Predicting Mortality Risk in Viral and Unspecified Pneumonia to Assist Clinicians with COVID-19 ECMO Planning

Helen Zhou*, Cheng Cheng*,
Zachary C. Lipton, George H. Chen, Jeremy C. Weiss

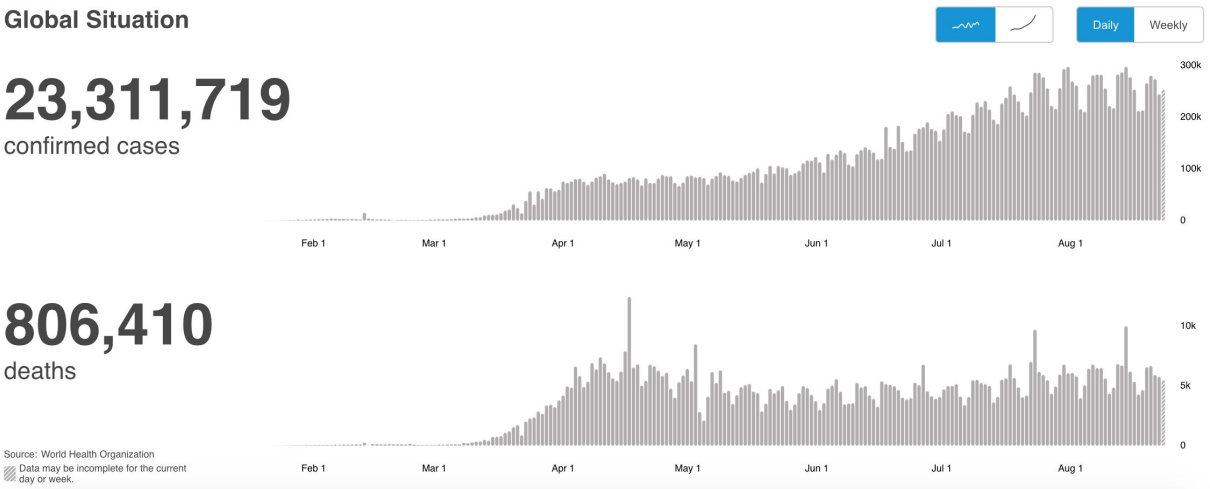
Machine Learning Department, Heinz College of Information Systems and Public Policy,
Carnegie Mellon University

* equal contribution

Motivation

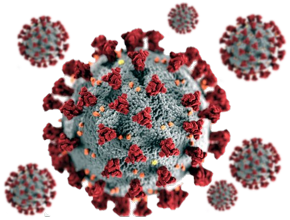
World Health Organization

Globally, as of **2:22pm CEST, 24 August 2020**, there have been **23,311,719 confirmed cases** of COVID-19, including **806,410 deaths**, reported to WHO.



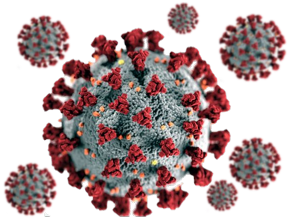
(WHO Coronavirus Disease Dashboard.)

Many concerning cases of COVID-19 progress from...



Severe Acute Respiratory Syndrome (SARS-CoV-2)

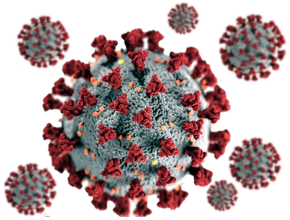
Many concerning cases of COVID-19 progress from...



Severe Acute Respiratory Syndrome (SARS-CoV-2)

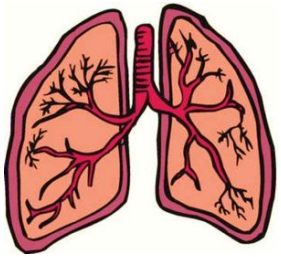
Viral pneumonia

Many concerning cases of COVID-19 progress from...



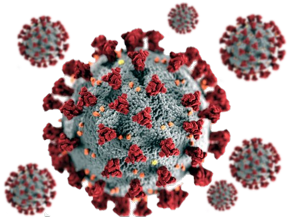
Severe Acute Respiratory Syndrome (SARS-CoV-2)

Viral pneumonia



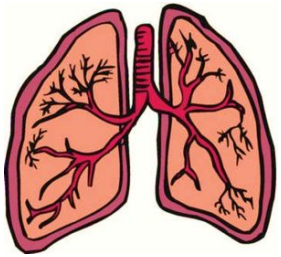
Acute respiratory distress syndrome (ARDS)

Many concerning cases of COVID-19 progress from...



Severe Acute Respiratory Syndrome (SARS-CoV-2)

Viral pneumonia

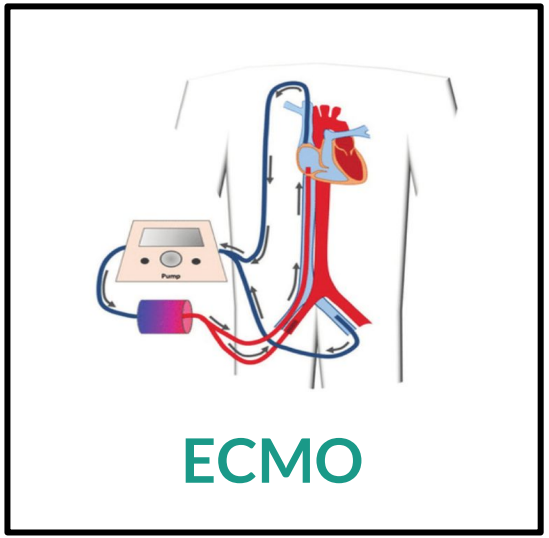


Acute respiratory distress syndrome (ARDS)

