision Making in the Face of Complexity*. Public Health Genomics. 2014 Sep 26; PMID: **25278172**
8. Smelser DT, Tromp G, Elmore JR, Kuivaniemi H, Franklin DP, et al. *Population risk factor estimates for abdominal aortic aneurysm from electronic medical records: a case control study*. BMC Cardiovasc Disord. 2014;14:174. PMID: **PMC4269847**
9. Segal MM, Abdellateef M, El-Hattab AW, Hilbush BS, De La Vega FM, et al. *Clinical Pertinence Metric Enables Hypothesis-Independent Genome-Phenome Analysis for Neurologic Diagnosis*. J Child Neurol. 2014 Aug 24; PMID: **PMC4339658**
10. Schully SD, Lam TK, Dotson WD, Chang CQ, Aronson N, et al. *Evidence synthesis and guideline development in genomic medicine: current status and future prospects*. Genet Med. 2014 Jun 19; PMID: **PMC4272332**
11. Lipsky EB, King BR, Tromp G. *Node-Oriented Workflow (NOW): A Command Template Workflow Management Tool for High Throughput Data Analysis Pipelines*. Journal of Datamining in Genomics & Proteomics. 2014;5(155).
12. Kuivaniemi H, Ryer EJ, Elmore JR, Hinterseher I, Smelser DT, et al. *Update on abdominal aortic aneurysm research: from clinical to genetic studies*. Scientifica (Cairo). 2014;2014:564734. PMID: **PMC4009235**
13. Caudle KE, Klein TE, Hoffman JM, Müller DJ, Whirl-Carrillo M, et al. *Incorporation of Pharmacogenomics into Routine Clinical Practice: the Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline Development Process*. Curr Drug Metab [Internet]. 2014 Jan 30. PMID: **PMC3977533**
14. Brownstein CA, Beggs AH, Homer N, Merriman B, Yu TW, et al. *An international effort towards developing standards for best practices in analysis, interpretation and reporting of clinical genome sequencing results in the CLARITY Challenge*. Genome Biol. 2014;15(3):R53. PMID: **PMC4073084**
15. Bock JA, Fairley KJ, Smith RE, Maeng DD, Pitcavage JM, et al. *Cost-Effectiveness of IL28B Genotype-Guided Protease Inhibitor Triple Therapy versus Standard of Care Treatment in Patients with Hepatitis C Genotypes 2 or 3 Infection*. Public Health Genomics. 2014 Sep 18; PMID: **PMC4275401**
16. Wang X, Morris NJ, Zhu X, Elston RC. *A variance component based multi-marker association test using family and unrelated data*. BMC Genet. 2013 Mar 4;14:17. PMID: **PMC3614458**
17. Sun X, Elston R, Morris N, Zhu X. *What Is the Significance of Difference in Phenotypic Variability across SNP Genotypes?* Am J Hum Genet [Internet]. 2013 Jul 31. PMID: **PMC3738833**
18. Muir AJ, Gong L, Johnson SG, Lee MTM, Williams MS, et al. *Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for IFNL3 (IL28B) Genotype and PEG Interferon- $\alpha$ -Based Regimens*. Clin Pharmacol Ther [Internet]. 2013 Oct 4; PMID: **PMC3904555**
19. Jones GT, Bown MJ, Gretarsdottir S, Romaine SPR, Helgadottir A, et al. *A sequence variant associated with sortilin-1 (SORT1) on 1p13.3 is independently associated with abdominal aortic aneurysm*. Hum Mol Genet [Internet]. 2013 Apr 9; PMID: **PMC3690970**
20. Bock J, Fairley K, Smith R, Maeng D, Pitcavage J, et al. *Cost-Effectiveness of Interleukin 28B Genotype-Guided Protease Inhibitor Therapy in Treatment-Naïve Patients with Hepatitis C Virus Genotype 2 or 3*. Clin Med Res. 2013 Sep;11(3):142. PMID: **PMC3788547**
21. Zhu X, Feng T, Elston RC. *Linkage-disequilibrium-based binning misleads the interpretation of genome-wide association studies*. Am J Hum Genet. 2012 Nov 2;91(5):965–968; author reply 969–970. PMID: **PMC3487138**
22. Wang X, Morris NJ, Schaid DJ, Elston RC. *Power of single- vs. multi-marker tests of association*. Genet Epidemiol. 2012 Jul;36(5):480–487. PMID: **PMC3708310**
23. Lu Q, Wei C, Ye C, Li M, Elston RC. *A likelihood ratio-based Mann-Whitney approach finds novel replicable joint gene action for type 2 diabetes*. Genet Epidemiol. 2012 Sep;36(6):583–593. PMID: **PMC3634342**

## GroupHealth/UW

1. Trinidad SB, Ludman EJ, Hopkins S, James RD, Hoeft TJ, et al. *Community dissemination and genetic research: Moving beyond results reporting*. Am J Med Genet A. 2015 Jul;167(7):1542–1550. PMID: **PMC4478235**
2. Trinidad SB, Coffin TB, Fullerton SM, Ralston J, Jarvik GP, et al. *“Getting off the Bus Closer to Your Destination”: Patients’ Views about Pharmacogenetic Testing*. Perm J. 2015;19(3):21–27. PMID: **PMC4500477**
3. Regier DA, Peacock SJ, Pataky R, van der Hoek K, Jarvik GP, et al. *Societal preferences for the return of incidental findings from clinical genomic sequencing: a discrete-choice experiment*. CMAJ. 2015 Apr 7;187(6):E190–197. PMID: **PMC4387060**
4. Gallego CJ, Shirts BH, Bennette CS, Guzauskas G, Amendola LM, et al. *Next-Generation Sequencing Panels for the Diagnosis of Colorectal Cancer and Polyposis Syndromes: A Cost-Effectiveness Analysis*. J Clin Oncol. 2015 Jun 20;33(18):2084–2091. PMID: **PMC4461806**
5. Carrell DS, Cronkite D, Palmer RE, Saunders K, Gross DE, et al. *Using natural language processing to identify problem usage of prescription opioids*. Int J Med Inform. 2015 Dec;84(12):1057–1064. PMID: **26456569**
6. Amendola LM, Dorschner MO, Robertson PD, Salama JS, Hart R, et al. *Actionable exomic incidental findings in 6503 participants*:

- challenges of variant classification*. Genome Res. 2015 Mar;25(3):305–315. **PMCID: PMC4352885**
7. Sherva R, Tripodis Y, Bennett DA, Chibnik LB, Crane PK, et al. *Genome-wide association study of the rate of cognitive decline in Alzheimer's disease*. Alzheimers Dement. 2014 Jan;10(1):45–52. **PMCID: PMC3760995**
  8. Burke W, Evans BJ, Jarvik GP. *Return of results: Ethical and legal distinctions between research and clinical care*. Am J Med Genet C Semin Med Genet. 2014 Mar;166(1):105–111. **PMCID: PMC4078921**
  9. Benitez BA, Jin SC, Guerreiro R, Graham R, Lord J, et al. *Missense variant in TREML2 protects against Alzheimer's disease*. Neurobiol Aging. 2014 Jun;35(6):1510.e19–26. **PMCID: PMC3961557**
  10. Shirts BH, Jacobson A, Jarvik GP, Browning BL. *Large numbers of individuals are required to classify and define risk for rare variants in known cancer risk genes*. Genet Med. 2013 Dec 19; **PMCID: PMC4063879**
  11. Ridge PG, Mukherjee S, Crane PK, Kauwe JSK, Alzheimer's Disease Genetics Consortium. *Alzheimer's disease: analyzing the missing heritability*. PLoS ONE. 2013 Nov 7;8(11):e79771. **PMCID: PMC3820606**
  12. Reitz C, Mayeux R, Alzheimer's Disease Genetics Consortium. *TREM2 and neurodegenerative disease*. N Engl J Med. 2013 Oct 17;369(16):1564–1565. **PMCID: PMC3980568**
  13. Reitz C, Jun G, Naj A, Rajbhandary R, Vardarajan BN, et al. *Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E ε4, and the risk of late-onset Alzheimer disease in African Americans*. JAMA. 2013 Apr 10;309(14):1483–1492. **PMCID: PMC3667653**
  14. Reitz C, Tosto G, Vardarajan B, Rogaeva E, Ghani M, et al. *Independent and epistatic effects of variants in VPS10-d receptors on Alzheimer disease risk and processing of the amyloid precursor protein (APP)*. Transl Psychiatry. 2013 May 14;3:e256. **PMCID: PMC3669917**
  15. Miyashita A, Koike A, Jun G, Wang L-S, Takahashi S, et al. *SORL1 is genetically associated with late-onset Alzheimer's disease in Japanese, Koreans and Caucasians*. PLoS ONE. 2013 Apr 2;8(4):e58618. **PMCID: PMC3614978**
  16. Lin C-F, Valladares O, Childress DM, Klevak E, Geller ET, et al. *DRAW+SneakPeek: analysis workflow and quality metric management for DNA-seq experiments*. Bioinformatics. 2013 Oct 1;29(19):2498–2500. **PMCID: PMC3777113**
  17. Larson EB. *Building trust in the power of "big data" research to serve the public good*. JAMA. 2013 Jun 19;309(23):2443–2444. **PMID: 23780455**
  18. Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, et al. *Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease*. Nat Genet. 2013 Dec;45(12):1452–1458. **PMCID: PMC3896259**
  19. Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, et al. *Variant of TREM2 associated with the risk of Alzheimer's disease*. N Engl J Med. 2013 Jan 10;368(2):107–116. **PMCID: PMC3677583**
  20. Holton P, Ryten M, Nalls M, Trabzuni D, Weale ME, et al. *Initial assessment of the pathogenic mechanisms of the recently identified Alzheimer risk Loci*. Ann Hum Genet. 2013 Mar;77(2):85–105. **PMCID: PMC3578142**
  21. Cruchaga C, Kauwe JSK, Harari O, Jin SC, Cai Y, et al. *GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease*. Neuron. 2013 Apr 24;78(2):256–268. **PMCID: PMC3664945**
  22. Burke W, Matheny Antommara AH, Bennett R, Botkin J, Clayton EW, et al. *Recommendations for returning genomic incidental findings? We need to talk!* Genet Med. 2013 Nov;15(11):854–859. **PMCID: PMC3832423**
  23. Zou F, Chai HS, Younkin CS, Allen M, Crook J, et al. *Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants*. PLoS Genet. 2012;8(6):e1002707. **PMCID: PMC3369937**
  24. Whitcomb DC, LaRusch J, Krasinskas AM, Klei L, Smith JP, et al. *Common genetic variants in the CLDN2 and PRSS1-PRSS2 loci alter risk for alcohol-related and sporadic pancreatitis*. Nat Genet. 2012 Dec;44(12):1349–1354. **PMCID: PMC3510344**
  25. Trinidad SB, Fullerton SM, Bares JM, Jarvik GP, Larson EB, et al. *Informed Consent in Genome-Scale Research: What Do Prospective Participants Think?* AJOB Primary Research. 2012;3(3):3–11. **PMCID: PMC3593675**
  26. Schellenberg GD, Montine TJ. *The genetics and neuropathology of Alzheimer's disease*. Acta Neuropathol. 2012 Sep;124(3):305–323. **PMCID: PMC3708460**
  27. Jun G, Vardarajan BN, Buros J, Yu C-E, Hawk MV, et al. *Comprehensive search for Alzheimer disease susceptibility loci in the APOE region*. Arch Neurol. 2012 Oct;69(10):1270–1279. **PMCID: PMC3579659**
  28. Floyd JS, Heckbert SR, Weiss NS, Carrell DS, Psaty BM. *Use of administrative data to estimate the incidence of statin-related rhabdomyolysis*. JAMA. 2012 Apr 18;307(15):1580–1582. **PMCID: PMC3483067**
  29. Coppola G, Chinnathambi S, Lee JJ, Dombroski BA, Baker MC, et al. *Evidence for a role of the rare p.A152T variant in MAPT in increasing the risk for FTD-spectrum and Alzheimer's diseases*. Hum Mol Genet. 2012 Aug 1;21(15):3500–3512. **PMCID: PMC3392107**
  30. Beskow LM, Burke W, Fullerton SM, Sharp RR. *Offering aggregate results to participants in genomic research: opportunities and challenges*. Genet Med. 2012 Apr;14(4):490–496. **PMCID: PMC3612423**
  31. Allen M, Zou F, Chai HS, Younkin CS, Crook J, et al. *Novel late-onset Alzheimer disease loci variants associate with brain gene expression*. Neurology. 2012 Jul 17;79(3):221–228. **PMCID: PMC3398432**

