DEEP, MULTI-OMICS, BLOOD-BASED BIOMARKER DISCOVERY IN CURRENT AND FORMER SMOKERS DEMONSTRATED HIGH EARLY DETECTION PERFORMANCE FOR NON-SMALL CELL LUNG CANCER

Ehdieh Khaledian[1], Saividya Ramaswamy[1], Manway Liu[1], Lonny Yarmus[2], Madhuvanthi Ramaiah[1], Bruce Wilcox[1], Chinmay Belthangady[1], Brian Koh[1], Philip Ma[1]
[1]PrognomiQ, San Mateo, California, USA; [2]Johns Hopkins University, Baltimore, Maryland, USA

INTRODUCTION

- Lung cancer is the leading cause of cancer-related deaths, with smoking as a major risk factor[1]
- Liquid biopsies are a promising tool to detect lung cancer early, identify genetic mutations that may be targeted with personalized therapies, monitor disease progression and treatment response, and detect cancer recurrence[2]
- However, there are still challenges to overcome before liquid biopsies can be used routinely in clinical practice, including the training and validation of machine learning classifiers able to demonstrate sufficient sensitivity and specificity for lung cancer
- PrognomiQ's multi-omics platform can comprehensively profile mRNA, proteins, and metabolites in blood samples, providing a powerful tool for developing liquid biopsy tests with high sensitivity and specificity for lung cancer

OBJECTIVE

- To investigate the potential of multi-omics assays for detecting and quantifying non-small cell lung cancer (NSCLC) biomarkers in blood samples
- To assess the performance of a machine learning classifier trained to identify NSCLC in a high-risk population

METHODS

Study Design

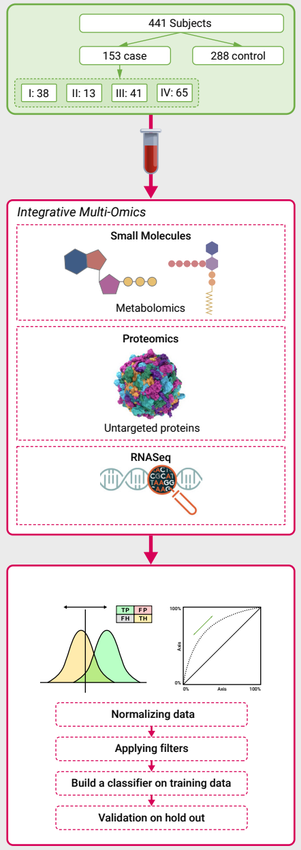
- We previously conducted a case-control study with 1031 subjects, including 361 subjects with untreated NSCLC, 340 subjects without cancer but with significant pulmonary and/or gastrointestinal co-morbidities, and 330 subjects without cancer or co-morbidities
- Using the subset of subjects from this case-control study who were either current or former smokers (441 subjects: 153 NSCLC cases and 288 controls), we developed a machine learning classifier for NSCLC

Sample and Data Processing

- Peripheral blood samples were collected from all subjects for multi-omics plasma analysis
- For this analysis, 3 different omics readouts were used: (1) transcriptomics, (2) liquid chromatography-mass spectrometry (LC-MS)-based proteomics, and (3) metabolomics
- All molecular data were quantified and normalized using standard methods specific to each assay and omics type
- Plasma used for proteomics analysis was processed using the Proteograph™ workflow with a Proteograph™ Assay (Seer Inc.)

