Abstract accepted for presentation at ATS 2024

Combination of multigene profiles and clinical factors improves prediction of short-term outcomes in idiopathic pulmonary fibrosis

Aparna C Swaminathan,[1,2] Megan L Neely,[1,2] Panagiotis Papavasileiou,[3] Christian Hesslinger,[3] Imre Noth,[4] Justin M Oldham,[5] Thomas Schlange,[3] Ramona Schmid,[3] Thomas B Leonard,[6] Jamie L Todd[1,2] on behalf of the IPF-PRO Registry investigators

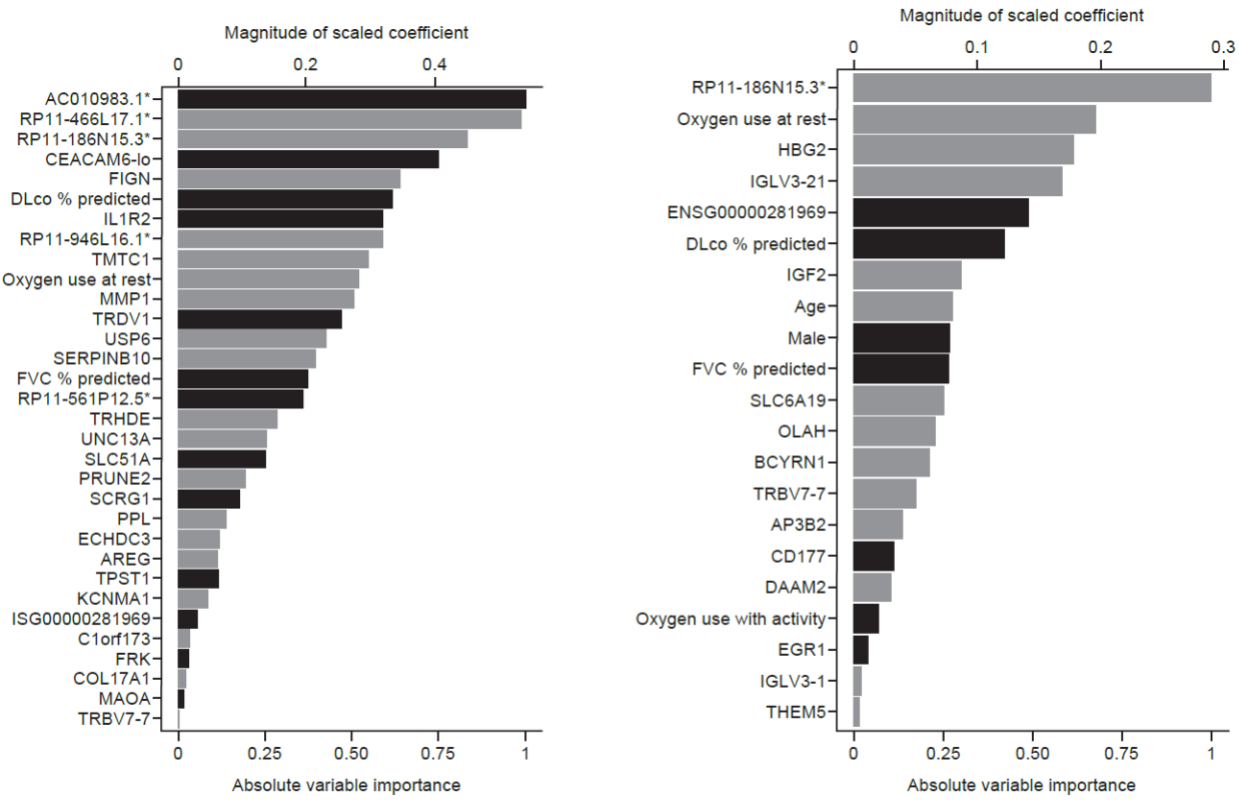
[1]Duke Clinical Research Institute, Durham, North Carolina, USA; [2]Duke University Medical Center, Durham, North Carolina, USA; [3]Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany; [4]University of Virginia, Charlottesville, Virginia, USA; [5]University of Michigan, Ann Arbor, Michigan, USA; [6]Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, Connecticut, USA.