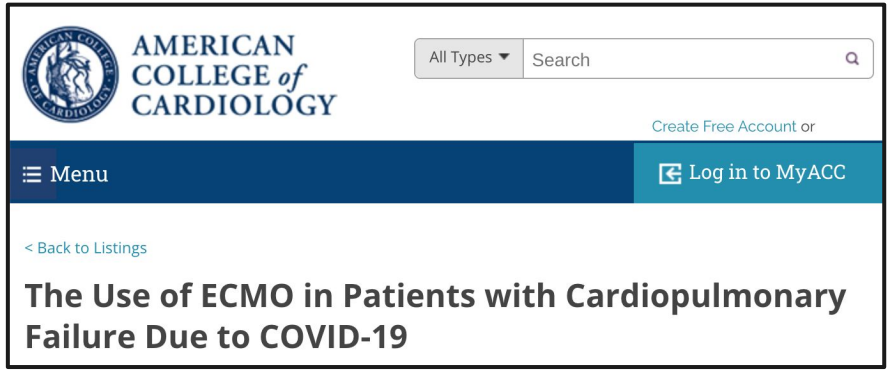
Death

When mechanical ventilation is insufficient to oxygenate the lungs, ECMO can temporarily sustain the patient.

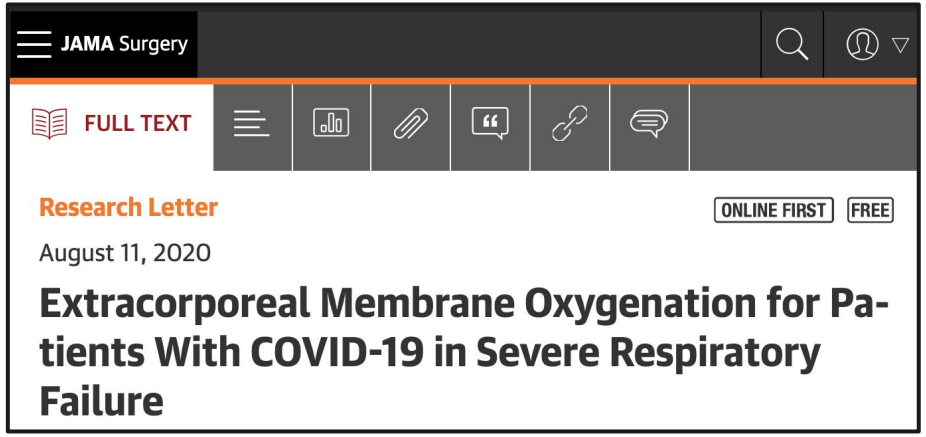
When mechanical ventilation is insufficient to oxygenate the lungs, ECMO can temporarily sustain the patient.



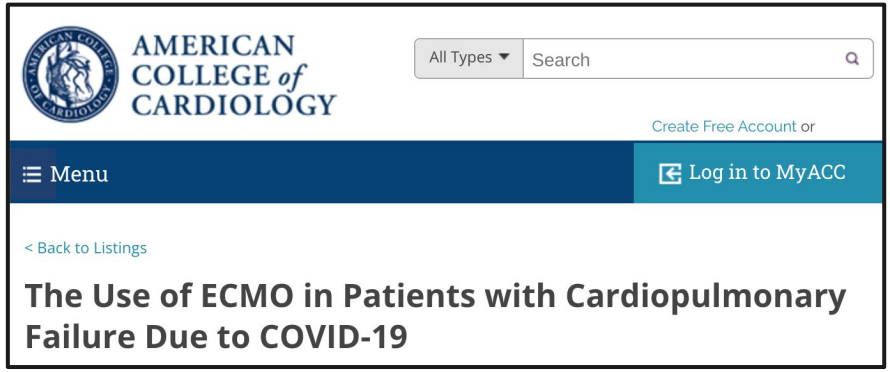
(Fadelelmoula, 2020)



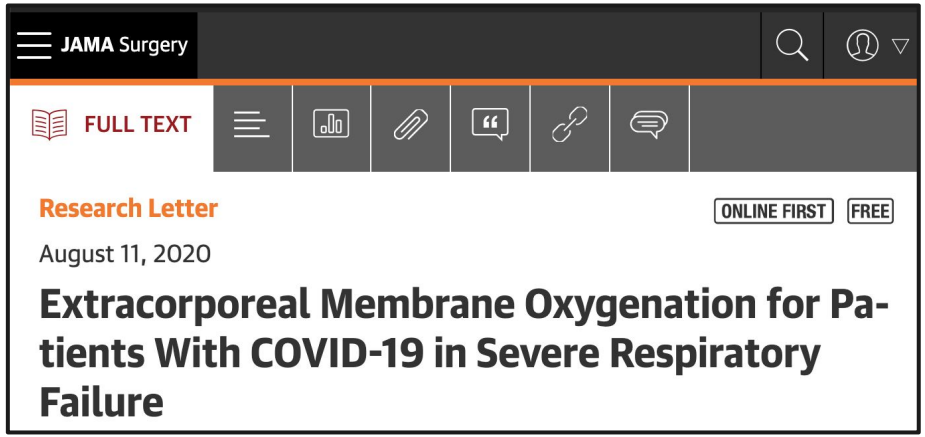
(Fadelelmoula, 2020)



