Background To date, predicting response to immune checkpoint blockade (ICB) therapy in non-small cell lung cancer (NSCLC) patients is based on tumor PD-L1 levels. However, available assays are only moderately predictive, and most require a tumor biopsy. Here, we describe a novel machine learning-based biomarker model that analyzes proteomic profiles in blood plasma to predict ICB response in NSCLC patients.

Methods We collected plasma samples and clinical data from 339 ICB-treated NSCLC patients via a multi-center clinical study (PROPHETIC; NCT04056247; approved by local IRB committees from each site), 60% of them received combination of ICB-chemotherapy and the rest received ICB alone. Patients displaying disease progression were classified as non-responders and the rest as responders. Proteomic profiling was performed by the SOMAscan assay. A machine-learning-based model for clinical response prediction was developed based on protein expression level in patient’s plasma. Using a proprietary algorithm, we identified Response Associated Proteins (RAPs), that serve as potential indicators of clinical response depending on their plasma level in the individual patient. The output of the model provides a patient-specific response probability for 3, 6, and 12 months after starting treatment.

Results The RAP-based model displayed strong predictive power over the first year of ICB-based therapy, as indicated by area under the curve (AUC) of the receiver operating characteristics (ROC) plot of 0.71, 0.77 and 0.78 for 3-, 6- and 12-months following treatment initiation, respectively, and a high goodness of fit between predicted response probability and observed response rate ($R^2 = 0.97$). Patients with low and high response probability predictions displayed a significant difference in overall survival and progression-free survival. The RAP-based model outperformed a PD-L1-based model (AUC of 0.5, 0.6 and 0.55 for 3-, 6- and 12-months, respectively). Notably, in a subgroup of patients with PD-L1-high tumors (>50% PD-L1) receiving monotherapy, patients with high response probability predictions survived significantly longer than patients with low response probability predictions (p-value 0.0002, 0.0036 and 0.0115 for 3-, 6- and 12-months, respectively) who had similar overall survival as patients with PD-L1-low tumors (<50% PD-L1).

Conclusions Altogether, we have developed a novel predictive model for ICB response in NSCLC patients based on proteomic profiling of blood plasma. The model offers two main clinical utilities. First, it provides response predictions for three time points over the first year of treatment. Second, it identifies a subgroup of high PD-L1 patients who may benefit more from combination therapy.