PEDF a Pleiotropic WTC-LI Biomarker: Machine Learning Biomarker Identification and Validation

G. Crowley[1], J. Kim[1], S. Kwon[2], R. Lam[3], D. J. Prezant[4], M. Liu[5], A. Nolan[6]; [1]Department of Medicine, NYU School of Medicine, New York, NY, United States, [2]Medicine, New York University School of Medicine, New York, NY, United States, [3]Department of Medicine, NYU School of Medicine, New York City, NY, United States, [4]Bureau of Health Services and Office of Medical Affairs, Fire Department of New York, Brooklyn, NY, United States, [5]Department of Population Health, Division of Biostatistics, NYU School of Medicine, New York, NY, United States, [6]Medicine, New York University, New York, NY, United States.

Corresponding author's email: george.crowley@nyumc.org

RATIONALE Metabolic Syndrome biomarkers expressed soon after exposure to World Trade Center-particulate matter (WTC-PM) predict future WTC-Lung Injury (WTC-LI), an obstructive airways disease. Previous panel-based investigation identified metabolic and inflammatory serum biomarkers of WTC-LI. However, there remains unaddressed multicollinearity in our panel-based and high-throughput platform datasets used to assess phenotype proximal to WTC-exposure. This study demonstrates the use of an automated, machine-learning-based, high-dimensional data pruning technique, and validates identified biomarkers in a larger cohort. METHODS Male, never-smoking WTC-LI cases ($FEV_{1,\%Pred}$<LLN at presentation) and controls with normal pre-9/11 lung function were obtained from symptomatic subjects referred for subspecialty pulmonary examination. The parent cohort consisted of cases of WTC-LI (n=100) and controls (n=127). Cases/controls (n=15/group) were chosen for untargeted metabolome assessment if they maintained stable case/control assignment. Feature selection was performed on metabolites, chemokines, cytokines, and clinical data in n=30 subjects. Hierarchical clustering identified potential mechanistic relationships between metabolites, chemokines, and cytokines. Chemokines/cytokines, and clinical biomarkers were validated in the parent cohort (complete cases; n=57 WTC-LI, n=85 controls) via binary logistic regression with 5-fold cross validation. Composite variables replaced significantly correlated variables. Model performance was assessed via receiver-operator-characteristic(ROC) curve analysis. RESULTS 580 metabolites (those detected in ≥80% of subjects/group with relative SD≥15%) were included in random forests analysis alongside clinical biomarkers and previously assayed cytokines and chemokines. The top 5% of biomarkers important to class separation by mean decrease accuracy included pigment epithelium-derived factor (PEDF), macrophage derived chemokine(MDC), systolic blood pressure, macrophage inflammatory protein-4(MIP-4), growth-regulated protein(GRO), monocyte chemoattractant protein-1(MCP-1), Apolipoprotein AII(Apo AII), cell membrane component metabolites (sphingolipids and phospholipids), and branched-chain amino acids. Validation of the top-performing serum chemokines, cytokines, and clinical biomarkers via confounder-adjusted binary logistic regression adjusted for age, body mass index at subspecialty pulmonary examination(SPE), and pre-9/11 lung function had ROC area-under-the-curve of 0.8603(0.7977-0.9228),Figure-1a,b. Decreased PEDF and MIP-4, and increased MDC and Apo AII were associated with increased odds of developing WTC-LI,Figure-1a. CONCLUSIONS An automated data pruning procedure identified several novel biomarkers of WTC-LI in a subcohort of WTC-LI cases and controls with metabolome assessed. These biomarkers' performance was validated in a parent cohort. The strongest predictor was PEDF, a pleiotropic, antiangiogenic agent that has not yet, to the authors' knowledge, been implicated in particulate-matter-related lung disease. Other biomarkers —GRO, MCP-1, MDC, MIP-4—reveal immune cell involvement in WTC-LI pathogenesis. This study represents an important advance in automated biomarker identification. The identified biomarkers warrant further investigation as potential pharmacotherapy targets.

PEDF a Pleiotropic WTC-LI Biomarker: Machine Learning Biomarker Identification and Validation

atsjournals.org/doi/abs/10.1164/ajrccm-conference.2020.201.1_MeetingAbstracts.A7124

American Thoracic Society 2020 International Conference

D55. RESPIRATORY HEALTH OUTCOMES OF WORLD TRADE CENTER EXPOSURES 18 YEARS LATER

American Journal of Respiratory and Critical Care Medicine 2020;201:A7124