(Krieger and Badulak, 2020)


AMERICAN COLLEGE of CARDIOLOGY
All Types
Search
Create Free Account or
Menu
Log in to MyACC
< Back to Listings
The Use of ECMO in Patients with Cardiopulmonary Failure Due to COVID-19

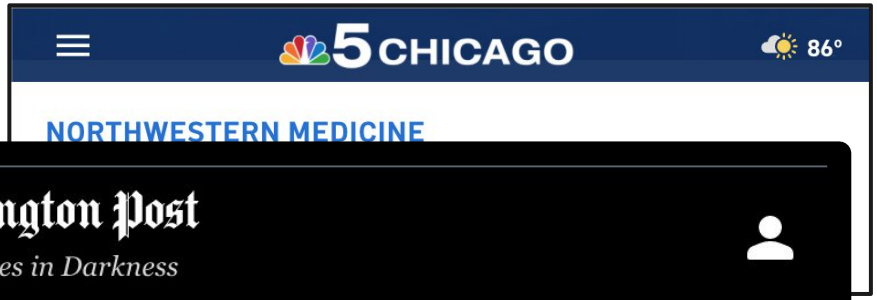

5 CHICAGO
86°
NORTHWESTERN MEDICINE
ECMO Machine Used to Save COVID-19 Patient's Life, Officials Say


JAMA Surgery
FULL TEXT
Research Letter
ONLINE FIRST
FREE
August 11, 2020
Extracorporeal Membrane Oxygenation for Patients With COVID-19 in Severe Respiratory Failure


HEALTH LAB
ROUNDS
Kara Gavin
April 21, 2020 3:28 PM
When Ventilators Don't Help COVID-19 Patients, This Might


25 NEWS
Weather
Live Stream
Video
News
Sports
Log In / Register
First COVID-19 patient saved in Boston with ECMO
Life support system 'last stop' if ventilators fail

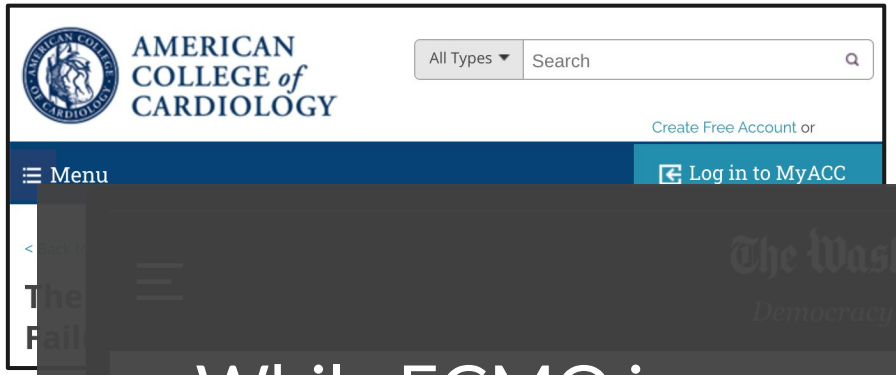

AMERICAN COLLEGE of CARDIOLOGY
All Types
Search
Create Free Account or
Menu
Log in to MyACC
< Back to
The
Fail

