Supplementary Table 1. Outlying values for continuous variables

Variable	Percentage of outlying values (%)
Age (yr)	2.65
Length of hospital stay (day)	5.63
No. of COVID-19 vaccines (n)	3.31
Duration of invasive ventilation (day)	0.33
Duration of oxygen therapy (day)	1.66
Creatine kinase (U/L)	1.99
Day of illness (day)	3.64
Initiation of oseltamivir (day)	0.99
Total white cell count ($\times10^9$/L)	3.64
Absolute lymphocyte count ($\times10^9$/L)	5.96
Platelet ($\times10^9$/L)	0.99
C-reactive protein (mg/L)	5.63
Total bilirubin (umol/L)	6.29
Albumin (d/L)	1.66
Alanine transaminase (U/L)	7.28
Alkaline phosphatase (U/L)	4.64

COVID-19, coronavirus disease 2019.
The dataset comprised 95 analysable variables, including 77 categorical (inclusive of 15 dummy variables) and 18 continuous variables. No duplicate records were identified. Data quality assessments included checks for outliers, missingness and data normality. Outlier analysis revealed that 16 out of the 18 continuous variables contained extreme values. Among these, alanine aminotransferase exhibited the highest proportion of outliers (7.3%). All detected outliers were retained in the dataset, as they were considered natural outliers within the clinical data.