Title: Interaction effects between established non-genetic risk factors and genetically regulated protein levels influencing the risk of coronary artery disease: A large-scale analysis of the All of Us cohort

Authors: Nubaira Rizvi (LSU-Health), Shuai Liu (LSU-Health), Lang Wu (LSU-Health)*, Jingjing Zhu (LSU-Health)* (*joint corresponding authors)

Background: While specific proteins have been identified as risk factors for coronary artery disease (CAD), whether these associations are modified by established non-genetic risk factors remains unknown. A better understanding of these potential interactions could substantially improve our knowledge of disease etiology and facilitate the development of improved risk assessment strategies. This study aimed to comprehensively characterize such potential interaction effects between genetically regulated protein levels and established risk factors including age, sex, and metabolic state on CAD risk in European populations.

Methods: Using MESA-derived protein genetic prediction models and large-scale CAD GWAS summary statistics, we identified 355 CAD-associated proteins in European populations. We applied these genetic prediction models to *All of Us* cohort whole-genome sequencing data to calculate genetically regulated protein levels for 162,509 participants (25,511 CAD cases; 136,998 controls). Multivariable logistic regression tested 254 mapped proteins for interactions with seven covariates: age, sex, BMI, smoking status, hypertension status, hyperlipidemia status, and type 2 diabetes status. Significance was determined via Likelihood Ratio Tests (LRT) comparing full interaction against reduced main-effects models, adjusting for baseline covariates and ten principal components to control for potential population stratification.

Results: Of 1,775 interaction models tested, 158 demonstrated significant interactions (LRT $p < 0.05$) across 97 unique proteins, forming two profiles. First, we identified 47 CAD risk associated proteins whose effects were amplified or attenuated by covariates. The strongest interaction occurred between seq.10613.33 and BMI ($p = 2.8 \times 10^{-5}$). Risk for proteins seq.3054.3 and seq.4154.57 shifted significantly based on age ($p = 0.0011$and 0.0021), while seq.9253.52 was altered by BMI and sex ($p = 1.4 \times 10^{-4}$, $p = 7.7 \times 10^{-4}$). Second, 50 proteins did not show significant associations with CAD risk in this dataset but became significant within specific strata of relevant covariates. Key interactions included seq.3290.50 with BMI ($p = 0.0026$), seq.9314.9 with smoking ($p = 0.0032$), and a distinct cluster (e.g., seq.22796.17, seq.15606.19) in Type 2 Diabetes patients. Overall, BMI was the dominant modulator, driving 56% (28/50) of the most significant interactions.

Conclusions: In the current work, associations of many CAD related proteins tend to show differences according to specific strata of key metabolic and demographic related risk factors. BMI acts as a primary modulator, unmasking potent risk factors within specific states like Type 2 Diabetes. If validated by future independent studies, such findings could improve the etiologic understanding of this common disease and facilitate its risk assessment by integrating these context-dependent interactions.