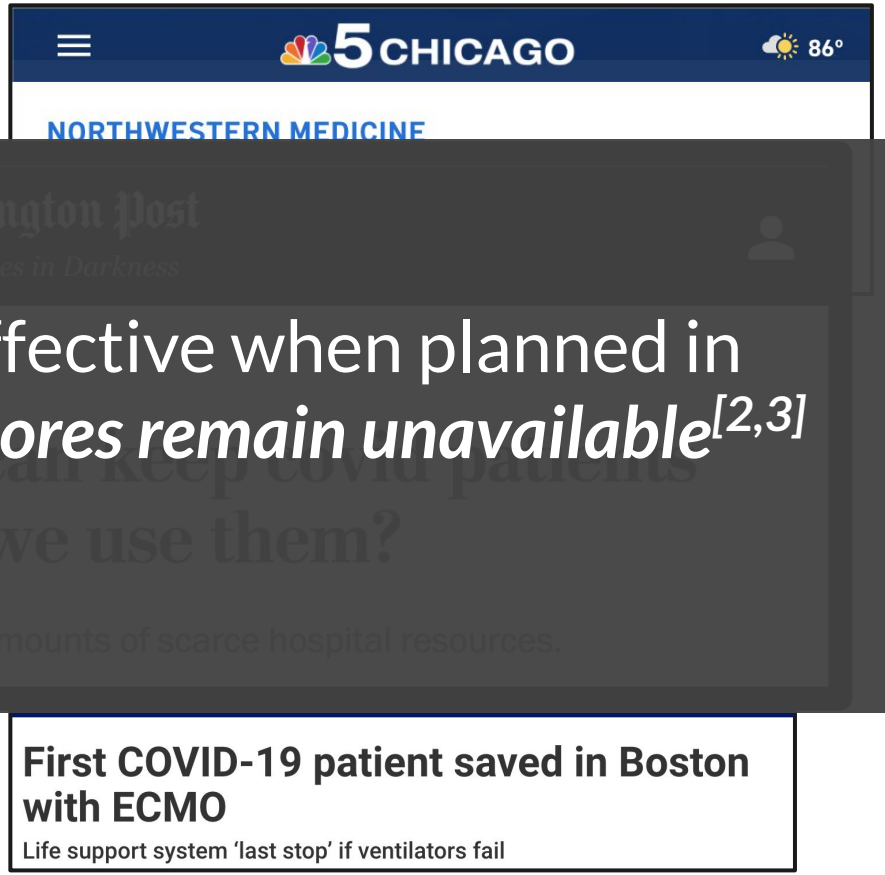

5 CHICAGO
86°
NORTHWESTERN MEDICINE

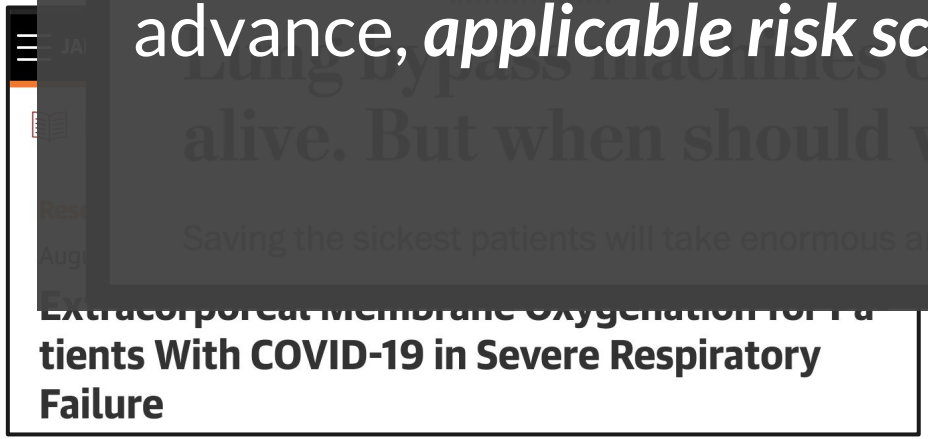
The Washington Post
Democracy Dies in Darkness
PostEverything • Perspective
Lung bypass machines can keep covid patients alive. But when should we use them?
Saving the sickest patients will take enormous amounts of scarce hospital resources.

JA
Rese
Augu
Extracorporeal Membrane Oxygenation for Pa-
tients With COVID-19 in Severe Respiratory
Failure


First COVID-19 patient saved in Boston with ECMO
Life support system 'last stop' if ventilators fail







While ECMO is more effective when planned in advance, ***applicable risk scores remain unavailable***[2,3]

Goal: to develop a risk score for patients eligible for ECMO to assist with early planning

Target

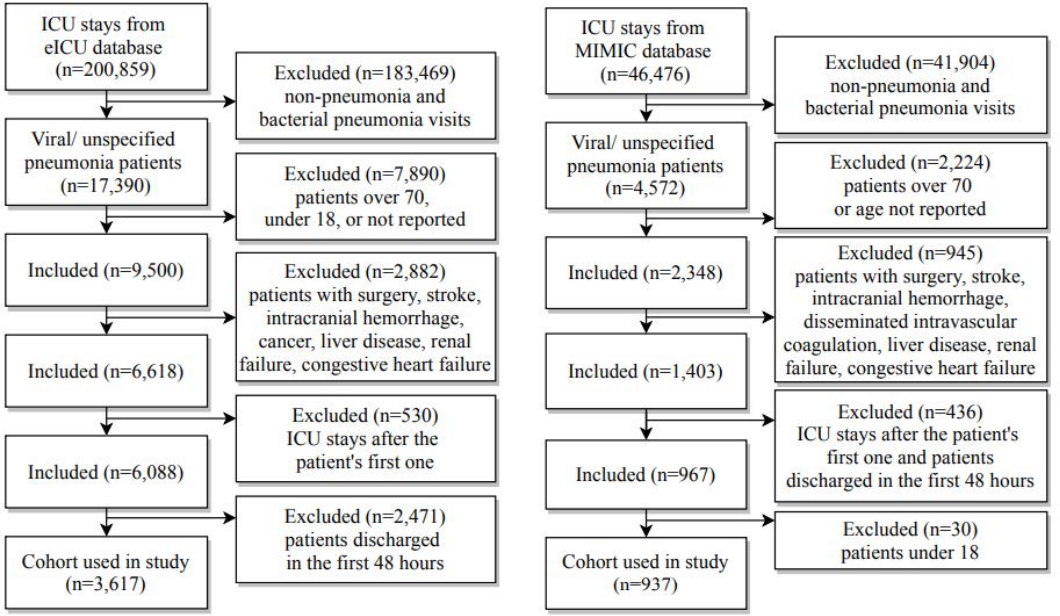
In-ICU mortality (survival analysis setting)

Secondary outcomes which indicate decompensation:

- Vasopressor use
- Ventilator use

Cohort

Critical care patients with viral or unspecified pneumonia, without contraindications for ECMO



(a) eICU cohort selection (b) MIMIC-III cohort selection

Fig. 1: Inclusion and exclusion criteria for cohorts extracted from eICU and MIMIC. Disseminated intravascular coagulation was highly missing from eICU.

Data

Two publicly available critical care databases:

- eICU Collaborative Research Database (n = 3617)
- MIMIC-III, Medical Information Mart for Intensive Care (n = 937)

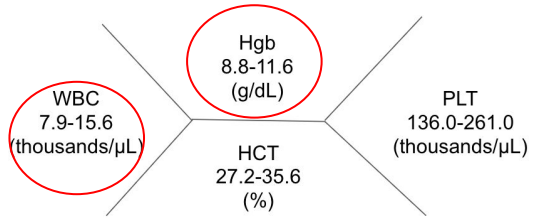
Table 1: Demographics and outcomes of patients with viral or unspecified pneumonia in eICU and MIMIC-III cohorts. Data are median (Q1-Q3) or count (% out of n).