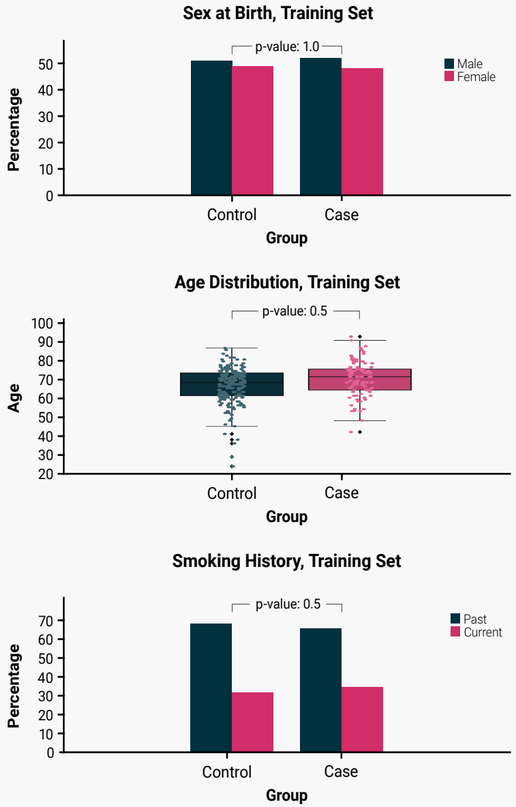
Classifier Training and Validation

- Current or former smokers were divided into a training set of 256 subjects, balanced for important confounders including sex, age, and smoking status (current/past), and used to train the classifier, while the remaining 185 subjects were used as a validation set



RESULTS

FIGURE 1. NSCLC case and control subjects in the training set were balanced for confounders.



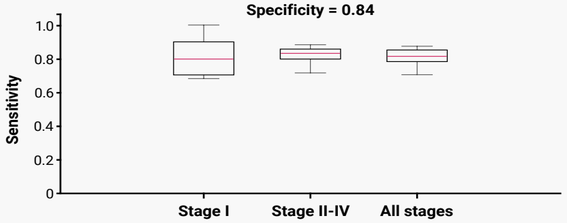
- To avoid bias from confounders, the distribution of subjects used in the training set was balanced for sex at birth, age, and smoking history
- After balancing, the differences between case vs control subjects in the training set were not statistically significant (**Figure 1**)

TABLE 1. Number of features identified for each omics type.

Omics Type	Average number of features per subject
RNA-seq	111,176 transcripts
Proteomics	4,440 proteins 30,063 unique peptides
Metabolomics	1307 metabolites

- A large number of omics data features were identified during classifier training (**Table 1**)
- The PrognomiQ proteomics platform, using the Seer Proteograph™ system, demonstrated performance superior to other published LC-MS-based plasma proteomic studies with ≥500 subjects, which typically detected <600 proteins on average per subject

FIGURE 2. The multi-omics classifier demonstrated high sensitivity for early- and late-stage NSCLC in validation.



- The validation set had a classification AUC of 0.93 with a 95% confidence interval (CI) of 0.89-0.96
- Across all stages, the combined multi-omic classifier had a sensitivity of 82% (CI: 69-90%) at a specificity of 84% (CI: 76-92%) on the validation set (**Figure 2**)
- Stage I sensitivity was 80% (CI: 67-100%) at a specificity of 84% (CI: 75-93%) and Stage II-IV sensitivity was 83% (CI: 76 d-92%) at a specificity of 84% (CI: 75-93%) (**Figure 2**)

CONCLUSIONS

- This NSCLC classifier, utilizing a multi-omics approach and an unprecedented proteomic interrogative depth and scale, demonstrated high performance in early- and all-stage NSCLC detection in a population of current and former smokers
- These encouraging data serve as the foundation for the development of a multi-omics assay for peripheral blood-based liquid biopsies for the early detection of lung cancer

REFERENCES

1. https://www.cancer.org/cancer/lung-cancer/about/key-statistics.html
2. Liquid biopsy in lung cancer: a perspective from members of the Pulmonary Pathology Society https://erj.ersjournals.com/content/54/3/1900871

DISCLOSURES

Study funded by PrognomiQ, Inc. All authors are current or former employees of PrognomiQ, Inc.

ACKNOWLEDGEMENTS

Funded by PrognomiQ, Inc (San Mateo, CA). Editorial and graphical assistance provided by Prescott Medical Communications Group (Chicago, IL).