Supplemental Material

Identifying early pulmonary arterial hypertension biomarkers in systemic sclerosis: Machine learning on proteomics from the DETECT cohort

Yasmina Bauer¹,², Simon de Bernard³, Peter Hickey⁴,⁵, Karri Ballard⁶, Jeremy Cruz⁶, Peter Cornelisse², Harbajan Chadha-Boreham⁷, Oliver Distler⁸, Daniel Rosenberg⁷, Martin Doelberg⁷, Sebastien Roux², Oliver Nayler² and Allan Lawrie⁴

¹Galapagos GmbH, Aeschengraben 20, CH-4051 Basel, Switzerland, ²Idorsia Pharmaceuticals Ltd, Hegenheimermattweg 91, CH-4123 Allschwil, Switzerland, ³AltraBio, 30 rue Pré-Gaudry, F-69007 Lyon, France, ⁴Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield, Medical School, Beech Hill Road, Sheffield, S10 2RX, UK, ⁵Sheffield Pulmonary Vascular Disease Unit, Royal Hallamshire Hospital, Glossop Road, Sheffield, S10 2JF, ⁶Myriad RBM, 3300 Duval Rd., #110, Austin, TX 78759, USA, ⁷Actelion Pharmaceuticals Ltd, Gewerbestrasse 16, CH-4123 Allschwil, Switzerland, ⁸University Hospital Zurich, Department of Rheumatology, Gloriastrasse 25, CH-8091 Zurich, Switzerland

Corresponding Author
Allan Lawrie
IICD, Medical School
Beech Hill Road, Sheffield, S10 2RX
United Kingdom
a.lawrie@sheffield.ac.uk | Telephone: +44 (0) 114 215 9536
**Supplemental Figure S1**: Comparison of age across the DETECT Discovery and Sheffield Confirmatory Cohorts.

**Supplemental Figure S2**: Comparison of sex across the DETECT Discovery and Sheffield Confirmatory Cohorts.
**Supplemental Figure S3:** Comparison of systemic sclerosis type across the DETECT Discovery and Sheffield Confirmatory Cohorts.

**Supplemental Figure S4:** Comparison of NT-proBNP across the DETECT Discovery and Sheffield Confirmatory Cohorts.
Supplemental Figure S5: Comparison of pulmonary vascular resistance across the DETECT Discovery and Sheffield Confirmatory Cohorts.

Supplemental Figure S6: Comparison of pulmonary arterial wedge pressure across the DETECT Discovery and Sheffield Confirmatory Cohorts.
Supplemental Figure S7: Comparison of mean pulmonary artery pressure across the DETECT Discovery and Sheffield Confirmatory Cohorts.

Supplemental Figure S8: Comparison of forced vital capacity across the DETECT Discovery and Sheffield Confirmatory Cohorts.
**Supplemental Figure S9**: Comparison of diffusion lung capacity for Carbon Monoxide across the DETECT Discovery and Sheffield Confirmatory Cohorts.

**Supplemental Figure S10**: Pearson correlation of RAGE protein levels with FVC % predicted in the DETECT cohort.
Supplemental Figure S11: Serum concentrations of the eight common analytes in DETECT serum patients with diffuse compared to limited SSc. Box and Whisker plots represent the interquartile range (box) with the line representing the median and whisker the full range of the data. Individual patient samples are represented by dots. The values in the top left corners represent the p-value of the Wilcoxon rank-sum test between the two patient groups.