1. Okula A, Wallace J, Peissig P, McCarty C, Brilliant M, et al. *Knowledge driven binning and PHEWAS analysis in Marshfield Personalized Medicine Research Project Using Biobin*. (In Press)
2. Zhang W, Yu Y, Hertwig F, Thierry-Mieg J, Zhang W, et al. *Comparison of RNA-seq and microarray-based models for clinical endpoint prediction*. *Genome Biol* [Internet]. 2015 [cited 2016 Jan 8];16(1). **PMCID: PMC4506430**
3. Springelkamp H, Mishra A, Hysi PG, Gharahkhani P, Höhn R, et al. *Meta-analysis of Genome-Wide Association Studies Identifies Novel Loci Associated With Optic Disc Morphology*. *Genet Epidemiol*. 2015 Mar;39(3):207–216. **PMCID: PMC4480365**
4. Shukla SK, Murali NS, Brilliant MH. *Personalized medicine going precise: from genomics to microbiomics*. *Trends Mol Med*. 2015 Aug;21(8):461–462. **PMCID: PMC4530069**
5. Schrodi SJ, DeBarber A, He M, Ye Z, Peissig P, et al. *Prevalence estimation for monogenic autosomal recessive diseases using population-based genetic data*. *Hum Genet*. 2015 Jun;134(6):659–669. **PMID: 25893794**
6. O'Brien SE, Schrodi SJ, Ye Z, Brilliant MH, Virani SS, et al. *Differential Lipid Response to Statins is Associated with Variants in the BUD13-APOA5 Gene Region*. *J Cardiovasc Pharmacol*. 2015 Apr 20; **PMID: 25900265**
7. Naidoo D, Wu AC, Brilliant MH, Denny J, Ingram C, et al. *A polymorphism in HLA-G modifies statin benefit in asthma*. *Pharmacogenomics J*. 2015 Jun;15(3):272–277. **PMCID: PMC4379135**
8. McCarty CA, Berg R, Waudby C, Foth W, Kitchner T, et al. *Long-Term Recall of Elements of Informed Consent: A Pilot Study Comparing Traditional and Computer-Based Consenting*. *IRB*. 2015 Feb;37(1):1–5. **PMID: 26247077**
9. He M, Person TN, Hebbbring SJ, Heinzen E, Ye Z, et al. *SeqHBase: a big data toolset for family based sequencing data analysis*. *J Med Genet*. 2015 Apr;52(4):282–288. **PMCID: PMC4382803**
10. Ye Z, Mayer J, Ivacic L, Zhou Z, He M, et al. *Phenome-wide association studies (PheWASs) for functional variants*. *Eur J Hum Genet*. 2014 Jul 30; **PMID: 25074467**
11. Springelkamp H, Höhn R, Mishra A, Hysi PG, Khor C-C, et al. *Meta-analysis of genome-wide association studies identifies novel loci that influence cupping and the glaucomatous process*. *Nat Commun*. 2014;5:4883. **PMCID: PMC4199103**
12. Seqc/Maqc-Iii Consortium. *A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium*. *Nat Biotech*. 2014 Sep;32(9):903–914. **PMCID: PMC4321899**
13. McCarty CA, Huggins W, Aiello AE, Bilder RM, Hariri A, et al. *PhenX RISING: real world implementation and sharing of PhenX measures*. *BMC Med Genomics*. 2014;7:16. **PMCID: PMC3994539**
14. McCarty CA, Berg R, Rottschreit CM, Waudby CJ, Kitchner T, et al. *Validation of PhenX measures in the personalized medicine research project for use in gene/environment studies*. *BMC Med Genomics*. 2014 Jan 14;7(1):3. **PMCID: PMC3896802**
15. Liu Y, Garrett ME, Yaspan BL, Bailey JC, Loomis SJ, et al. *DNA Copy Number Variants of Known Glaucoma Genes in Relation to Primary Open-Angle Glaucoma*. *Invest Ophthalmol Vis Sci*. 2014 Dec;55(12):8251–8258. **PMCID: PMC4271633**
16. Land ME, Cooper RF, Young J, Berg E, Kitchner T, et al. *Cone structure in subjects with known genetic relative risk for AMD*. *Optom Vis Sci*. 2014 Aug;91(8):939–949. **PMCID: PMC4111779**
17. Hall MA, Dudek SM, Goodloe R, Crawford DC, Pendergrass SA, et al. *Environment-wide association study (EWAS) for type 2 diabetes in the Marshfield Personalized Medicine Research Project Biobank*. *Pac Symp Biocomput*. 2014;200–211. **PMCID: PMC4037237**
18. Barrie ES, Weinshenker D, Verma A, Pendergrass S, Lange L, et al. *Regulatory Polymorphisms in Human DBH Affect Peripheral Gene Expression and Sympathetic Activity*. *Circ Res*. 2014 Oct 17; **PMCID: PMC4258174**
19. Bailey JNC, Yaspan BL, Pasquale LR, Hauser MA, Kang JH, et al. *Hypothesis-independent pathway analysis implicates GABA and acetyl-CoA metabolism in primary open-angle glaucoma and normal-pressure glaucoma*. *Hum Genet*. 2014 Oct;133(10):1319–1330. **PMCID: PMC4273559**
20. Samwald M, Freimuth R, Luciano JS, Lin S, Powers RL, et al. *An RDF/OWL knowledge base for query answering and decision support in clinical pharmacogenetics*. *Stud Health Technol Inform*. 2013;192:539–542. **PMCID: PMC4028612**
21. Pendergrass SA, Verma SS, Holzinger ER, Moore CB, Wallace J, et al. *Next-generation analysis of cataracts: determining knowledge driven gene-gene interactions using biofilter, and gene-environment interactions using the phenx toolkit*. *Pac Symp Biocomput*. 2013;147–158. **PMCID: PMC3615413**
22. Pendergrass SA, Frase A, Wallace J, Wolfe D, Katiyar N, et al. *Genomic analyses with biofilter 2.0: knowledge driven filtering, annotation, and model development*. *BioData Min*. 2013 Dec 30;6(1):25. **PMCID: PMC3917600**
23. Pasquale LR, Loomis SJ, Weinreb RN, Kang JH, Yaspan BL, et al. *Estrogen pathway polymorphisms in relation to primary open angle glaucoma: an analysis accounting for gender from the United States*. *Mol Vis*. 2013;19:1471–1481. **PMCID: PMC3712669**
24. Pasquale LR, Loomis SJ, Kang JH, Yaspan BL, Abdrabou W, et al. *CDKN2B-AS1 genotype-glaucoma feature correlations in primary open-angle glaucoma patients from the United States*. *Am J Ophthalmol*. 2013 Feb;155(2):342–353.e5. **PMCID: PMC3544983**
25. Mahnke AN, Plasek JM, Hoffman DG, Partridge NS, Foth WS, et al. *A rural community's involvement in the design and usability testing of a computer-based informed consent process for the personalized medicine research project*. *Am J Med Genet A*. 2013 Nov 22; **PMCID: PMC4145717**



26. Maenner MJ, Baker MW, Broman KW, Tian J, Barnes JK, et al. *FMR1 CGG expansions: prevalence and sex ratios*. *Am J Med Genet B Neuropsychiatr Genet*. 2013 Jul;162B(5):466–473. **PMCID: PMC3885228**
27. Hebbring SJ, Schrodri SJ, Ye Z, Zhou Z, Page D, et al. *A PheWAS approach in studying HLA-DRB1\*1501*. *Genes Immun*. 2013 Apr;14(3):187–191. **PMCID: PMC3637423**
28. Wiggs JL, Yaspan BL, Hauser MA, Kang JH, Allingham RR, et al. *Common variants at 9p21 and 8q22 are associated with increased susceptibility to optic nerve degeneration in glaucoma*. *PLoS Genet*. 2012;8(4):e1002654. **PMCID: PMC3343074**
29. Rasmussen LV, Peissig PL, McCarty CA, Starren J. *Development of an optical character recognition pipeline for handwritten form fields from an electronic health record*. *Journal of the American Medical Informatics Association: JAMIA*. 2012 Jun 1;19(e1):e90–e95. **PMCID: PMC3392858**
30. Haas RJ, McCarty CA, Payseur BA. *Genetic ancestry inference using support vector machines, and the active emergence of a unique American population*. *Eur J Hum Genet [Internet]*. 2012 Dec 5; **PMCID: PMC3641388**
31. Foth W, Waudby C, Brilliant MH. *Certificates of confidentiality and the Marshfield Clinic's Personalized Medicine Research Project*. *Virtual Mentor*. 2012;14(8):653–656. **PMCID: PMC3698606**
32. Waudby CJ, Berg RL, Linneman JG, Rasmussen LV, Peissig PL, et al. *Cataract research using electronic health records*. *BMC Ophthalmol*. 2011 Nov 11;11:32. **PMCID: PMC3226674**