Rationale: The course of idiopathic pulmonary fibrosis (IPF) is variable. Potential biomarkers of death or progression have been identified, but their predictive value beyond known clinical risks is uncertain. We sought to derive gene-inclusive profiles that discriminate risk for short-term outcomes in patients with IPF and compare their performance with clinical factors alone.
Methods: The cohort comprised 261 patients with IPF from the multicenter IPF-PRO Registry who had whole blood total RNA sequencing at enrollment that met quality parameters (n=261). The outcomes of death or IPF progression (composite of ≥10% decline in FVC % predicted, death, or lung transplant) were assessed at 12 months post-enrollment. DESeq2 was used to determine differential gene expression among patients with versus without each outcome. To detect pathways enriched for differentially expressed genes, over-representation analysis was conducted using clusterProfiler. Elastic net logistic regression was used to derive predictive models for death and progression. Each model considered clinical factors at enrollment (lung function, oxygen use, age, sex) and differentially expressed genes with false discovery rate-adjusted p≤0.05 and $\log_2$fold-change >1 as potential predictors. The variable importance of the predictors selected was plotted. Models including only clinical factors were also constructed. Model performance was assessed using the C-index and the optimism-corrected C-index.
Results: At 12 months post-enrollment, 27 patients (10.3%) had died and 66 (25.3%) had experienced progression. Analyses identified 2,005 and 70 genes that were differentially expressed (FDR-adjusted p≤0.05) for death and progression, with 94 of 2,005 and 15 of 70,

respectively, also having a $\log_2$fold-change >1. Differentially expressed genes were enriched in pathways related to the innate immune system (e.g., in neutrophil degranulation [death] and in the complement cascade [progression]). Multivariable risk prediction models selected a set of 29 genes and 3 clinical factors for the outcome of death (optimism-corrected C-index 0.92) and 15 genes and 6 clinical factors for the outcome of progression (optimism-corrected C-index 0.67). Models considering only clinical factors showed worse risk discrimination for both death and progression (optimism-corrected C-indices 0.51 and 0.49, respectively). Long non-coding RNAs were identified as the most important variables in models for both death and progression (Figure).

Conclusions: In patients with IPF, models that include both circulating gene expression and clinical measures have better discriminatory ability for short-term risk of death or progression than models considering only clinical factors. Long non-coding RNAs were important in outcome discrimination, warranting further evaluation of their potential regulatory functions in IPF.

Figure. Variables selected by logistic regression models with Elastic Net (ENet) penalty that considered clinical factors and differentially expressed genes at enrollment as potential predictors for the outcomes of death (left) or IPF progression (right). Variable importance measures were calculated as the absolute value of the scaled regression coefficients divided by the largest coefficient in absolute value. Black bars indicate variables with negative ENet coefficients and grey bars indicate those with positive ENet coefficients.



*Long non-coding RNAs.

Disclosures: The IPF-PRO/ILD-PRO Registry is supported by Boehringer Ingelheim Pharmaceuticals, Inc (BIPI) and run in collaboration with the Duke Clinical Research Institute (DCRI) and enrolling centers. Writing assistance, which was contracted and funded by BIPI, was provided by Fleishman-Hillard, London, UK. Aparna C Swaminathan, Megan L Neely and Jamie L Todd are faculty members of the Duke Clinical Research Institute (DCRI), which receives funding support from BIPI to coordinate the IPF-PRO/ILD-PRO Registry. Panagiotis Papavasileiou, Christian Hesslinger, Thomas Schlange, Ramona Schmid and Thomas B Leonard are employees of Boehringer Ingelheim. Imre Noth reports patents pending on transcriptomic signatures for progression of IPF and proteomic signatures for disease classification. Justin M Oldham reports no disclosures.