

## Supplementary Tables

Table S1: All variables that were used to create the clusters.

Patient characteristics/co-morbidities	Questionnaires	Lung function	Long COVID symptom categories	Acute phase characteristics
BMI	FSS	FEV1	Fatigue	Oxygen received (type)
Age	PROMIS	FVC	Respiratory	Hospitalized
Sex	PC-PTSD-5	FEV1/FVC	Neurological	Hospital duration
Asthma	EQ5D	DLCO	Cardiovascular	Thrombosis
COPD	CLCIC		Gastrointestinal	Pulmonary embolism
CVD			Other	
Diabetes				

BMI: body mass index, COPD: chronic obstructive pulmonary disease, CVD: cardiovascular disease, FEV1: forced expiratory volume, FVC: forced vital capacity, DLCO: diffusion capacity of the lungs for carbon monoxide.

## Supplementary Figures

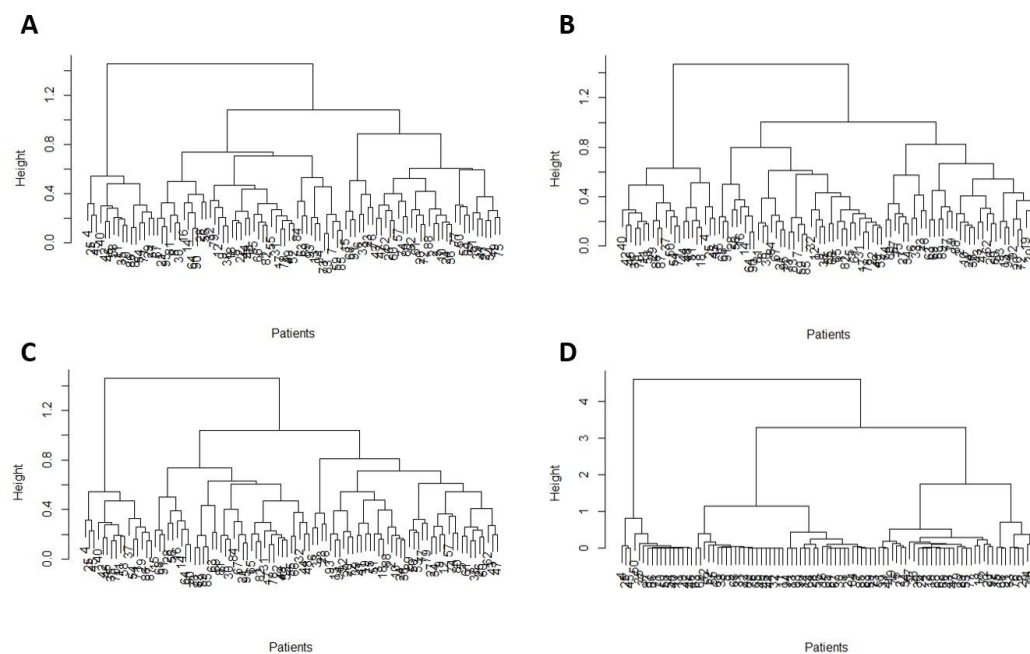


Figure S1: Dendrograms from hierarchal clustering of individual imputed data sets (A-C) and the dendrogram from the consensus clustering (D).

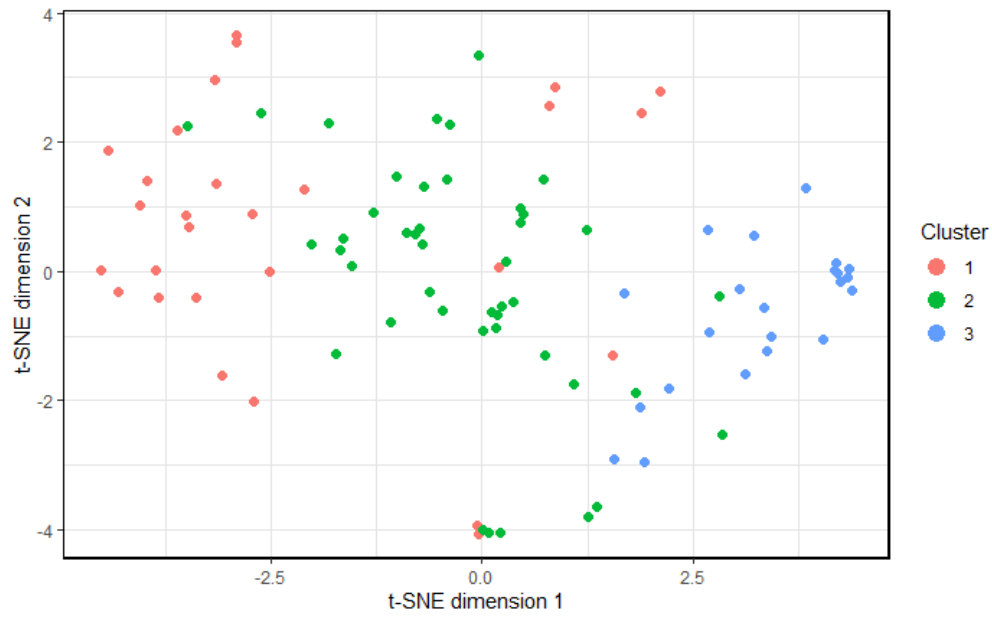


Figure S2: t-SNE plot showing the separation of the clusters created by the PAM clustering method projected on a 2-Dimensional space.