## Mayo

1. Robinson CL, Jouni H, Krusselbrink TM, Austin EE, Christensen KD, et al. *Disclosing genetic risk for coronary heart disease: effects on perceived personal control and genetic counseling satisfaction*. *Clin Genet*. 2015 Feb 23; **PMID: 25708169**
2. Ritchie MD, de Andrade M, Kuivaniemi H. *The foundation of precision medicine: integration of electronic health records with genomics through basic, clinical, and translational research*. *Front Genet [Internet]*. 2015 Mar 17 [cited 2015 Jun 25];6. **PMCID: PMC4362332**
3. Panahiazar M, Taslimitehrani V, Pereira NL, Pathak J. *Using EHRs for Heart Failure Therapy Recommendation Using Multidimensional Patient Similarity Analytics*. *Stud Health Technol Inform*. 2015;210:369–373. **PMID: 25991168**
4. Nikpay M, Goel A, Won H-H, Hall LM, Willenborg C, et al. *A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease*. *Nat Genet*. 2015 Oct;47(10):1121–1130. **PMCID: PMC4589895**
5. Lindor NM, Schahl KA, Johnson KJ, Hunt KS, Mensink KA, et al. *Whole-Exome Sequencing of 10 Scientists: Evaluation of the Process and Outcomes*. *Mayo Clin Proc*. 2015 Oct;90(10):1327–1337. **PMID: 26434960**
6. Kullo IJ, Leeper NJ. *The genetic basis of peripheral arterial disease: current knowledge, challenges, and future directions*. *Circ Res*. 2015 Apr 24;116(9):1551–1560. **PMCID: PMC4410432**
7. Kullo IJ, Jouni H, Olson JE, Montori VM, Bailey KR. *Design of a randomized controlled trial of disclosing genomic risk of coronary heart disease: the Myocardial Infarction Genes (MI-GENES) study*. *BMC Med Genomics*. 2015;8:51. **PMCID: PMC4536729**
8. Graves KD, Sinicrope PS, McCormick JB, Zhou Y, Vadaparampil ST, et al. *Public perceptions of disease severity but not actionability correlate with interest in receiving genomic results: nonalignment with current trends in practice*. *Public Health Genomics*. 2015;18(3):173–183. **PMID: 25790929**
9. Cliff KE, Halverson CME, Fiksdal AS, Kumbamu A, Sharp RR, et al. *Patients' views on incidental findings from clinical exome sequencing*. *Applied & Translational Genomics*. 2015 Mar;4:38–43.
10. Caraballo PJ, Parkulo M, Blair D, Elliott M, Schultz C, et al. *Clinical Decision Support to Implement CYP2D6 Drug-Gene Interaction*. *Stud Health Technol Inform*. 2015;216:946. **PMID: 26262248**
11. Breitenstein MK, Simon G, Ryu E, Armasu SM, Weinshilboum RM, et al. *Using EHR-Linked Biobank Data to Study Metformin Pharmacogenomics*. *Stud Health Technol Inform*. 2015;210:914–918. **PMID: 25991289**
12. Barajas MR, Formea CM, McCormick JB, Abdalrhim AD, Han LC, et al. *A patient-centered approach to the development and pilot of a warfarin pharmacogenomics patient education tool for health professionals*. *Curr Pharm Teach Learn*. 2015 Apr;7(2):249–255. **PMCID: PMC4339072**
13. Allyse MA, McCormick JB, Sharp RR. *Prudentia Populo: Involving the Community in Biobank Governance*. *Am J Bioeth*. 2015;15(9):1–3. **PMID: 26305740**
14. Peterson KJ, Pathak J. *Scalable and High-Throughput Execution of Clinical Quality Measures from Electronic Health Records using MapReduce and the JBoss® Drools Engine*. *AMIA Annu Symp Proc*. 2014;2014:1864–1873. **PMCID: PMC4419873**
15. Olson JE, Bielinski SJ, Ryu E, Winkler EM, Takahashi PY, et al. *Biobanks and personalized medicine*. *Clin Genet*. 2014 Jul;86(1):50–55. **PMID: 24588254**
16. Khaleghi M, Isseh IN, Jouni H, Sohn S, Bailey KR, et al. *Family history as a risk factor for carotid artery stenosis*. *Stroke*. 2014 Aug;45(8):2252–2256. **PMID: 25005442**
17. Khaleghi M, Isseh IN, Bailey KR, Kullo IJ. *Family history as a risk factor for peripheral arterial disease*. *Am J Cardiol*. 2014 Sep 15;114(6):928–932. **PMCID: PMC4206951**
18. Jadhav A, Sheth A, Pathak J. *Analysis of Online Information Searching for Cardiovascular Diseases on a Consumer Health Information Portal*. *AMIA Annu Symp Proc*. 2014 Nov 14;2014:739–748. **PMCID: PMC4419905**

19. Jadhav A, Andrews D, Fiksdal A, Kumbamu A, McCormick JB, et al. *Comparative analysis of online health queries originating from personal computers and smart devices on a consumer health information portal*. J Med Internet Res. 2014;16(7):e160. **PMCID: PMC4115262**
20. Bobo WV, Pathak J, Kremers HM, Yawn BP, Brue SM, et al. *An electronic health record driven algorithm to identify incident antidepressant medication users*. J Am Med Inform Assoc. 2014 Oct;21(5):785–791. **PMCID: PMC4147111**
21. Bielinski SJ, Olson JE, Pathak J, Weinshilboum RM, Wang L, et al. *Preemptive genotyping for personalized medicine: design of the right drug, right dose, right time-using genomic data to individualize treatment protocol*. Mayo Clin Proc. 2014 Jan;89(1):25–33. **PMCID: PMC3932754**
22. Zhu Q, Freimuth RR, Pathak J, Chute CG. *Using clinical element models for pharmacogenomic study data standardization*. AMIA Summits Transl Sci Proc. 2013 Mar 18;2013:292–296. **PMCID: PMC3845787**
23. Wooten EC, Hebl VB, Wolf MJ, Greytak SR, Orr NM, et al. *Formin homology 2 domain containing 3 variants associated with hypertrophic cardiomyopathy*. Circ Cardiovasc Genet. 2013 Feb;6(1):10–18. **PMCID: PMC3578062**
24. Tao C, Li D, Shen F, Lian Z, Pathak J, et al. *Phenotyping on EHR Data Using OWL and Semantic Web Technologies*. Proceedings of the 2013 International Conference on Smart Health. Berlin, Heidelberg: Springer-Verlag; 2013. p. 31–32.
25. Takahashi PY, Ryu E, Olson JE, Anderson KS, Hathcock MA, et al. *Hospitalizations and emergency department use in Mayo Clinic Biobank participants within the employee and community health medical home*. Mayo Clin Proc. 2013 Sep;88(9):963–969. **PMCID: PMC4151531**
26. Sohn S, Ye Z, Liu H, Chute CG, Kullo IJ. *Identifying Abdominal Aortic Aneurysm Cases and Controls using Natural Language Processing of Radiology Reports*. AMIA Summits Transl Sci Proc. 2013 Mar 18;2013:249–253. **PMCID: PMC3845740**
27. Seedorff M, Peterson KJ, Nelsen LA, Cocos C, McCormick JB, et al. *Incorporating expert terminology and disease risk factors into consumer health vocabularies*. Pac Symp Biocomput. 2013;421–432. **PMCID: PMC3587774**
28. Ridgeway JL, Han LC, Olson JE, Lackore KA, Koenig BA, et al. *Potential bias in the bank: what distinguishes refusers, nonresponders and participants in a clinic-based biobank?* Public Health Genomics. 2013 Apr 12;16(3):118–126. **PMCID: PMC3821039**
29. Rea S, Bailey KR, Pathak J, Haug PJ. *Bias in recording of body mass index data in the electronic health record*. AMIA Summits Transl Sci Proc. 2013 Mar 18;2013:214–218. **PMCID: PMC3845755**
30. Pathak J, Kiefer RC, Chute CG. *Using linked data for mining drug-drug interactions in electronic health records*. Stud Health Technol Inform. 2013;192:682–686. **PMCID: PMC3909652**
31. Pathak J, Kiefer RC, Chute CG. *Mining Anti-coagulant Drug-Drug Interactions from Electronic Health Records Using Linked Data*. In: Baker CJO, Butler G, Jurisica I, editors. Data Integration in the Life Sciences. Springer Berlin Heidelberg; 2013. p. 128–140.
32. Pathak J, Bailey KR, Beebe CE, Bethard S, Carrell DC, et al. *Normalization and standardization of electronic health records for high-throughput phenotyping: the SHARPN consortium*. J Am Med Inform Assoc. 2013 Nov 4; **PMCID: PMC3861933**
33. Olson JE, Ryu E, Johnson KJ, Koenig BA, Maschke KJ, et al. *The Mayo Clinic Biobank: a building block for individualized medicine*. Mayo Clin Proc. 2013 Sep;88(9):952–962. **PMCID: PMC4258707**
34. McGuire AL, Joffe S, Koenig BA, Biesecker BB, McCullough LB, et al. *Point-counterpoint. Ethics and genomic incidental findings*. Science. 2013 May 31;340(6136):1047–1048. **PMCID: PMC3772710**
35. McCormick J. *Whole Genome Sequencing*. Lahey Clinic Medical Ethics Journal. 2013 Winter;20(1):1–2.
36. Ludvigsson JF, Pathak J, Murphy S, Durski M, Kirsch PS, et al. *Use of computerized algorithm to identify individuals in need of testing for celiac disease*. J Am Med Inform Assoc. 2013 Aug 16; **PMCID: PMC3861918**
37. Liu H, Bielinski SJ, Sohn S, Murphy S, Waghlikar KB, et al. *An information extraction framework for cohort identification using electronic health records*. AMIA Jt Summits Transl Sci Proc. 2013;2013:149–153. **PMCID: PMC3845757**
38. Li D, Simon G, Chute CG, Pathak J. *Using association rule mining for phenotype extraction from electronic health records*. AMIA Summits Transl Sci Proc. 2013 Mar 18;2013:142–146. **PMCID: PMC3845788**
39. Kiefer RC, Freimuth RR, Chute CG, Pathak J. *Mining Genotype-Phenotype Associations from Public Knowledge Sources via Semantic Web Querying*. AMIA Summits Transl Sci Proc. 2013 Mar 18;2013:118–122. **PMCID: PMC3845769**
40. Jouni H, Shameer K, Asmann YW, Hazin R, Andrade M de, et al. *Clinical Correlates of Autosomal Chromosomal Abnormalities in an Electronic Medical Record-Linked Genome-Wide Association Study A Case Series*. Journal of Investigative Medicine High Impact Case Reports. 2013 Oct;1(4):2324709613508932. **PMCID: PMC4130164**
41. Fan J, Arruda-Olson AM, Leibson CL, Smith C, Liu G, et al. *Billing code algorithms to identify cases of peripheral artery disease from administrative data*. J Am Med Inform Assoc. 2013 Dec;20(e2):e349–354. **PMCID: PMC3861931**
42. Zhu Q, Freimuth RR, Lian Z, Bauer S, Pathak J, et al. *Harmonization and semantic annotation of data dictionaries from the Pharmacogenomics Research Network: A case study*. J Biomed Inform [Internet]. 2012 Nov 29; **PMCID: PMC3606279**
43. Thompson WK, Rasmussen LV, Pacheco JA, Peissig PL, Denny JC, et al. *An Evaluation of the NQF Quality Data Model for Representing Electronic Health Record Driven Phenotyping Algorithms*. AMIA Annu Symp Proc. 2012;2012:911–920. **PMCID: PMC3540514**
44. Tao C, Pathak J, Solbrig HR, Wei W-Q, Chute CG. *Terminology representation guidelines for biomedical ontologies in the semantic web notations*. J Biomed Inform [Internet]. 2012 Sep 28; **PMCID: PMC3563768**

