

Extended Data Table 1 Characteristics of participants included in this study.

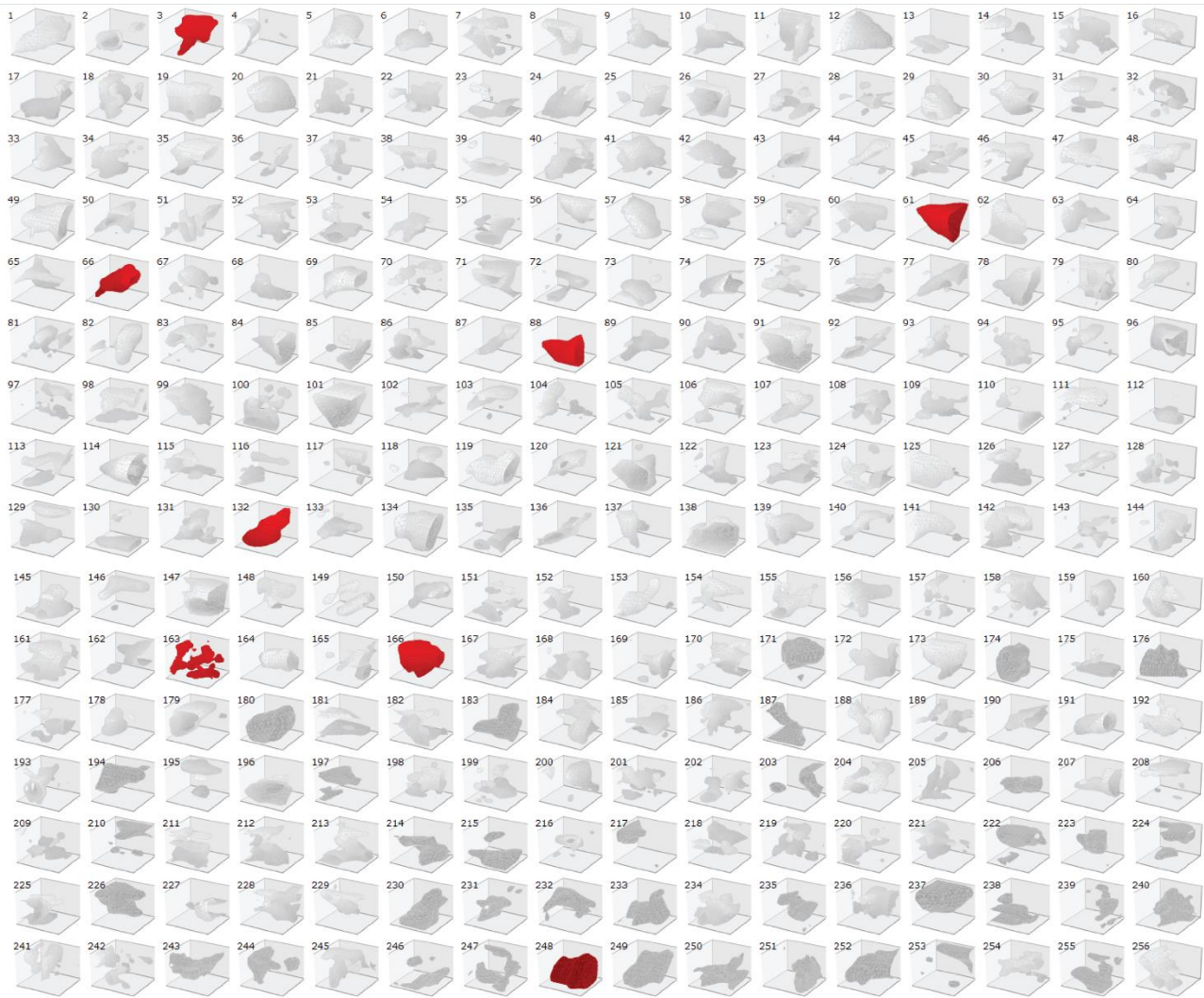
Variables	Training				Validation			
	COVID-19 (n=47)	healthy (n=20)	CAP (n=49)	P value	COVID-19 (n=153)	healthy (n=15)	CAP (n=37)	P value
Age~ Median [Min, Max]	53.0 [31.0, 74.0]	29.0 [14.0, 50.0]	37.0 [16.0, 76.0]	<0.001	64.0 [20.0, 89.0]	31.0 [19.0, 52.0]	32.0 [20.0, 81.0]	<0.001
Gender								
Female	24 (51.1%)	7 (35.0%)	22 (44.9%)	0.477	81 (52.9%)	6 (40.0%)	23 (62.2%)	0.3274
Male	23 (48.9%)	13 (65.0%)	27 (55.1%)		72 (47.1%)	9 (60.0%)	14 (37.8%)	

Abbreviations: COVID-19, Coronavirus Disease 2019; CAP, community acquired pneumonia.

