Use of EHR for Understanding disease

Marylyn D Ritchie, PhD, FACMI
Department of Genetics
Institute for Biomedical Informatics
Penn Center for Precision Medicine
University of Pennsylvania
Disclaimer
This certifies that the views expressed in this presentation are those of the author and do not reflect the official policy of the NIH.

Disclosure
This certifies that I, Marylyn D Ritchie, have no financial relationship that is relevant to the subject matter of this presentation.
Electronic Health Records (EHR) have become mainstream

- **Purpose**
- Medical care
- Billing insurance
- Ordering procedures
- Ordering medications
- Scheduling
High quality phenotypes/traits for clinical research

- Diagnosis codes
- Medications
- Vital signs
- Lab measurements
- Procedure codes/measures
- Clinical notes

PheKB - a knowledgebase for discovering phenotypes from electronic medical records

What is the Phenotype KnowledgeBase?

Health data is becoming an increasing important source for clinical and genomic research. Researchers create and refine algorithms using structured and unstructured data to better identify cohorts of subjects within the health data.

The Phenotype KnowledgeBase, PheKB, is a collaborative environment to building and validating electronic algorithms to identify characteristics of patients within health data. PheKB was functionally designed to enable...
Utility of medical biobanks

Biospecimens

Genomics

Transcriptomics

Proteomics

Metabolomics
Utility of medical biobanks

GWAS/PheWAS

Polygenic Risk Scores (PRS)
EHR-linked Biobank versus Epidemiology Study

Heart failure cases  Heart failure controls  Sleep apnea cases  Sleep apnea controls

Schizophrenia cases  Schizophrenia controls  Type II Diabetes cases  Type II Diabetes controls
EHR-linked Biobank versus Epidemiology Study

- Heart failure cases
- Sleep apnea cases
- Schizophrenia cases
- Type II Diabetes cases
- Breast cancer cases
- Rheumatoid arthritis cases
- Chronic kidney disease cases
Discovery science using biobanks + EHR

Phenotype-first approach

Disease of interest → Phenotype algorithm → Genetic analysis → Biological follow-up

Genotype/genome-first approach

Gene of interest → PheWAS → Chart review → Biological follow-up
Phenotype-first approach - Lipids

• Goal: to identify genes associated with lipid traits and investigate the relationship between lipids, genotypes, gene expression, and complex human disease

- Lipid GWAS
- Transcriptome-wide association study (TWAS)
- Xpress-PheWAS on Lipid associated genes
- Putative causal associations with lipids
  • Mendelian randomization
- Identified 67 novel lipid associated genes
- Evidence of pleiotropy between lipid and diseases

Veturi Y, ... Ritchie MD. Nature Genetics volume 53, pages 972–981 (2021)
Phenotype-first approach - Lipids

- 16 chromosomes show evidence of pleiotropy
- Diseases/phenotypes
  - Skin and subcutaneous tissue
  - Respiratory
  - Nervous system
  - Neoplasms
  - Musculoskeletal
  - Mental/behavioral traits
  - Genitourinary
  - Endocrine and metabolic
  - Eye
  - Digestive
  - Circulatory

Veturi Y, ... Ritchie MD. Nature Genetics volume 53, pages 972–981 (2021)
What about the other data?

We have these

Diagnoses

Imaging

Clinical lab measures

Genomics
What about the other data?

We have these

Diagnoses
Imaging
Clinical lab measures
Genomics

We want these

99% of a person’s life experiences happen outside of health care system
Genotype-first approach

Genes from Biological Pathway

**ARTICLE**

Mendelian pathway analysis of laboratory traits reveals distinct roles for ciliary subcompartments in common disease pathogenesis

Theodore George Drivas,1,2,4,* Anastasia Lucas,1 Xinyuan Zhang,1 and Marylyn DeRiggi Ritchie1,3,5,*

**Summary**

Rare monogenic disorders of the primary cilium, termed ciliopathies, are characterized by extreme presentations of otherwise common diseases, such as diabetes, hepatic fibrosis, and kidney failure. However, despite a recent revolution in our understanding of the cilium’s role in rare disease pathogenesis, the organelle’s contribution to common disease remains largely unknown. Hypothesizing that common genetic variants within Mendelian ciliopathy genes might contribute to common disease pathogenesis, we performed association studies of 16,874 common genetic variants across 122 ciliary genes with 12 quantitative laboratory traits characteristic of ciliopathy syndromes in 452,993 individuals in the UK Biobank. We incorporated tissue-specific gene expression analysis, expression quantitative trait loci, and Mendelian disease phenotype information into our analysis and replicated our findings in meta-analysis. 101 statistically significant associations were identified across 42 of the 122 examined ciliary genes (including eight novel replicating associations). These ciliary genes were widely expressed in tissues relevant to the phenotypes being studied, and eQTL analysis revealed strong evidence for correlation between ciliary gene expression levels and laboratory traits. Perhaps most interestingly, our analysis identified different ciliary subcompartments as being specifically associated with distinct sets of phenotypes. Taken together, our data demonstrate the utility of a Mendelian pathway-based approach to genomic association studies, challenge the widely held belief that the cilium is an organelle important mainly in development and in rare syndromic disease pathogenesis, and provide a framework for the continued integration of common and rare disease genetics to provide insight into the pathophysiology of human diseases of immense public health burden.


---

Genes from Across the Exome

**LETTERS**

https://doi.org/10.1038/s41591-020-1133-8

Exome-wide evaluation of rare coding variants using electronic health records identifies new gene-phenotype associations

Joseph Park1,2,3, Anastasia M. Lucas1,3, Xinyuan Zhang1,3, Kumardeep Chaudhary2,4,6, Judy H. Cho4,5,6, Girish Nadkarni4,5,6, Amanda Dobbny4,5,6, Geetha Chittoor1,2, Navya S. Josyula1,2, Nathan Katz6, Joseph H. Breyer4,7, Shadi Ahmadmehrabi1, Theodore G. Drivas8, Venkata R. M. Chavali9, Maria Fasolino10,11, Hisashi Sawada12,13, Alan Daugherty11,2, Yannming Li14,15, Chen Zhang15,14, Yuki Bradford13, JoEllen Weaver16, Anurag Verma13, Renae L. Judy16, Rachel L. Kemer1, John D. Overton17, Jeffrey G. Reid17, Manuel A. R. Ferreira17, Alexander H. Li17, Aris Baras17, Scott A. LeMaire17, Ying H. Shen1,14, Ali Naji18, Klaus H. Kaestner19, Golnaz Vahedi19, Todd L. Edwards18, Jinbo Chen19, Scott M. Damrauer16, Anne E. Justice17, Ron Do4,5,6, Marylyn D. Ritchie13 and Daniel J. Rader1,2,13

10.1038/s41591-020-1133-8
EHR-linked biobanks are facilitating tremendous opportunities in translational research for precision health.
Acknowledgements

- Penn Medicine Biobank Team
- Penn TURBO team
- Penn Center for Precision Medicine
- Institute for Biomedical Informatics
- ITMAT
- Penn Genetics
- Penn Medicine Information Services
- Ritchie Lab Members