45. Tao C, Jiang G, Oniki TA, Freimuth RR, Zhu Q, et al. *A semantic-web oriented representation of the clinical element model for secondary use of electronic health records data*. J Am Med Inform Assoc [Internet]. 2012 Dec 25. **PMCID: PMC3628064**
46. Savova GK, Olson JE, Murphy SP, Cafourek VL, Couch FJ, et al. *Automated discovery of drug treatment patterns for endocrine therapy of breast cancer within an electronic medical record*. J Am Med Inform Assoc. 2012 Jun;19(e1):e83–89. **PMCID: PMC3392847**
47. Pathak J, Kiefer RC, Chute CG. *Using Semantic Web Technologies for Cohort Identification from Electronic Health Records for Clinical Research*. AMIA Summits Transl Sci Proc. 2012 Mar 19;2012:10–19. **PMCID: PMC3392057**
48. Pathak J, Kiefer RC, Bielinski SJ, Chute CG. *Mining the Human Phenome using Semantic Web Technologies: A Case Study for Type 2 Diabetes*. AMIA Annu Symp Proc. 2012;2012:699–708. **PMCID: PMC3540447**
49. Pathak J, Kiefer RC, Bielinski SJ, Chute CG. *Applying semantic web technologies for phenome-wide scan using an electronic health record linked Biobank*. J Biomed Semantics. 2012;3(1):10. **PMCID: PMC3554594**
50. Pathak J, Kiefer R, Freimuth R, Chute C. *Validation and discovery of genotype-phenotype associations in chronic diseases using linked data*. Stud Health Technol Inform. 2012;180:549–553. **PMID: 22874251**
51. Li D, Endle CM, Murthy S, Stancl C, Suesse D, et al. *Modeling and Executing Electronic Health Records Driven Phenotyping Algorithms using the NQF Quality Data Model and JBoss® Drools Engine*. AMIA Annu Symp Proc. 2012;2012:532–541. **PMCID: PMC3540464**
52. Clifford L, Singh A, Wilson GA, Toy P, Gajic O, et al. *Electronic health record surveillance algorithms facilitate the detection of transfusion-related pulmonary complications*. Transfusion. 2012 Aug 31; **PMID: 22934792**
53. Pathak J, Chute CG. *Further revamping VA's NDF-RT drug terminology for clinical research*. J Am Med Inform Assoc. 2011 May;18(3):347–348. **PMCID: PMC3078669**
54. Ding K, Bailey KR, Kullo IJ. *Genotype-informed estimation of risk of coronary heart disease based on genome-wide association data linked to the electronic medical record*. BMC Cardiovasc Disord. 2011 Nov 3;11:66. **PMCID: PMC3269823**
55. Chute CG, Pathak J, Savova GK, Bailey KR, Schor MI, et al. *The SHARPN project on secondary use of Electronic Medical Record data: progress, plans, and possibilities*. AMIA Annu Symp Proc. 2011 Oct 22;2011:248–256. **PMCID: PMC3243296**
56. Bielinski SJ, Chai HS, Pathak J, Talwalkar JA, Limburg PJ, et al. *Mayo Genome Consortia: A Genotype-Phenotype Resource for Genome-Wide Association Studies With an Application to the Analysis of Circulating Bilirubin Levels*. Mayo Clin Proc. 2011 Jul;86(7):606–614. **PMCID: PMC3127556**

## Mount Sinai

1. Zhu X, Feng T, Tayo BO, Liang J, Young JH, et al. *Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension*. Am J Hum Genet. 2015 Jan 8;96(1):21–36. **PMCID: PMC4289691**
2. Singh A, Nadkarni G, Gottesman O, Ellis SB, Bottinger EP, et al. *Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration*. J Biomed Inform. 2015 Feb;53:220–228. **PMCID: PMC4520404**
3. Pers TH, Karjalainen JM, Chan Y, Westra H-J, Wood AR, et al. *Biological interpretation of genome-wide association studies using predicted gene functions*. Nat Commun [Internet]. 2015 Jan 19 [cited 2015 Jun 26];6. **PMCID: PMC4420238**
4. Huopaniemi I, Nadkarni G, Nadukuru R, Lotay V, Ellis S, et al. *Disease progression subtype discovery from longitudinal EMR data with a majority of missing values and unknown initial time points*. AMIA Annu Symp Proc. 2014 Nov 14;2014:709–718. **PMCID: PMC4419979**
5. Fort D, Wilcox AB, Weng C. *Could Patient Self-reported Health Data Complement EHR for Phenotyping?* AMIA Annu Symp Proc. 2014;2014:1738–1747. **PMCID: PMC4419899**
6. Overby CL, Weng C, Haerian K, Perotte A, Friedman C, et al. *Evaluation considerations for EHR-based phenotyping algorithms: A case study for drug-induced liver injury*. AMIA Summits Transl Sci Proc. 2013 Mar 18;2013:130–134. **PMCID: PMC3814479**

## Northwestern

1. Rasmussen LV. *The electronic health record for translational research*. J Cardiovasc Transl Res. 2014 Aug;7(6):607–614. **PMCID: PMC4147395**
2. Gawron AJ, Thompson WK, Keswani RN, Rasmussen LV, Kho AN. *Anatomic and advanced adenoma detection rates as quality metrics determined via natural language processing*. Am J Gastroenterol. 2014 Dec;109(12):1844–1849. **PMID: 24935271**
3. Smith ME, Aufox S. *Biobanking: The Melding of Research with Clinical Care*. Curr Genet Med Rep. 2013 Jun 1;1(2):122–128. **PMCID: PMC3804314**
4. Chisholm RL. *The opportunities and challenges of implementing genomics-informed personalized medicine*. Clin Pharmacol Ther. 2013 Aug;94(2):181–182. **PMID: 23872829**
5. Edwards KL, Lemke AA, Trinidad SB, Lewis SM, Starks H, et al. *Genetics researchers' and IRB professionals' attitudes toward genetic research review: a comparative analysis*. Genet Med. 2012 Feb;14(2):236–242. **PMCID: PMC3448270**
6. Lemke AA, Smith ME, Wolf WA, Trinidad SB. *Broad Data Sharing in Genetic Research: Views of Institutional Review Board Professionals*. IRB. 2011 Jun;33(3):1–5. **PMCID: PMC3394177**

7. Edwards KL, Lemke AA, Trinidad SB, Lewis SM, Starks H, et al. *Attitudes toward Genetic Research Review: Results from a Survey of Human Genetics Researchers*. Public Health Genomics. 2011 Sep;14(6):337–345. **PMCID: PMC3221257**

## Vanderbilt

1. Wei W-Q, Denny JC. *Extracting research-quality phenotypes from electronic health records to support precision medicine*. Genome Med [Internet]. 2015 Apr 30 [cited 2015 Jun 23];7(1). **PMCID: PMC4416392**
2. Weeke P, Denny JC, Basterache L, Shaffer C, Bowton E, et al. *Examining rare and low-frequency genetic variants previously associated with lone or familial forms of atrial fibrillation in an electronic medical record system: a cautionary note*. Circ Cardiovasc Genet. 2015 Feb;8(1):58–63. **PMCID: PMC4334677**
3. Wan Z, Vorobeychik Y, Xia W, Clayton EW, Kantarcioglu M, et al. *A Game Theoretic Framework for Analyzing Re-Identification Risk*. PLoS ONE. 2015 Mar 25;10(3):e0120592. **PMCID: PMC4373733**
4. Unertl KM, Field JR, Price L, Peterson JF. *Clinician Perspectives on Using Pharmacogenomics in Clinical Practice*. Per Med. 2015;12(4):339–347. **PMCID: PMC4664195**
5. Peterson JF, Field JR, Shi Y, Schildcrout JS, Denny JC, et al. *Attitudes of clinicians following large-scale pharmacogenomics implementation*. Pharmacogenomics J. 2015 Aug 11; **PMID: 26261062**
6. Dumitrescu L, Goodloe R, Bradford Y, Farber-Eger E, Boston J, et al. *The effects of electronic medical record phenotyping details on genetic association studies: HDL-C as a case study*. BioData Min [Internet]. 2015 May 6 [cited 2016 Jan 8];8. **PMCID: PMC4428098**
7. Xu H, Aldrich MC, Chen Q, Liu H, Peterson NB, et al. *Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality*. J Am Med Inform Assoc. 2014 Jul 22; **PMCID: PMC4433365**
8. Xie W, Kantarcioglu M, Bush WS, Crawford D, Denny JC, et al. *SecureMA: protecting participant privacy in genetic association meta-analysis*. Bioinformatics. 2014 Dec 1;30(23):3334–3341. **PMCID: PMC4296153**
9. Weeke P, Roden DM. *Applied pharmacogenomics in cardiovascular medicine*. Annu Rev Med. 2014;65:81–94. **PMCID: PMC4081861**
10. Stitzel N, Hong-Hee W, Morrison AC, Peloso GM, Do R, et al. *Inactivating Mutations in NPC1L1 and Protection from Coronary Heart Disease*. New England Journal of Medicine. 2014 Nov 12;371(22):2072–2082. **PMID: 25390462**
11. Sattar AHMS, Li J, Liu J, Heatherly R, Malin B. *A probabilistic approach to mitigate composition attacks on privacy in non-coordinated environments*. Knowledge-Based Systems. 2014; **PMCID: PMC4295121**
12. Ramsey LB, Johnson SG, Caudle KE, Haidar CE, Voora D, et al. *The Clinical Pharmacogenetics Implementation Consortium Guideline for SLCO1B1 and Simvastatin-Induced Myopathy: 2014 Update*. Clin Pharmacol Ther. 2014 Oct;96(4):423–428. **PMCID: PMC4169720**
13. Ogbogu U, Burningham S, Ollenberger A, Calder K, Du L, et al. *Policy recommendations for addressing privacy challenges associated with cell-based research and interventions*. BMC Med Ethics. 2014 Feb 3;15(1):7. **PMCID: PMC3914710**
14. Naveed M, Ayday E, Clayton EW, Fellay J, Gunter CA, et al. *Privacy in the Genomic Era*. arXiv:14051891 [cs] [Internet]. 2014 May 8 [cited 2016 Jan 8];
15. Li M, Carrell D, Aberdeen J, Hirschman L, Malin BA. *De-identification of clinical narratives through writing complexity measures*. Int J Med Inform. 2014 Oct;83(10):750–767. **PMCID: PMC4215974**
16. Kawai VK, Cunningham A, Vear SI, Van Driest SL, Oginni A, et al. *Genotype and risk of major bleeding during warfarin treatment*. Pharmacogenomics. 2014 Dec;15(16):1973–1983. **PMCID: PMC4304738**
17. Karnes JH, Cronin RM, Rollin J, Teumer A, Pouplard C, et al. *A genome-wide association study of heparin-induced thrombocytopenia using an electronic medical record*. Thromb Haemost. 2014 Dec 11;113(3). **PMCID: PMC4433536**
18. Jeff JM, Armstrong LL, Ritchie MD, Denny JC, Kho AN, et al. *Admixture mapping and subsequent fine-mapping suggests a biologically relevant and novel association on chromosome 11 for type 2 diabetes in African Americans*. PLoS ONE. 2014;9(3):e86931. **PMCID: PMC3940426**
19. Xia W, Heatherly R, Ding X, Li J, Malin B. *Efficient discovery of de-identification policy options through a risk-utility frontier*. Proceedings of the third ACM conference on Data and application security and privacy. New York, NY, USA: ACM; 2013. p. 59–70. **PMCID: PMC4266184**
20. Wei W-Q, Cronin RM, Xu H, Lasko TA, Basterache L, et al. *Development and evaluation of an ensemble resource linking medications to their indications*. J Am Med Inform Assoc. 2013 Sep 1;20(5):954–961. **PMCID: PMC3756263**
21. Van Driest S, Shi Y, Bowton E, Schildcrout J, Peterson J, et al. *Clinically Actionable Genotypes Among 10,000 Patients With Preemptive Pharmacogenomic Testing*. Clin Pharmacol Ther. 2013 Nov 19; **PMCID: PMC3961508**
22. Rosenbloom ST, Madison JL, Brothers KB, Bowton EA, Clayton EW, et al. *Ethical and practical challenges to studying patients who opt out of large-scale biorepository research*. J Am Med Inform Assoc. 2013 Jul 25; **PMCID: PMC3861935**
23. Oetjens MT, Denny JC, Ritchie MD, Gillani NB, Richardson DM, et al. *Assessment of a pharmacogenomic marker panel in a polypharmacy population identified from electronic medical records*. Pharmacogenomics. 2013 May;14(7):735–744. **PMCID: PMC3725600**

