Huan Mo, MD, MS; Wei-Qi Wei, MD, PhD; Joshua C Denny, MD, MS.

Department of Biomedical Informatics, Vanderbilt University, TN.
Contact: huan.mo@vanderbilt.edu

OBJECTIVE:
1. Introduce EHR-driven phenotyping.
2. Identify causes of current hurdles of cross-institutional collaborations in phenotyping.
3. Develop and implement strategies.

FUNDING
PhEMA: R01 GM105688; iPGx: R01 GM103859.

BACKGROUND
Electronic health records (EHR) are widely used for clinical and translational researches to identify subjects with various phenotypes (e.g., diseases, response to treatments).[1,2] National networks have adopted EHR as their core phenotype data repositories, including Electronic Medical Records and Genomics (eMERGE), Pharmacogenomics Research Network (PGRN), and the National Patient Centered Clinical Research Network (PCORnet).

Although the current EHR infrastructures are reliable for daily healthcare services and data researches within individual institutions, the interoperability is generally poor across different institutes, and phenotype algorithms often require laborious implementation in each secondary site.

What is EHR-Driven Phenotyping

HURDLES
1. EHR data schemas heavily vary in different institutes;
2. Most of these schemas are proprietary and inaccessible by external collaborators;
3. EHR data are often poorly structured, and the complexity makes them difficult to normalize.

STRATEGIES
With other national efforts, we are implementing these strategies to reduce the hurdles for cross-institutional collaborations:
1. Adopt open-source common data models (e.g., i2b2, OMOP) as secondary data repositories;
2. Develop and adopt open APIs of existing EHR data infrastructure to offer connectivity for external applications (e.g., i2b2 messages, SHRINE, FHIR and other HL-7 protocols);[3]
3. Partition phenotype algorithms to common logical workflows and site-specific mounting queries for data elements, such as our KNIME solutions.[4]

REFERENCES