	Variable	eICU (n = 3617)	MIMIC (n = 937)
Demographics	Age, years	58.0 (48.0-64.0)	54.5 (44.1-62.7)
	18-30	225 (6.2%)	83 (8.9%)
	30-39	277 (7.7%)	94 (10.0%)
	40-49	500 (13.8%)	159 (17.0%)
	50-59	1064 (29.4%)	281 (30.0%)
	60-70	1546 (42.7%)	320 (34.2%)
	Male	1949 (53.9%)	542 (57.8%)
	Female	1663 (46.0%)	395 (42.2%)
Out.	Deceased	270 (7.5%)	94 (10.0%)
	Vasopressors administered	589 (16.3%)	389 (41.5%)
	Ventilator used	1835 (50.7%)	758 (80.9%)

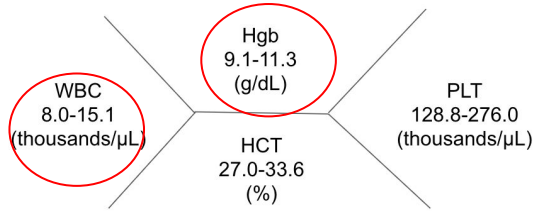
Characteristics of patients

Complete blood count

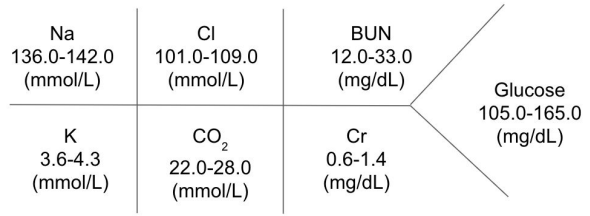
eICU

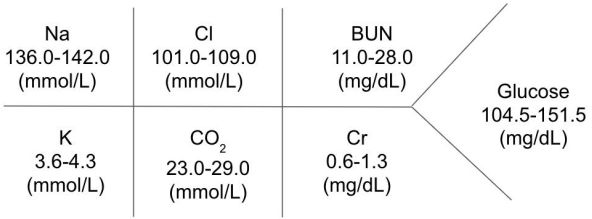


MIMIC



Chem-7





Other

SaO2 (mmHg)	Temp. (° C)	GCS
94.0-99.0	36.6-37.3	10.0-15.0
95.0-98.0	36.6-37.7	9.0-15.0

The PEER Score

The Viral or Unspecified Pneumonia ECMO-Eligible Risk Score

Target: **in-ICU mortality**, vasopressor use, ventilator use.

Data: two publicly available critical care databases:
eICU (n = 3617) and MIMIC-III (n = 937)

Cohort: critical care patients with viral or unspecified pneumonia, without contraindications for ECMO

Methods

Cox proportional hazards model:

$$\lambda(t|X_i) = \lambda_0(t)\exp(\beta_1 X_{i1} + \cdots + \beta_p X_{ip}) = \lambda_0(t)\exp(X_i \cdot \beta)$$

- L1 regularization
- 10-fold cross val. & grid search, max. concordance s.t. sufficient sparsity

Hyperparameter Selection

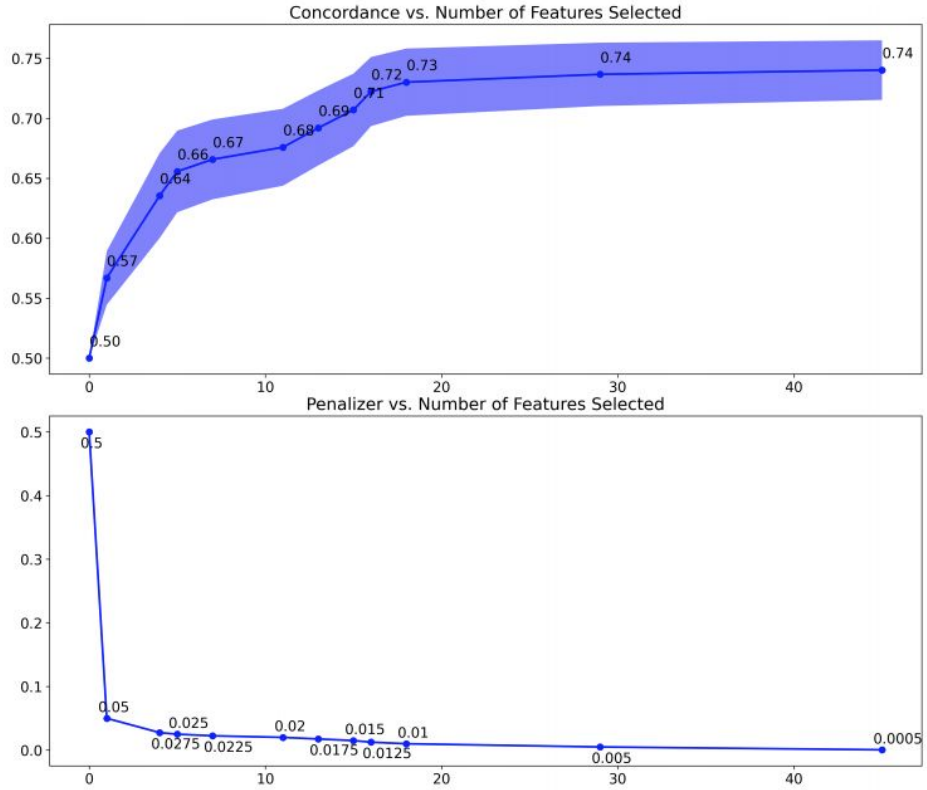


Figure F.1: Tradeoff from controlling the penalty hyperparameter λ in Lasso-Cox. As λ decreases, more features are selected and concordance increases. Beyond $\lambda = 0.01$, the gain in performance levels off.