24. Heatherly RD, Loukides G, Denny JC, Haines JL, Roden DM, et al. *Enabling genomic-phenomic association discovery without sacrificing anonymity*. PLoS ONE. 2013 Feb 6;8(2):e53875. **PMCID: PMC3566194**
25. Brothers KB, Westbrook MJ, Wright MF, Myers JA, Morrison DR, et al. *Patient awareness and approval for an opt-out genomic biorepository*. Personalized Medicine. 2013 Jun;10(4):349–359. **PMCID: PMC3882901**
26. Atreya RV, Smith JC, McCoy AB, Malin B, Miller RA. *Reducing patient re-identification risk for laboratory results within research datasets*. J Am Med Inform Assoc. 2013 Jan 1;20(1):95–101. **PMCID: PMC3555327**
27. Altman RB, Clayton EW, Kohane IS, Malin BA, Roden DM. *Data re-identification: societal safeguards*. Science. 2013 Mar 1;339(6123):1032–1033. **PMCID: PMC3740512**
28. Wilke RA, Ramsey LB, Johnson SG, Maxwell WD, McLeod HL, et al. *The clinical pharmacogenomics implementation consortium: CPIC guideline for SLCO1B1 and simvastatin-induced myopathy*. Clin Pharmacol Ther. 2012 Jul;92(1):112–117. **PMCID: PMC3384438**
29. Westbrook MJ, Wright MF, Van Driest SL, McGregor TL, Denny JC, et al. *Mapping the incidentalome: estimating incidental findings generated through clinical pharmacogenomics testing*. Genet Med [Internet]. 2012 Nov 29; **PMCID: PMC3648626**
30. Tamersoy A, Loukides G, Nergiz ME, Saygin Y, Malin B. *Anonymization of Longitudinal Electronic Medical Records*. IEEE Transactions on Information Technology in Biomedicine. 2012 May;16(3):413–423. **PMCID: PMC3779068**
31. Schildcrout JS, Denny JC, Bowton E, Gregg W, Pulley JM, et al. *Optimizing Drug Outcomes Through Pharmacogenetics: A Case for Preemptive Genotyping*. Clinical Pharmacology & Therapeutics. 2012;92(2):235–242. **PMCID: PMC3785311**
32. Pulley JM, Denny JC, Peterson JF, Bernard GR, Vnencak-Jones CL, et al. *Operational Implementation of Prospective Genotyping for Personalized Medicine: The Design of the Vanderbilt PREDICT Project*. Clinical Pharmacology & Therapeutics. 2012;92(1):87–95. **PMCID: PMC3581305**
33. Loukides G, Gkoulalas-Divanis A. *Utility-preserving transaction data anonymization with low information loss*. Expert Syst Appl. 2012 Aug;39(10):9764–9777. **PMCID: PMC3340604**
34. Clayton EW. *Sharing individual research results with biospecimen contributors: counterpoint*. Cancer Epidemiol Biomarkers Prev. 2012 Feb;21(2):260–261. **PMCID: PMC3815582**
35. Clayton EW, McGuire AL. *The legal risks of returning results of genomics research*. Genet Med. 2012 Apr;14(4):473–477. **PMCID: PMC3779603**
36. Brothers KB, Clayton EW. *Parental Perspectives on a Pediatric Human Non-Subjects Biobank*. AJOB Primary Research. 2012;3(3):21–29. **PMCID: PMC3501745**
37. Xu H, AbdelRahman S, Lu Y, Denny JC, Doan S. *Applying semantic-based probabilistic context-free grammar to medical language processing—a preliminary study on parsing medication sentences*. J Biomed Inform. 2011 Dec;44(6):1068–1075. **PMCID: PMC3226929**
38. Tatonetti NP, Denny JC, Murphy SN, Fernald GH, Krishnan G, et al. *Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels*. Clin Pharmacol Ther. 2011 Jul;90(1):133–142. **PMCID: PMC3216673**
39. Malin B, Loukides G, Benitez K, Clayton E. *Identifiability in biobanks: models, measures, and mitigation strategies*. Human Genetics. 2011 Sep;130(3):383–392. **PMCID: PMC3621020**
40. Langanke M, Brothers KB, Erdmann P, Weinert J, Krafczyk-Korth J, et al. *Comparing different scientific approaches to personalized medicine: research ethics and privacy protection*. Per Med. 2011 Jul;8(4):437–444. **PMCID: PMC3164515**
41. El Emam K, Jonker E, Arbuckle L, Malin B. *A Systematic Review of Re-Identification Attacks on Health Data*. PLoS One. 2011 Dec 2;6(12). **PMCID: PMC3229505**
42. Brothers KB, Morrison DR, Clayton EW. *Two large-scale surveys on community attitudes toward an opt-out biobank*. Am J Med Genet A. 2011 Dec;155A(12):2982–2990. **PMCID: PMC3222722**

# OVERVIEW of eMERGE TOOLS

## Main resources and tools produced and supported:

- [CDS KB](#): A repository of clinical decision support knowledge designed to support clinical processes from diagnosis and investigation through treatment and long term care.
- [eleMAP](#): A tool that allows researchers to harmonize their local phenotype data dictionaries to existing metadata and terminology standards such as the caDSR (Cancer Data Standards Registry and Repository), NCIT (NCI Thesaurus) and SNOMED-CT (Systematized Nomenclature of Medicine-Clinical Terms).
- [eMERGE Infobutton Project](#): A template containing important content topics for genomic medicine.
- [eMERGE Record Counter](#): A web-based research tool that provides exploratory data figures for research planning purposes and feasibility assessment.
- [Model consent language](#): A publication representing the compiled work of eMERGE I investigators and consultants on consent language for the collection and storage of human biospecimens and data for future research, particularly those collections that have an electronic medical records component.
- [MyResults.org](#): A website to educate providers and patients about genetic test results and the related medications.
- [PheKB](#): An online collaborative environment for building and validating electronic algorithms to identify characteristics of patients within health data.
- [PheWAS Catalog](#): A catalog that contains the PheWAS results for 3,144 single-nucleotide polymorphisms (SNPs) present in the NHGRI GWAS Catalog.
- [SPHINX](#): A data exploring tool for genetics related drug response hypothesis generation.
- [Synthesis-View](#), [PheWAS-View](#) and Phenogram: Visualization tools for genome & phenome-wide data.

# SITE SUMMARIES

## The Center for Applied Genomics (CAG) at Children's Hospital of Philadelphia

- **Aim 1: Phenotype Characterization in Four Pediatric Cohorts**
  - We have been active in the Phase II Phenotyping workgroup, leading the asthma (4,498 CHOP samples, 3,680 additional eMERGE samples), asthma severity (4,498; >1000 eMERGE), ADHD (1,013 CHOP samples, +335 eMERGE), atopic dermatitis (1,695 CHOP, +383 eMERGE), and GERD (1,573 CHOP, eMERGE in progress) algorithms. For Phase III, we propose to develop algorithms for the following primary phenotypes: intellectual disability, epilepsy, autism and obesity, all of which have multiple covariate phenotypes to pursue (i.e. obesity patients are enriched for asthma, obstructive sleep apnea, diabetes and metabolic syndrome). We will also mine EHRs to determine pharmacogenomic (PGx) response profiles (efficacy and adverse events), and search for polymorphisms impacting variation in response to commonly-used drugs in the existing pediatric dataset, leveraging CHOP and eMERGE data. *We hypothesize that data-mining of information from GWAS and EHRs from >60,000 children of different ethnic backgrounds will result in discovery of biomarkers and disease/PGx profiles that can inform clinical care.* (Note: Here, we collapse Pediatric/PGx workgroups under Phenotyping)
- **Aim 2: Characterization of Common and Rare Variants in >3,000 Pediatric Patients**
  - CAG has built a comprehensive infrastructure for detecting novel rare variants, yielding more than 375 publications since our founding in 2006. Under eMERGE-II, we (re-)developed a suite of genomics tools including ANNOVAR (950 citations), PennCNV (751), SNVER (57), ParseCNV (11), and PennCNV-Seq (10). While continuing to drive discovery in the complex phenotypes (as per eMERGE-II), these tools will be further optimized for detecting rare variants from existing genotype and sequencing data, as well as prospective data generated under eMERGE Phase III. We aim to build upon extensive preliminary data and public resources to catalog and validate rare variants in those phenotypes proposed above, as well as those proposed by other eMERGE sites. *We hypothesize that continued optimization of CAG's analysis pipeline will uncover a range of actionable variants of diagnostic/clinical significance in a range of pediatric and adult diseases.*
- **Aim 3: Return of Results (RoR) to 160+ Patients with Actionable Genomic Variants**
  - Under eMERGE Phase II, we extended our Clinical Laboratory Improvement Act (CLIA)/College of American Pathologists (CAP) certified workflow status. We are CLIA-approved for genotyping, clinical pathology-, whole exome- and whole genome- sequencing (ThermoFisher and Illumina). We have developed a highly efficient genomics pipeline for sample-tracking, quality-control, base calling, and annotating variants all of which are being conducted in-house in a CLIA-approved environment. Our analysis team and state-of-the-art genotyping/sequencing resources, affords flexibility to efficiently validate actionable variants in a wide range of phenotypes. *We hypothesize that state-of-the-art informatics analyses will accelerate characterization of actionable variants, maximize validation success and clinical utility, and ultimately improve patient care.*
- **Aim 4: Exploration of Challenges Facing At-Risk Families**
  - We propose a series of formative and summative analyses to assess the health impact, cost-effectiveness, and ELSI implications of the site for patients. We will follow patients from our RoR program, leveraging CAG's re-contact protocol to assess the health consequences of patients for whom results have been returned under SA3. Ultimately, this will feed-forward to improving guidelines and governance rules for CAGs biorepository, databases, and RoR procedures. In addition, we propose a series of developments of the CAG-developed *MyResults.org* website, established

under eMERGE-II. *We hypothesize that these programs will allow refinement to RoR protocols and yield successful and improved healthcare-related outcomes for CHOP patients.*