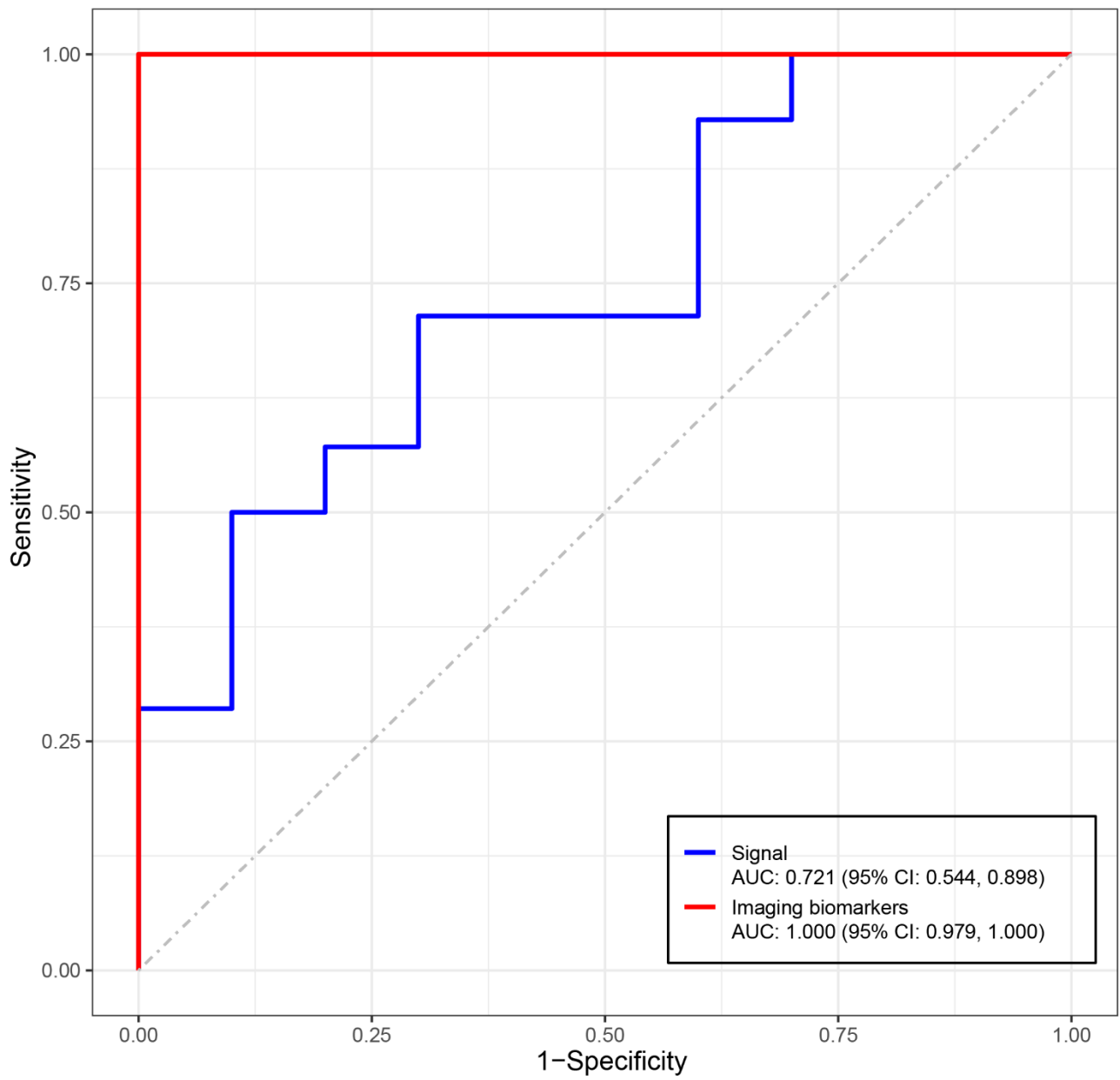
Extended Data Table 2 Significant dictionary elements positively correlated with COVID-19.

