

Supplementary material for submission #7091:
Deep survival analysis with longitudinal X-rays for COVID-19

1 Parameter sensitivity study

In the main paper, we investigated both the number of images and the size of mini batch, and concluded that 1) the model improves as the number of images increases for the majority of tasks (admission, ICU discharge, and discharge) and 2) mini-batch size does not affect the model performance for sufficiently large values of B ($B \geq 20$).

Table 1 summarizes the results.

Admission	$k = 8$	$k = 4$	$k = 2$	$k = 1$
$B = 20$	0.195	0.191	0.180	0.272
$B = 40$	-	0.171	0.201	0.303
$B = 80$	-	-	0.198	0.232
$B = 142$ (All)	-	-	0.193	0.252
ICU Admission	$k = 8$	$k = 4$	$k = 2$	$k = 1$
$B = 20$	0.259	0.258	0.258	0.285
$B = 40$	-	0.234	0.241	0.287
$B = 80$	-	-	0.254	0.278
$B = 160$	-	-	0.237	0.277
ICU Discharge	$k = 8$	$k = 4$	$k = 2$	$k = 1$
$B = 20$	0.345	0.331	0.306	0.427
$B = 40$	-	0.335	0.337	0.366
$B = 80$	-	-	0.327	0.341
$B = 160$	-	-	0.330	0.358
Discharge	$k = 8$	$k = 4$	$k = 2$	$k = 1$
$B = 20$	0.230	0.231	0.229	0.240
$B = 40$	-	0.229	0.232	0.233
$B = 80$	-	-	0.231	0.235
$B = 160$	-	-	0.228	0.238
Death	$k = 8$	$k = 4$	$k = 2$	$k = 1$
$B = 20$	0.268	0.271	0.279	0.262
$B = 40$	-	0.247	0.346	0.239
$B = 80$	-	-	0.232	0.249
$B = 160$	-	-	0.238	0.239

Table 1: Parameter sensitivity study on mini-batch size B and maximum number of x-rays per patient k for Discharge and ICU Discharge Predictions. Performance is measured by concordance error, lower is better. The experiments are conducted on hospitalization (admission) prediction. When $B = 142$, all available patients from the training set are included in the mini-batch.

	ICU Admit	ICU Discharge	Discharge
Event #	420	252	1197
Censored #	1099	173	333

Table 2: Details on the data distribution for each type of events for unfiltered data ablation test.

Method	ICU Admit	ICU Discharge	Discharge
CoxPH (Full)	0.414	0.373	0.273
Weibull (Full)	0.331	0.324	0.264
Gempertz (Full)	0.345	0.326	0.185
DeepSurv	0.358	0.337	0.186
Survival+SVM	0.355	0.397	0.215
Survival+Random Forest	0.286	0.307	0.206
Ours	0.302	0.238	0.173

Table 3: Concordance error for time-to-event predictions. Lower is better, colors encode the best performing 3 methods for each event. Full models take input concatenated across all available days, while cross-sectional (CS) models only take time-dependent data from the baseline date, which satisfies the PH assumption.

2 Non-Image Data Only Comparisons

We conduct an ablation test on data filtering. As mentioned in our data pipeline, to compare with models that only take images as input, we removed patients that do not have any x-rays taken during the target interval. As a sanity check, the study below uses unfiltered data and compares non-image only models.

Table 3 compares results of ICU Admit, ICU Discharge and (Hospital) Discharge predictions, where there are substantial population changes when we applied the filter. We keep the same event-censor ratio, and split the data randomly into train, eval, and test by 60 : 20 : 20.

The results provide evidence that our technique works well even in the absence of images, independent of the data filtering.

Type	Abbreviation	Unit
Alanine Transferase	ALT	<i>U/L</i>
Glucose Level	-	<i>mg/dL</i>
Albumin Level	Alb	<i>g/dL</i>
Aspartase Transaminase	AST	<i>U/L</i>
Ala Aminotransferase	ALA	<i>U/L</i>
Aspartate Aminotransferase	ASA	<i>U/L</i>
Creatinine	Cr	<i>mg/dL</i>
Lymphocyte	-	<i>cells/uL</i>
Creatine Kinase	CK	<i>U/L</i>
Lactate Dehydrogenase	LDH	<i>U/L</i>
Ferritin Level	-	<i>ng/mL</i>
B-Type Natriuretic Peptide	BNP	<i>pg/mL</i>
Troponin I	Trop	<i>ng/mL</i>
Sedimentation Rate	ESR	<i>mm/hr</i>
C-Reactive Protein	CRP	<i>mg/dL</i>
D-Dimer,	-	<i>ng/mL</i>
Interleukin-6	IL-6	<i>pg/mL</i>
Platelets	-	$\times 10^3/uL$
Thromboplastin Time	PTT	<i>s</i>
vs_hr_hr	-	-
xp_resp_spo2	-	-
xp_resp_rate_pt	-	-
vs_bp_noninvasive (s)	-	-
vs_bp_noninvasive (d)	-	-
vs_bp_noninvasive (m)	-	-
HCO3 (Arterial)	-	<i>mmol/L</i>
pO2 (Arterial)	-	<i>mmHg</i>
pCO2 (Arterial)	-	<i>mmHg</i>
pH (Arterial)	-	-

Table 4: List of time-dependent features used in our technique.

	Admit	ICU Admit	ICU Discharge	Discharge	Death
Mean	2.72	1.78	6.65	2.88	4.61
Std	3.39	1.11	4.61	3.39	4.00

Table 5: Details on the number of images per patient within data range