- **Aim 5: Integration of Genomic Data into CHOP's EHR**

- Under eMERGE II, we worked extensively with the IRB and EpicCare teams to develop a protocol for returning results to CAG patients, and our relevant infrastructure is live and operational. We propose to dramatically expand our EHRI platform in eMERGE-III, and to integrate actionable results identified under SA3. In addition *MyResults* will be redeveloped to improve transparency and include a physician education sub-section, with tie-in to clinical decision support (CDS) interfaces at CHOP, and other eMERGE sites. *We hypothesize that these efforts will leave CHOP poised for large-scale integration of genomic results into its Epic EHR platform, which will leverage the cumulative efforts of SA1-4.*

## Cincinnati Children's Hospital Medical Center (CCHMC)

- **Aim 1: Genomics**

- Rank at least 100 genes for DNA sequence analysis and develop an eMERGE III Targeted Gene Panel (eTGP) with the eMERGE III Steering Committee and other eMERGE III network members.
- Provide >2000 DNA samples for DNA sequencing on the eMERGE III Targeted Gene Panel and an aggregate of 6,544 data records to dbGAP or the Short Read Archive (SRA).
- Identify candidate causal somatic variants in 4,000 next generation sequencing (NGS) targeted gene panels obtained in the course of clinical care for patients at CCHMC.
- Further develop and provide user access to advanced genomic sequence analysis informatics (CASSI).

- **Aim 2: Phenotypes**

- Develop, validate, and distribute new electronic health record (EHR)-based phenotype algorithms informed by natural language processing using heuristic and machine learning methods.
  - Phenotypes underway include: Appendicitis, Methylphenidate response in Attention Deficit Hyperactivity Disorder (ADHD) and Malignant Hyperthermia.
  - Phenotypes for the theme of pain and its management include: Hypermobility (including Ehlers-Danlos Syndrome), Migraine, Fibromyalgia, Neonatal Abstinence Syndrome (NAS) and Narcotic Dependence.
  - Other phenotypes of present interest include Familial Hypercholesterolemia, Primary Pulmonary Hypertension, and Pyloric Stenosis.
- Explore the variants in the eMERGE III Gene Panel (eTGP) for candidate phenotypic expression.
- Pilot demonstration project: Translate pertinent EHR-based algorithms into "executable code" that will run against the PCORnet Common Data Model.
- Pilot demonstration project: Apply an eMERGE algorithm (e.g., Abdominal Aortic Aneurysm (AAA)) to the Million Veterans Project data and develop an algorithm in the Veterans Affairs EMR (e.g., Chronic Obstructive Lung Disease (COPD)) and bring to eMERGE III network sites.

- **Aim 3: Implementation and Evaluation**

- Determine the perspectives of patients, guardians, and physicians.
  - Pilot demonstration project: Development of tools to assess adolescent references.
  - Pilot demonstration project: Assessment of adolescent preferences for return of negative incidental genomic results and of electronic portal effectiveness (MyChart).
- Pilot demonstration project: Evaluate the genetics of outpatient pain after tonsillectomy.



- Pilot demonstration project: Assess the economic impact of CYP3A5 genotype guided tacrolimus dosing.
- Integrate clinical decision support processes into the EHR for identified cohorts, for follow-up recommendations, and for using specific discrete genetic information to trigger consequences.
  - Pilot demonstration project: *PTEN* analysis in patients who are evaluated for possible autism.
  - Pilot demonstration project: Assess outlier yield of testing for Primary Ciliary Dyskinesia (PCD).
  - Demonstration project: Evaluate >2000 eMERGE III Targeted Gene Panel (eTgP) data records for clinically actionable and returnable results; develop and apply workflows for returning results to physicians, and patients or guardians; and provide specific recommendations and clinical decision support tools to physicians and patients.
- Explore and integrate ethical, legal, and social considerations into EHR and clinical care decision making processes as genomic medicine advances, particularly re-evaluation of genomic variation.

## Columbia University

- **Aim 1: Advance next-generation phenotyping**
  - Designing, validating, and sharing high-throughput, data quality-aware, standards-based phenotyping methods.
- **Aim 2: Perform genetic association studies of rare variants with diverse clinical phenotypes**
  - Broad collaboration with the eMERGE network and other phenotyping research communities.
- **Aim 3: Develop practical, scalable learning mechanisms for returning results**
  - Leveraging a genomic patient portal and genetic providers to dynamically elicit and incorporate patient preferences for return of genomic results, returning results, and studying patient understanding of returned results.
- **Aim 4: Provide genomic decision support**
  - Enhancing and validating our clinical and informatics infrastructure for genomic decision support with learning mechanisms for tailored shared decision-making between patients and health care providers.

## Geisinger

- **Aim 1: Use existing biospecimens, genotype and sequence data and EMR-generated phenotypes for discovery in the proposed disorders: familial hypercholesterolemia (FH) and chronic rhinosinusitis (CRS).**
  - For FH discovery activities will include: characterization of variants in FH genes chosen for sequencing and correlation with lipid levels across a large, diverse US population; exploration of variants in these genes for extreme lipid phenotypes and Phenome Wide Association Studies (PheWAS) to identify other associated disorders; use of machine learning techniques to develop the most effective population screening algorithm to identify patients likely to have FH; Develop and test economic models of population based screening informed by data accruing from the project. For CRS discovery activities will include: Perform large scale replication studies of the small genetic epidemiology studies of CRS; study the impact of variants in the set of CRS candidate genes selected for sequencing and use for PheWAS; gene environment interactions including Environment Wide Association Studies (EWAS); study similarities and differences between CRS in children and adults.
- **Aim 2: Develop and test approaches for implementation of genomic information in clinical practice.**

- For FH implementation activities will include: Implemented screening algorithm with return of genomic results and study the impact on outcomes; study implementation strategies, enablers and barriers across all sites; implement familial case finding for individuals diagnosed with FH. For CRS implementation activities will include: Return of *CFTR* variants associated with a range of phenotypes; pilot the impact on treatment outcomes of treatment response or prognostic factors identified in the discovery aim.
- **Aim 3: Explore, develop and implement novel approaches for family-centered communication around clinically relevant genomic results.**
  - Utilizing mixed-methods, explore attitudes of biobank participants toward the familial implications of returned results, including factors that may impede or facilitate communication with other family members. We will subsequently develop and test alternative approaches for notifying family members of genomic results. Deliberative engagement will also be used to inform the envisioned two-phase study and achieve this specific aim. The best practices that result will be implemented for all GHS return of results.

## Group Health Cooperative/University of Washington

- **Aim 1: Genomic medicine discovery and implementation focused on CRC/P, TG, and NPC.**
  - Aim 1a: Collaboratively develop phenotype algorithms for the network, including our own CRC/P and TG phenotypes. This work builds on our E1/E2-derived expertise, including natural language processing (NLP). Our other site-specific phenotype, NPC, is easily derived from WBC/differential phenotypes established in E2.
  - Aim 1b: We will have 1000 participants with CRC/P and 1000 Asian ancestry participants sequenced. We will then use eMERGE-wide phenotype and sequence data, as well as collaborators' CRC/P data, to classify the pathogenicity of variants and to determine the age-dependent penetrance of pathogenic variants for CRC/P genes. Further, using a phenotype-wide association (PheWas) approach, we will determine whether pathogenic variants in these genes are associated with previously unrecognized cancer phenotypes.
  - Aim 1c: We propose to investigate rare variants in 2 novel genes for association with TG and NPC and explore new imputation methods to assess rare variant association with these quantitative traits.
- **Aim 2: Integrate genomic information into GH-wide clinical care and the EMR.**
  - Aim 2a: We will develop intuitive, comprehensive reports to return genes deemed actionable by the American College of Medical Genetics and Genomics (ACMG). If sequencing should be Clinical Laboratory Improvement Amendment (CLIA) compliant, negative results can be incorporated into the EMR. Medical Genetics will care for participants with actionable results.
  - Aim 2b: We will incorporate formal stakeholder input to implement integrated processes and tools to ensure that patients and providers are informed of appropriate actions based on genomic results, developing CDS and online information links for all providers including support for pathogenic variants for CRC/P and long QT syndrome (LQTS). These models can generalize to other genetic disorders.
  - Aim 2c: We will also implement processes and tools to support education for primary care and specialty providers, including development and evaluation of educational outreach and online resources.
- **Aim 3: Evaluate the effectiveness and economic impact of result return to patients and their families.**

- Aim 3a: We will evaluate the economic impact and cost effectiveness of returning IFs from the ACMG6 list of actionable genes.
- Aim 3b: We will test, in a nested randomized controlled trial (RCT), the effectiveness and social and economic impact of an innovative online tool to increase family communication about CRC/P risk and screening aimed at maximizing benefits and reducing harm.