Methods

Cox proportional hazards model:

$$\lambda(t|X_i) = \lambda_0(t)\exp(\beta_1 X_{i1} + \cdots + \beta_p X_{ip}) = \lambda_0(t)\exp(X_i \cdot \beta)$$

- L1 regularization
- 10-fold cross val. & grid search, max. concordance s.t. sufficient sparsity

Methods

Cox proportional hazards model:

$$\lambda(t|X_i) = \lambda_0(t)\exp(\beta_1 X_{i1} + \cdots + \beta_p X_{ip}) = \lambda_0(t)\exp(X_i \cdot \beta)$$

- L1 regularization
- 10-fold cross val. & grid search, max. concordance s.t. sufficient sparsity
- MissForest imputation

Methods

Cox proportional hazards model:

$$\lambda(t|X_i) = \lambda_0(t)\exp(\beta_1 X_{i1} + \cdots + \beta_p X_{ip}) = \lambda_0(t)\exp(X_i \cdot \beta)$$

- L1 regularization
- 10-fold cross val. & grid search, max. concordance s.t. sufficient sparsity
- MissForest imputation

Experimental setup:

- Created training and test set for eICU data (70-30 split)
- Entirety of the MIMIC cohort is used for model evaluation

Results: Hazard Ratios

18 nonzero out of 52 variables

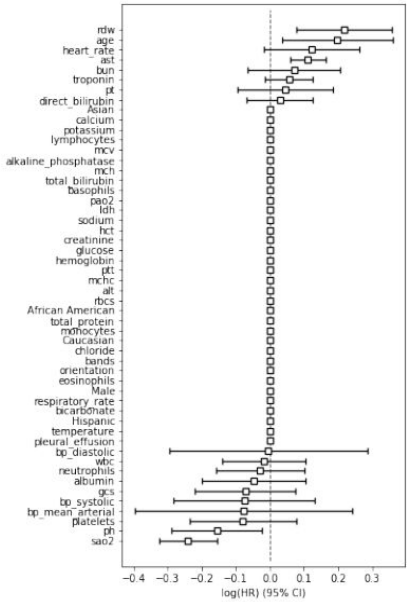


Figure E.1: Coefficients of the learned Cox model with penalty 0.01 and L1 regularization, with 95% confidence intervals (equivalent to what is reported in Table 2).

Table 2: Hazard ratios for the Lasso-Cox model, i.e. the PEER score, excluding hazard ratios equal to 1 (since they do not contribute to the model). Hazards ratios (HR) and 95% confidence intervals (CI) are reported on normalized data.

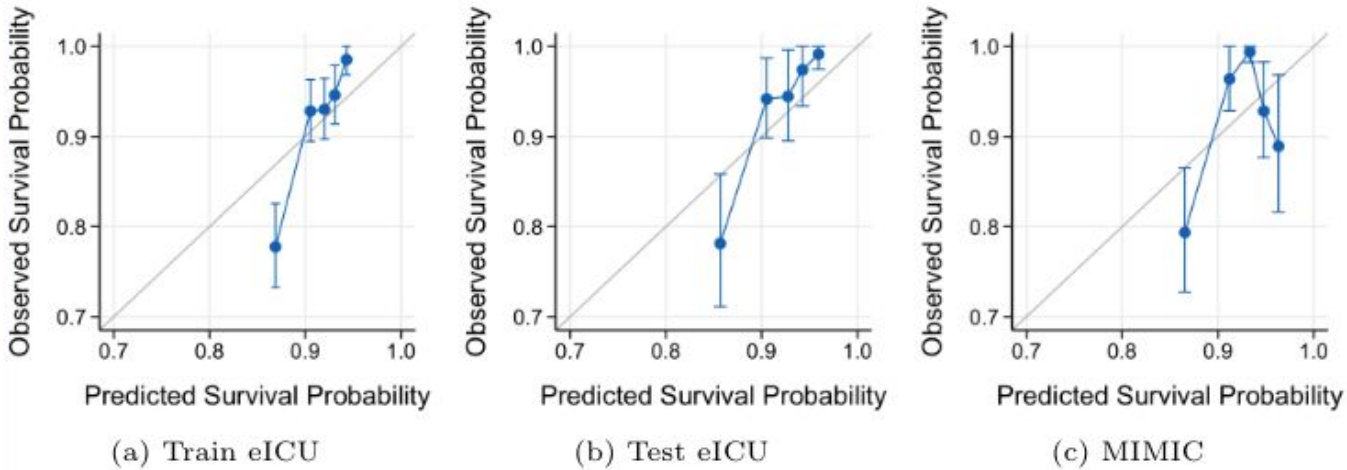
Feature	HR (95% CI)
Age (years)	1.22 (1.04 – 1.43)
Heart rate (beats per minute)	1.13 (0.984 – 1.3)
Systolic blood pressure (mmHg)	0.928 (0.755 – 1.14)
Diastolic blood pressure (mmHg)	0.996 (0.745 – 1.33)
Mean arterial pressure (mmHg)	0.926 (0.673 – 1.27)
Glasgow Coma Scale	0.93 (0.803 – 1.08)
White blood cells (thousands/μL)	0.984 (0.871 – 1.11)
Platelets (thousands/μL)	0.924 (0.79 – 1.08)
Red blood cell dist. width (%)	1.24 (1.08 – 1.43)
Neutrophils (%)	0.972 (0.853 – 1.11)
Blood urea nitrogen (mg/dL)	1.07 (0.937 – 1.23)
Aspartate aminotransferase (units/L)	1.12 (1.06 – 1.18)
Direct bilirubin (mg/L)	1.03 (0.935 – 1.13)
Albumin (g/dL)	0.954 (0.82 – 1.11)
Troponin (ng/mL)	1.06 (0.985 – 1.14)
Prothrombin time (sec)	1.05 (0.909 – 1.2)
pH	0.856 (0.75 – 0.977)
Arterial oxygen saturation (mmHg)	0.787 (0.723 – 0.856)