Dictionary element	OR	95% CI		P value	FDR
		Lower	Upper		
IB-3	7.71E+36	7.7352E+18	2.26E+57	0.00015289	0.000430108
IB-61	3.25153E+18	359.9820936	2.60E+35	0.026274294	0.0431216
IB-66	9.30E+46	3.97E+27	2.08E+69	8.61E-06	4.24E-05
IB-88	1.37E+21	13586381.21	3.57E+36	0.004540289	0.008673985
IB-132	4.12E+73	8.99E+44	1.32E+107	2.94E-06	1.79E-05
IB-163	4.27E+151	6.10E+101	3.47E+214	1.08E-07	2.55E-06
IB-166	3.09E+28	1.71393E+15	3.88E+43	7.19E-05	0.000224565
IB-248	3.55897E+18	5413396.828	5.74E+31	0.003411235	0.006769584

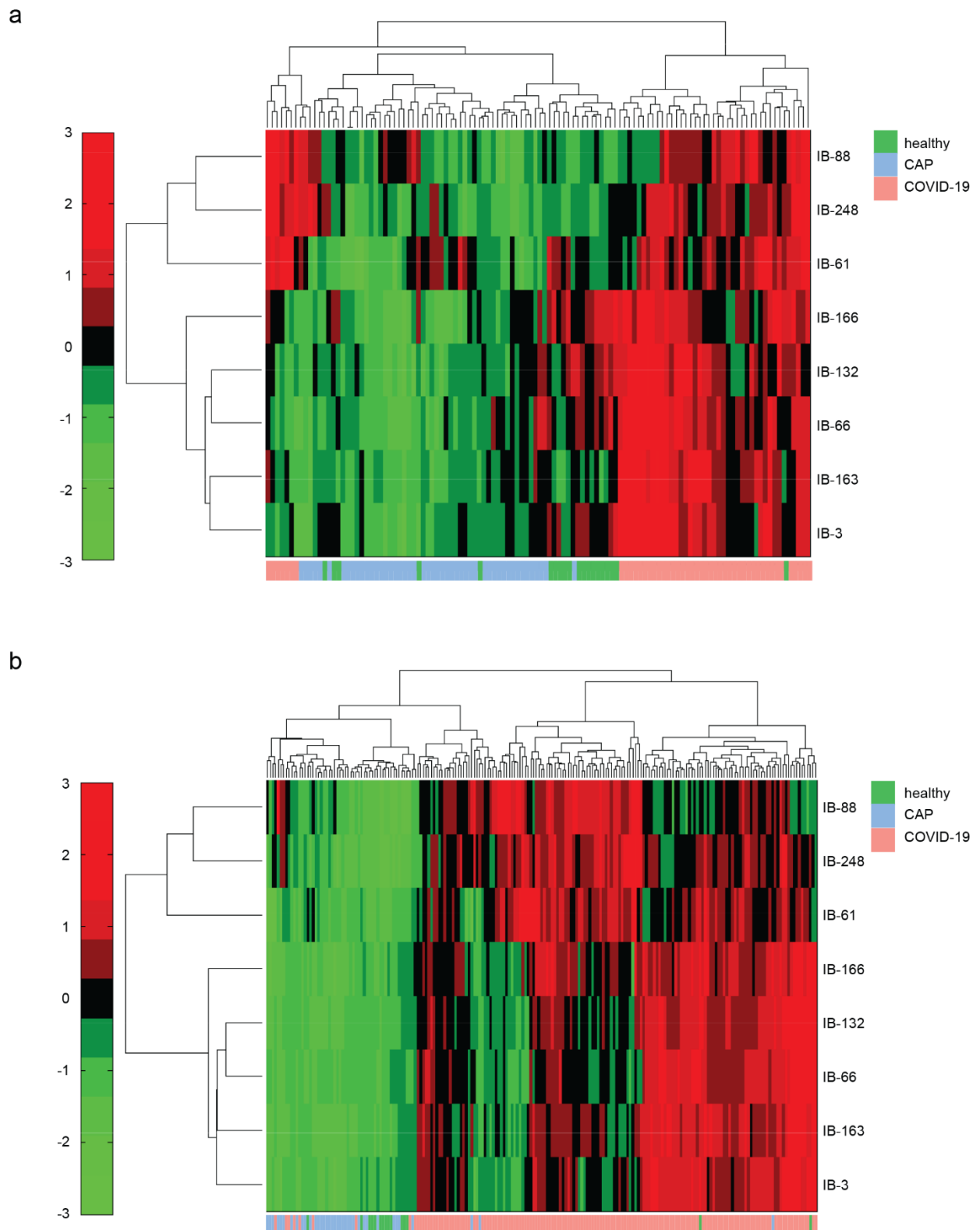
Abbreviation: IB, Imaging biomarker; OR, Odds ratio; FDR, False discovery rate.



Extended Data Fig. 1 Visualization of 256 dictionary elements (i.e., signal patterns) learned from training cohort, where 8 (in red) dictionary elements have significant positive correlation with COVID-19. These 8 dictionary elements, served as our imaging biomarkers, have also been demonstrated in Figure 1.



Extended Data Fig. 2 COVID-19 screening performance of signal-based model (blue line) and imaging biomarker-based model (red line) on training cohort.



Extended Data Fig. 3. The relationship between 8 imaging biomarkers and COVID-19. a-b. Heatmap of relative abundance of 8 imaging biomarkers in both training (a) and validation (b) cohorts.