## Harvard

- **Aim 1 (Discovery): Hypothesis: Common and rare variants from a custom chip including 50,000 loss of function (LoF) alleles will be associated with cardiovascular, neuropsychiatric and immune-mediated phenotypes derived from the EMR.**
  - We have analyzed 63,000 whole exome sequences, collected all LoF alleles observed among these individuals, bioinformatically filtered them, and designed a custom genotyping array that includes 50,000 novel LoF alleles as well as an enhanced common marker panel. In 25,000 Partners HealthCare Biobank subjects that will have been genotyped on this custom chip, we will test the association of variants with EMR phenotypes for cardiovascular, neuropsychiatric and immune-mediated diseases. Our goal is to discover clinically actionable genes and variants in Partners HealthCare Biobank subjects.
- **Aim 2 (Penetrance and Pleiotropy): Hypothesis: Sequencing a set of established genes or loci will allow us to discover additional variation, and define penetrance and pleiotropy using EMR phenotypes.**
  - Genes for sequencing will be chosen from five sources: associations found in specific aim 1, the ACMG list of actionable genes, the PGRN list of actionable pharmacogenetic genes, high impact common CVAS variants, and other eMERGE groups' nominations. Examples of these genes in our three disease areas might be: a) *LDLR*, *PCSK9*, and *APOB* for association with hypercholesterolemia, coronary heart disease, and stroke; b) *DRD2*, *CACNA1C*, *TCF4*, *CHD8* for association with ADHD, BD, depression, and schizophrenia; and c) *HLA-DRB1*, *IL23RA*, *TYK2*, *TNFRSF1A*, and *CTLA4* for association with RA, IBD, and MS, and asthma. We will study the penetrance and pleiotropic effects of rare variants in these genes by PheWAS, chart review and recall of subjects for additional phenotyping. Finally, genes will be ranked in terms of effect size and population-attributable risk for prioritization of return of research results.
- **Aim 3 (Implementation): Hypothesis: Physicians will alter their surveillance and treatment of patients based upon voluntary return of actionable variants to provide safe and cost-effective benefits to patients.**
  - To better understand the implications of returning genetic information from a Biobank into the clinical care stream for patients, we will screen our entire Biobank population of 25,000 individuals for pathogenic variants in the gene *LDLR*, the leading genetic cause of premature coronary artery disease, and conduct an exploratory trial in disclosing this information. Biobank participants with pathogenic variants in *LDLR* will be offered enrollment into a randomized trial in which their finding will be CLIA confirmed, and in one arm, this result will be communicated to their physicians through the EMR. Over one year, we will collect the following outcomes through participant surveys and EMR queries: physician visits, laboratory testing, changes in medication prescriptions, low-density lipoprotein cholesterol levels, medical costs and the number of family members screened and treated as a result of the intervention. We will collaborate with the entire eMERGE III Network to incorporate what we learn from this pilot trial into large-scale implementation protocols for the genes selected by the Network.

## Mayo Clinic

- **Aim 1: Classify variant pathogenicity and develop clinical decision support for actionable variants discovered by sequencing of a panel of 100 genes including genes related to FH/CRC.**
  - We will: a) use state-of-the-art bioinformatics pipelines for variant annotation and assign pathogenicity and actionability to variants through an expert multidisciplinary variant curation group; b) evaluate variants of uncertain significance (VUS) by investigating associations with EHR-ascertained phenotypes; and c) develop EHR-based clinical decision support for actionable variants to facilitate clinical decision-making at point of care.
- **Aim 2: Return clinically actionable results including those relevant to FH/CRC and assess the subsequent patient outcomes, costs and utilization of health care resources, and behavioral and psychosocial effects.**
  - Building on methods we developed in eMERGE II, we will assess: a) impact of genomic results on patient outcomes, including diagnostic and therapeutic interventions and new case detection in families; b) costs and utilization of health care resources following return of results; and c) understanding of genomic results, concerns about including genomic data in the EHR, worries about future employability or insurance coverage, and views regarding sharing genomic results with family members and others.
- **Aim 3: Perform EHR-based phenotyping and genomic discovery**
  - We will use validated electronic algorithms including natural language processing, to extract phenotypes from the Mayo and MPHIC EHRs. For the sequence data set (n= 25,000 across the network), our phenotypes of interest are plasma lipid levels. We will work with the FH Foundation and other eMERGE sites to develop an EHR-based algorithm for identifying FH cases. For the high-density genotyping data set (n= ~50,000 across the network) our phenotypes of interest include adverse drug reactions and treatment response. We will: a) perform common and rare variant association analyses for phenotypes of interest; b) explore penetrance using family data from cascade screening for FH/CRC; and c) assess the pleiotropic effects of genetic variants using PheWAS and additional novel approaches.

## Northwestern University

- **Aim 1: Perform electronic phenotyping and genomic analysis on at least four novel phenotypes of medical, scientific and public health importance.**
  - See [Section 3](#) for contact information. We will use both the eMERGE GWAS dataset (common variants) and deep DNA sequence of at least 100 genes likely to contain medically actionable variants (rare variants) to identify novel phenotype-genotype associations. New phenotyping algorithms for cardiovascular, gastrointestinal and neurological phenotypes, plus chronic rhinosinusitis will be developed at NU. Those developed at other sites will be implemented in the NU environment. We propose deep sequence analysis of candidate genes associated with our phenotypes and those recommended for reporting by the ACMG and some CPIC guidelines. We will deploy modular phenotyping algorithms and a PheWAS (2) approach that may allow sub-classification of conditions associated with the sequenced genes in addition to the broader collection of imputed variants available in the combined eMERGE I and II dataset.

- **Aim 2: Consent at least 2000 ethnically diverse participants who agree to receive results about clinically actionable variants found in their genotyped genes and return appropriate information to them.**
  - Participants will be enriched for the possibility of carrying an actionable genetic variant in one or more of the sequenced genes. We will implement strategies to identify actionable results using resources such as ClinGen and ClinVar, and ACMG and CPIC guidelines and return them in a clinical setting. A learning healthcare systems approach will be used in returning results that engages participants through capturing family history and other valuable outcome information that may aid in clinical care.
- **Aim 3: Expand our ancillary genomics system (AGS) to incorporate sequence data, including whole genome and exome sequence, and develop strategies for transferring actionable variants into EHR to enable decision support.**
  - We will assess the implications of this model on institutional policy, clinical care and research.
- **Aim 4: Develop and share best practices, expertise and experience for:**
  - Creating tools and methods for rapid and accurate phenotyping; (b) assisting patients and families in making choices about genomic results; (c) sharing of clinical and family history data with families; (d) storing and representing genomic data in the EHR; and (e) understanding downstream effects of genomic sequencing results on the healthcare system, patient and family.

## Vanderbilt University

- **Aim 1: To leverage EMR-based phenotyping to define genetics of disease and disease subsets.**
  - We will expand the network’s phenotyping library by creating increasingly granular phenotype definitions that identify specific *subsets* of disease with predictable clinical courses or response to therapies. These will be analyzed by GWAS using existing eMERGE genotypes. The increased phenotype granularity will also enable PheWAS for replication here and for discovery in Aim 2.
- **Aim 2: To identify rare variants with strong associations with human traits by sequencing 100 genes in 25,000 subjects across eMERGE-III.**
  - We propose that the 100 genes be drawn from (1) the ACMG list of 56 genes with variants known to confer large effect sizes for important human diseases; (2) eMERGEPGx genes; and (3) genes with very strong associations in PheWAS we have conducted using exome chip data in nearly 30,000 BioVU records to analyze associations with disease codes and lab values. We will use both genotype- and phenotype-driven methods to search for new associations between sequence variants and disease. Actionable sequence variants will be returned to subjects and their EMRs in Aim 3.
- **Aim 3: To expand PREDICT, our pre-emptive pharmacogenomic implementation program, to deliver actionable variants to patients and providers and to assess their response.**

Specific Aims 1 and 2 will expand the range of genomic or phenotypic information that could be considered relevant to an individual patient’s care (“actionable”). We have developed an initial framework for delivering such information in the flow of healthcare within the PREDICT program at Vanderbilt. Our initial focus has been on pharmacogenomic variation, and we will collaborate across eMERGE to develop, implement, and assess tools to deliver new information, measuring impact to ensure optimal benefit to patients.

## Partners – Broad Institute

- **Aim 1: Assay Design, Sequence Generation and Analysis.**
  - We will design and clinically validate a targeted eMERGE III gene panel using Broad's well-established clinical next generation sequencing (NGS) platform as well as additional guidance from LMM's many years of experience in rapid development and validation of CLIA-certified NGS gene panel tests. The LMM will also leverage its expertise in evaluating gene-disease relationships as well as its linkage with relevant ClinGen expert committees to advise on the contents of this gene panel. CLIA grade sequencing, alignment and variant calling will be performed at the Broad using software that is most widely used and at the leading edge of NGS analysis. Raw data as well as DNA aliquots will be immediately available to LMM for downstream processing.
- **Aim 2: Orthogonal Confirmation, Interpretation and Report Generation.**
  - Variant calls and raw data will flow from the Broad Institute to the eMERGE Coordinating Center (CC) as well as into LMM's clinical grade NGS tertiary analysis pipeline. Variant filtering will be done using criteria established with the eMERGE Steering Committee (SC). Variants will be validated to reach diagnostic confidence using orthogonal technologies as necessary (Sanger sequencing, ddPCR) and clinical reports will be created using LMM's CLIA processes, including its gene/disease/variant knowledge base, as well as automated report drafting infrastructure. We propose to create a default clinical report for all genes in the eMERGE panel. We propose to report all variants with established clinical significance but are prepared to customize reports according to site specific needs and adapt them over time as gene content and return of results approaches evolve.
- **Aim 3: Return of results, novel discovery approaches and interaction with the scientific and clinical genomics community.**
  - Standard clinical reports and subsequent knowledge updates will be transmitted electronically to the sites via site specific instances of GIC, which will link to LMM's variant knowledgebase. An eMERGE-specific instance of GIL will **a)** provide sites with direct real-time access to all genomic knowledge within LMM's genomic knowledge base, **b)** provide linkage to an existing variant and knowledge sharing network of multiple US and Canadian labs, and **c)** enable reciprocal exchange of variant data between eMERGE and ClinVar. To enable novel discovery analyses, primary data and variant annotations will be disseminated to the CC and/or to a portal located at the Broad Institute housing genomic analysis tools. We will liaise with the greater scientific and clinical genomics community via our established roles in the following NIH consortia: Clinical Genome Resource (**ClinGen**), Clinical Sequencing Exploratory Research (**CSER**), Newborn Sequencing In Genomic medicine and public Health (**NSIGHT**), Undiagnosed Diseases Network (**UDN**) and through our roles in organizations developing standards in genomics or coordinating the rollout of clinical genetic IT support including the American College of Medical Genetics and Genomics (**ACMG**), Association for Molecular Pathology (**AMP**), College of American Pathology (**CAP**), Global Alliance for Genomics and Health (**GA4GH**), Health Level Seven International (**HL7**), Human Gene Variation Society (**HGVS**), International Rare Diseases Research Consortium (**IRDRC**) and Institute of Medicine (**IOM**) to establish knowledge feeds and share genomic data and best practices between eMERGE and these bodies.

## Baylor College of Medicine Human Sequencing Center

- **Aim 1: Provide DNA sequence from approximately 100 selected gene targets in a CAP and CLIA certified environment to advance the aims of the eMERGE network.**
  - DNA samples will be received, archived and genotyped to enable sample tracking and QA/QC. Complete sequences of the defined loci will be generated by a custom DNA capture panel and high quality DNA sequence variants, including structural variants, will be identified by the BCM clinical analytical pipeline. Actionable variants, including but not limited to currently known pathological disease variants and the ACMG recommended incidental findings within the targeted areas, will be validated by Sanger sequencing. A second round of validation will be offered for variants that are nominated by eMERGE Network members. Designs can be iterated, at least annually.
- **Aim 2: Communicate data to the eMERGE coordination center, with standardized clinical annotations and advanced data security measures.**
  - Raw data and derived files will be archived for 2 years for the benefit of the consortia and submissions to the public databases such as dbGAP and CLinGen will be offered.
- **Aim 3: Progressively provide alternative strategies for sequencing so that the eMERGE Steering Committee can determine the most efficient and cost-effective pathway to molecular diagnoses.**
  - The sequencing core will apprise the eMERGE network of technical developments that may influence the Networks' choice of technologies and directions and so economics and other factors that favor alternate approaches can be considered. As a part of this collaborative effort, we will contribute expertise and leadership in the areas of efficient high throughput sequencing, data analysis/interpretation, and clinical reporting, while participating actively with the Steering Committee and network community to continually improve data generation.