Results: Concordance

Score	Train eICU concordance	Test eICU concordance	MIMIC concordance
PEER (ours)	**0.77 (0.72 - 0.81)**	**0.77 (0.69 - 0.83)**	**0.66 (0.57 - 0.74)**
CURB-65 [Lim et al., 2003b]	0.66 (0.61 - 0.70)	0.62 (0.55 - 0.69)	0.59 (0.52 - 0.66)
PSI/PORT [Fine et al., 1997]	0.71 (0.66 - 0.76)	0.71 (0.63 - 0.78)	0.62 (0.55 - 0.69)
SMART-COP [Charles et al., 2008]	0.69 (0.64 - 0.73)	0.73 (0.67 - 0.80)	**0.66 (0.59 - 0.72)**
GOQ [Gong et al., 2020b]	0.67 (0.63 - 0.71)	0.62 (0.54 - 0.70)	0.58 (0.50 - 0.66)

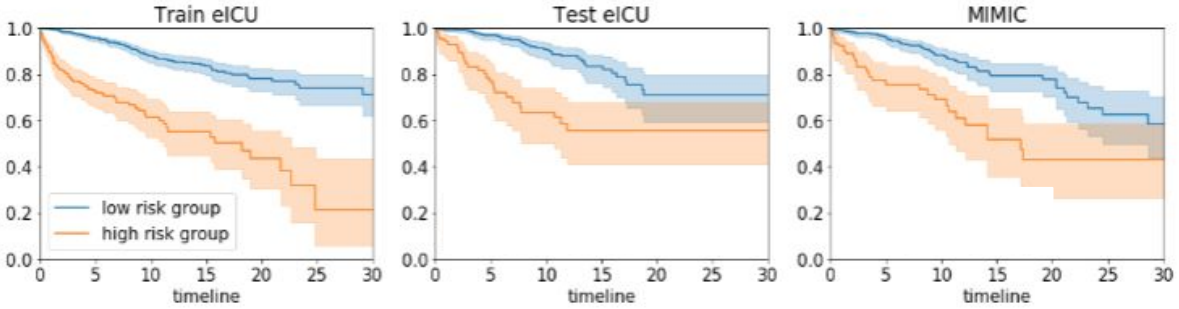
Bootstrapping with 1000 replicates was used to compute 95% confidence intervals (in parentheses)

Results: Calibration

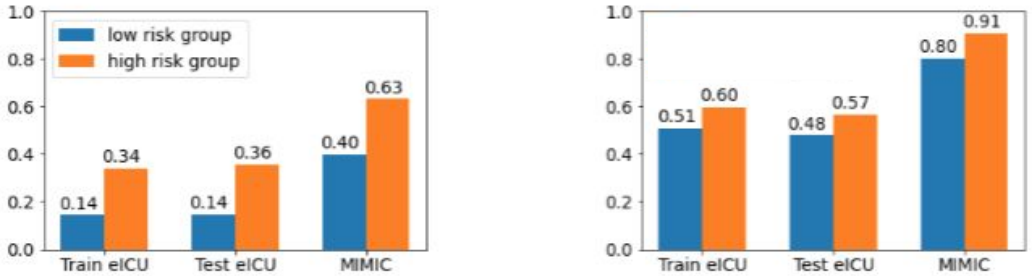


Calibration plots with 95% confidence intervals.

Results: Low and High Risk Groups



Kaplan-Meier survival curves of high vs. low risk groups in train eICU, test eICU, and MIMIC. Shaded regions are the 95% confidence intervals.



