

Data Inputs

Electronic Health
Record (EHR)

eMERGE-seq Panel
Genetic Data



Pre-processing

Constructing
Training and Test

Selecting Relevant
Features



Deriving QDRS

PheRS

Eigen

PC1,2,...

LPC

NMF1

LVS



Genetic Association Studies

Comprehensive Gene-
based Association Tests

Identifying Significant
Association



Clinical Risk Stratification

Plotting Distribution and
Prevalence of Cases

Pinpointing
Undiagnosed Cases



Performance Analysis

Evaluating
AUROC, AUPRC

Robustness to unrelated
features

Feature Weights Analysis