## Vanderbilt University Serving as the Coordinating Center

- **Aim 1: Accelerate EMR phenotyping and enable end-to-end genomic medicine research.**
  - In Phase III, we propose advancing current infrastructure by expanding upon current tools and exploring new resources for electronic phenotype algorithm development and data sharing. We propose enabling broader PheWAS use with additional services and new methods. In addition, we propose establishing the Genomic Actionability and Implementation Repository (GenAIR), a collaborative development space and central repository for collection and dissemination of best practices and standards related to genomic medicine implementation.
- **Aim 2: Integrate high quality genomic information across eMERGE sites.**
  - The University of Washington (UW) team (led by Drs. Crosslin and Jarvik) will manage genomic data for eMERGE III. The group has extensive expertise providing sequencing data production and analysis support for many sequencing and genotyping studies. We have interacted and propelled many different studies through our involvement in the NIGMS Pharmacogenomics Research Network, the NHGRI Clinical Sequencing Exploratory Network, NHGRI/NHLBI Centers for Mendelian Genomics, the NHLBI Resequencing and Genotyping Project and Exome Sequencing Project (ESP), and many others. For Phase III, we will manage the following activities: 1) provide continued access to all existing genotype data for eMERGE data by download and improve on these data by providing new imputations to the expanding whole genome sequencing datasets; 2) transition the existing SPHINX database to a format similar to the exome variant server that is easy to use and familiar; 3) aggregate PGRNseq data to perform new network wide studies; 4) aggregate the sequence data produced in phase III by recalling the data and delivering multiple sample variant calls and

annotation for single nucleotide variant, indel, and structural variant analysis; and 5) assist with rare variation analysis for network wide projects using this new sequencing data.

- **Aim 3: Secure EMR and genomic data sharing risk mitigation.**
  - In Phase III, we propose to share tools and expertise that enable secure genomic medicine implementation while maintaining robust data access for research. We will continue to pursue research methods, tools and novel approaches that inform privacy and security policies around genomic sequencing and EMR data sharing. To bring privacy and data security issues and opportunities to the diverse investigators in the eMERGE Network, we propose a workshop for eMERGE participants on EMR and Genomic Data Privacy and Security and Clinical Data Management to be led by Brad Malin, PhD and Paul Harris, PhD (see Aim 5). By focusing efforts on privacy and anonymization policy guidance, novel methods research and real-world site implementation guidance, we will ensure eMERGE retains its reputational strength and leadership position within the field of EMR and genomic data sharing risk mitigation.
- **Aim 4: Provide excellent logistical support to the entire Network; including committees, work groups, NHGRI, and the ESP.**
  - We will draw heavily upon existing infrastructure, personnel and processes to assist the Network in accomplishing Phase III program goals. Specifically, we will continue to refine Network communications, organizational and functional support models to optimize engagement of a broad range of investigators, facilitate work, enhance learning and catalyze new ideas.
- **Specific Aim 5: Synergize with other related networks; to eliminate redundancy, to promote crosspollination of best practices, and to share eMERGE tools more broadly.**
  - The productive interactions with PGRN and CSER in Phase II resulted in large part from fortuitous collegiality among like-minded investigators involved in related research networks. We will use existing information-sharing tools to encourage, foster, and possibly seed more inter-network interaction. Project managers will follow up with project planning assistance for nascent collaborations or additional tool implementation advice to foster cross-network interactions. All network projects will continue to be governed according to each network's process, so for eMERGE, the Steering Committee will still need to approve any work deriving from these activities. As such, the success of this work will depend on cooperation between the operations groups and investigators from the different networks.



# ESP MEETING MINUTES: *May 2015*

## eMERGE Network

### ESP Conference Call Minutes

May 4, 2015 – 4:00 p.m. ET (3:00 p.m. CT/1:00 p.m. PT)

#### Attendees:

**CCHMC/BCH:** John Harley, Ingrid Holm; **CHOP:** Hakon Hakonarson; **Geisinger:** Marc Williams, David Carey; **Group Health/UW:** Gail Jarvik, Eric Larson; **Marshfield/Essentia/PSU:** Cathy McCarty, Peggy Peissig; **Mayo:** Iftikhar Kullo; **Mt. Sinai:** Erwin Bottinger; **Northwestern:** Rex Chisholm, Maureen Smith, Laura Rasmussen-Torvik; **VU:** Dan Roden, Sarah Stallings; **NHGRI:** Rongling Li, Teri Manolio, Ken Wiley, Jackie Odgis; **CC:** Paul Harris, Melissa Basford, Adam Hardebeck, Mollie Bodin-Claar; **ESP:** Howard McLeod (Moffitt Cancer Center), Eta Berner (University of Alabama-Birmingham), Gerardo Heiss (UNC), Lisa Parker (University of Pittsburgh)

#### Opening Remarks

- Rongling and Rex welcomed the ESP members, and Howard McLeod complimented the group for their work on the eMERGE Network ESP preparation packet.

#### Network Overview & Response to ESP Recommendations

- Recommendations from the ESP Packet were reviewed.
- The Network's publications have been cited 7,786 times since Phase I (2,704) since Phase II. The network's focus on network wide projects during 2015 was reaffirmed, as over 100 network manuscripts are being developed.
- The ESP noted that local IT integration can be a major challenge.
- ESP members inquired about any feedback received from clinicians related to the Infobutton project, and Marc explained that the project is in the CDS data collection process, but early analysis points to minimal pushback from clinicians.
- The ESP also inquired about eMERGE's experience with the correlation of genomic CDS to non-genomic CDS. They noted that the alert override audience doesn't feel as knowledgeable about CDS. Marc noted that besides taking longer to explain genomic matters to clinicians/admin., the group hasn't noticed significant difference between the two. The group will have more detailed answers after the CDS data analysis.

#### Network Workgroup Update: CERC Survey

- Maureen Smith and Ingrid Holm provided an overview of the CERC Survey project. On April 8<sup>th</sup>, 90,000 participants were mailed a survey. The survey currently has a response rate of 7%. A second wave of reminder cards and surveys will be circulated to those that did not initially respond. The survey analysis will extend beyond the end of Phase II.
- The pilot survey was also reviewed. The pilot had an overall response rate of 12% out of 1,500 targeted patients. It was also noted that oversampling of minority and low education groups helped provide a more demographically accurate representation of the study.
- The ESP inquired about the CERC Survey target cohort. The workgroup co-chairs confirmed that the survey recipients are patients from eMERGE institutions, but are not necessarily participants in eMERGE biobanks.
- It is not possible to know the overlap between survey recipients and eMERGE participants at all sites, but this

information is collected where possible. This info was not collected in the pilot survey.

#### Network Workgroup Update: PGx

- Sequencing is on track for completion by the end of Phase II.
- The ESP expressed interest in evidence of the comfort level with precision prescribing for providers. Provider Education metrics are largely qualitative, and implementation has taken place at various sites, making it difficult to gather consistent quantitative data.

**ACTION ITEM: The Provider Education team was advised to investigate methods to collect more qualitative data about providers' precision prescribing comfort level.**

#### Conclusion

- The ESP complimented the group for their productivity throughout Phase II, and for continuing to make progress throughout the grant.

#### Action Items

1. The Provider Education team was advised to investigate methods to collect more qualitative data about providers' precision prescribing comfort level.

## NHGRI-ESP Executive Session Minutes

5:00-5:30 p.m. ET (4:00-4:30 p.m. CT/2:00-2:30 p.m. PT)

#### Attendees:

**ESP:** Eta Berner (UAB) Gerardo Heiss (UNC), Howard McLeod (Moffitt, Chair), Lisa Parker (Pittsburgh); **NHGRI:** Ronling Li, Teri Manolio, Jacqueline Odgis

#### Executive Session

- The External Scientific Panel (ESP) met with members of NHGRI staff in Executive Session after the ESP teleconference held on May 4, 2015.
- The ESP was pleased with how responsive the Network has been to the ESP's recommendations from the last joint ESP-Steering Committee meeting. The ESP was also impressed with how well Network investigators were able to anticipate the ESP's questions in their presentations.
- The ESP appreciated the quality of the packet for this teleconference, as well as the increased amount of information on the progress of network-wide projects in the Steering Committee's report.
- The ESP noted that the Network seems to have many activities to complete before Phase II concludes.
- Regarding the CERC Survey project, the ESP was concerned about the low (~7%) response rate and the complexity of the concepts assessed, given the anticipated knowledge and reading level of the participants. NHGRI staff noted that the CERC Survey Workgroup has addressed the projected low response rate by piloting and modifying the survey multiple times throughout its development to refine the survey language

and increase the number of responders, particularly participants with low education levels. Recognizing that the survey has already been implemented, it was suggested that, for future surveys, the CERC Survey Workgroup consider simplifying the biobank scenario vignettes presented in the survey to maximize participation. If information on participants' biobank status is available, the ESP wanted the CERC Survey Workgroup to compare responses from survey participants enrolled in biobanks to responses from participants who have never been enrolled in biobank research. As a potential Phase III project, the ESP also recommended integrating a study of consent approaches into actual enrollment into biobanks, i.e. by randomizing participants to different currently used approaches for consent.

- The eMERGE-PGx Workgroup seems to be moving forward in conducting this project; however, overall success for the project will depend on whether the participants in this study are using the selected drugs or have the desired variants. The workgroup has effectively demonstrated its ability to obtain and integrate results into EHR for clinical use. The ESP agreed that it will be worthwhile for the eMERGE-PGx Workgroup to continue harmonizing multiple platforms across multiple sites as this reflects real-world platform diversity among hospital systems. It was suggested that data comparing EHR integration of results across different platforms should be made publically available.

### ESP Recommendations

1. The CERC Survey Workgroup should consider, for future surveys, simplifying the biobank scenario vignettes to maximize participation.
2. If information on participants' biobank status is available, the CERC Survey Workgroup should compare responses from survey participants enrolled in biobanks to responses from participants who have never been enrolled in biobank research.
3. For Phase III, the CERC Survey Workgroup should consider integrating a study of consent approaches into actual enrollment into biobanks, e.g., by randomizing participants to different currently used approaches for consent. The primary recommendation is to study actual attempts to recruit and enroll biobank participants using different approaches, rather than engaging in study of hypothetical scenarios.
4. The eMERGE-PGx Workgroup should continue to harmonize multiple platforms across multiple sites as this reflects real-world platform diversity among hospital systems.
5. The eMERGE-PGx Workgroup should make the data comparing EHR integration of results across different platforms publically available

