

Table S3a. COVID-SOFA score's goodness-of-fit for 60-day mortality based on the 28-day mortality model

60-day mortality	
	COVID-SOFA ICU admission
Hosmer-Lemeshow Test	$\chi^2 = 11.45$, $p = 0.17$
Observed <i>vs.</i> Predicted Mortality Probability	0.58 <i>vs.</i> 0.46
Overall percentage correct	75.5%
Akaike Information Criterion (AIC) (value <i>vs.</i> AIC for SOFA)	455 <i>vs.</i> 528
Bayesian Information Criterion (BIC) (value <i>vs.</i> BIC for SOFA)	492 <i>vs.</i> 556
Likelihood ratio test	$p < 0.01$

COVID-SOFA = coronavirus disease 2019 – sequential organ failure assesment, ICU = intensive care unit

Table S3b. A comparative analysis of COVID-SOFA score *vs.* SOFA score discriminative ability for 60-day mortality based on the 28-day mortality model

60-day mortality								
	C-index 95% CI	C-index diff.	Error rate improvement	AUROC 95% CI	AUROC diff.	p	AUPRC 95% CI	AURPC diff. 95% BC(a) CI
COVID-SOFA ICU admission	0.707 0.674-0.739	0.061	17.23%	0.802 0.761-0.839	0.102	< 0.001*	0.831 0.779-0.873	0.076 0.063-0.09
SOFA ICU admission	0.646 0.614-0.678			0.699 0.653-0.743			0.755 0.697-0.805	
COVID-SOFA 48 hours	0.747 0.716-0.778	0.049	16.22%	0.872 0.837-0.902	0.079	< 0.001*	0.888 0.842-0.922	0.084 0.068-0.103
SOFA 48 hours	0.698 0.665-0.732			0.793 0.751-0.831			0.803 0.749-0.848	

*DeLong (12); COVID-SOFA = coronavirus disease 2019 – sequential organ failure assessment, ICU = intensive care unit, diff. = difference, AUROC = area under the receiver operating characteristic curve, AUPRC = area under the precision-recall curve, 95% CI = 95% confidence interval, BC(a) = bias corrected accelerated bootstrapping