A070 · Replicating Pharmacogenomic Associations in All of Us: An EHR-Based Pipeline for FDA-Labeled Drug-Gene Pairs

Presenter: Julia James — University of Massachusetts Lowell

Authors: Julia James, Rachel Melamed

Session: Day 1 · Thu Oct 1, 2026 · 2:15–4:15 PM

Pharmacogenomics identifies genetic variants that predict drug response, yet most studies rely on randomized trials with homogeneous populations. The All of Us Research Program, with whole genome sequencing (WGS) and longitudinal Electronic Health Record (EHR) data for over 414,000 participants, offers an opportunity to evaluate whether established pharmacogenomics associations replicate in a diverse real-world cohort. We developed an EHR-based pipeline and applied it to two FDA labelled drug-gene pairs: CYP2C19 loss-of-function variants with clopidogrel-associated major adverse cardiovascular events (MACE), and SLCO1B1 rs4149056 with statin-associated myopathy. For each pair, we defined first-prescription cohorts with WGS availability and matched one year of prior EHR data, extracted genotypes, defined outcomes using ICD-coded diagnoses, and ran logistic regression adjusting for age, sex, and ancestry. For the statin analysis, we incorporated sex-specific creatine kinase elevation as a secondary outcome. We also ran a chromosome 12 candidate region GWAS on the SLCO1B1 locus. The clopidogrel cohort (n=7,656, 10.5% carriers) showed no significant MACE association (OR 0.85, p=0.21) but a significant opposite-direction stroke signal (OR 0.68, p=0.04). The atorvastatin cohort (n=55,906, 23.7% carriers) yielded 202 composite myopathy cases; logistic regression showed OR 0.78, opposite to expectation. Rs4149056 was not significant in the GWAS (OR=1.38, p=0.13). Neither association replicated as expected, illustrating key challenges in translating pharmacogenomics findings to biobank-scale cohorts. We are currently running a GWAS to identify novel pharmacogenetic variants associated with statin-induced myopathy beyond the SLCO1B1 locus.

Replicating Pharmacogenomic Associations in All of Us: An EHR-Based Pipeline for FDA-Labeled Drug-Gene Pairs

Julia James*, Dr. Rachel Melamed
University of Massachusetts Lowell

Pharmacogenomics identifies genetic variants that predict drug response, yet most studies rely on randomized trials with homogeneous populations. The All of Us Research Program, with whole genome sequencing (WGS) and longitudinal Electronic Health Record (EHR) data for over 414,000 participants, offers an opportunity to evaluate whether established pharmacogenomics associations replicate in a diverse real-world cohort. We developed an EHR-based pipeline and applied it to two FDA labelled drug-gene pairs: CYP2C19 loss-of-function variants with clopidogrel-associated major adverse cardiovascular events (MACE), and SLCO1B1 rs4149056 with statin-associated myopathy. For each pair, we defined first-prescription cohorts with WGS availability and matched one year of prior EHR data, extracted genotypes, defined outcomes using ICD-coded diagnoses, and ran logistic regression adjusting for age, sex, and ancestry. For the statin analysis, we incorporated sex-specific creatine kinase elevation as a secondary outcome. We also ran a chromosome 12 candidate region GWAS on the SLCO1B1 locus. The clopidogrel cohort (n=7,656, 10.5% carriers) showed no significant MACE association (OR 0.85, p=0.21) but a significant opposite-direction stroke signal (OR 0.68, p=0.04). The atorvastatin cohort (n=55,906, 23.7% carriers) yielded 202 composite myopathy cases; logistic regression showed OR 0.78, opposite to expectation. Rs4149056 was not significant in the GWAS (OR=1.38, p=0.13). Neither association replicated as expected, illustrating key challenges in translating pharmacogenomics findings to biobank-scale cohorts. We are currently running a GWAS to identify novel pharmacogenetic variants associated with statin-induced myopathy beyond the SLCO1B1 locus.