(a) vasopressor (b) ventilator

Proportion of each subgroup that received vasopressors or ventilators

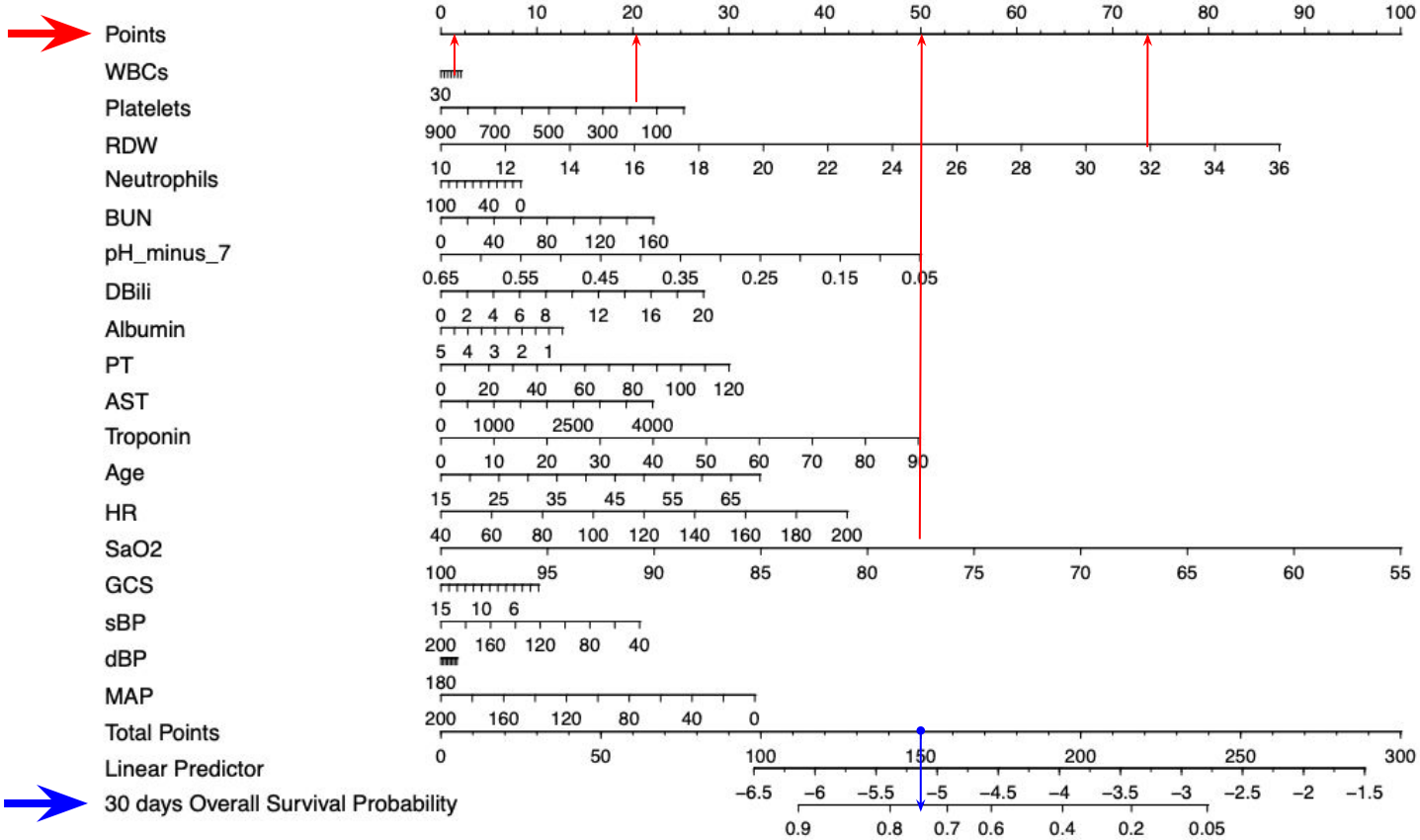


Fig. 2: Nomogram for manual calculation of the PEER score.

Discussion

- The PEER score achieves greater or comparable concordance to baselines on the eICU (in-domain) and MIMIC (out-of-domain) test sets.
- 18 features, easy to calculate.
- Stratifying each cohort into low risk and high risk subpopulations based the PEER score, there is a clear separation in their survival curves.
- However, our cohort is defined not by COVID-19 positive pneumonia patients but instead by viral or unspecified pneumonia patients who are ECMO-eligible

Future Work

- Apply our model to MIMIC-IV data
- Apply our model to COVID-19 data once we get access to it
- Making the model more accessible for actual use by clinicians (web application, additional cross-site validation, clinical trials)
- Develop or apply more flexible models (e.g. Random Survival Forest, DeepSurv)
- Handle performance degradation when tested on new dataset
    - Incorporate additional data sources
    - Methods to handle covariate shift

Thank you!

Link to paper: https://arxiv.org/abs/2006.01898

Stay in touch: {hlzhou,ccheng2}@cmu.edu

References (full version in paper):

WHO Coronavirus Disease (COVID-19) Dashboard. (n.d.). June 30, 2020.

Tarig, Fadelelmoula. "Extracorporeal membrane oxygenation therapy in adult patients: a narrative review of literature." May, 2020.

Krieger, J and Badulak, J. The Use of ECMO in Patients with Cardiopulmonary Failure Due to COVID-19.American College of Cardiology. Published online August 4, 2020.

Appendix: KM SE

Greenwood's formula for getting SE for Kaplan-Meier Survival Curve:

$$SE\left(S_t\right) = S_t\sqrt{\Sigma\frac{D_t}{N_t(N_t - D_t)}}$$

The quantity $\frac{D_t}{N_t(N_t - D_t)}$ is summed for numbers at risk (N_t) and numbers of deaths (D_t) occurring through the time of interest

Appendix: Missingness

Variable	eICU (n = 3617)	MIMIC (n = 937)
Age	0.001 (5)	0.0 (0)
Gender	0.001 (5)	0.0 (0)
Pleural effusion	0.0 (0)	0.0 (0)
Orientation	0.334 (1209)	0.48 (450)
Temperature (°C)	0.006 (20)	0.182 (171)
Heart rate (beats per minute)	0.009 (32)	0.018 (17)
Respiratory rate (breaths per minute)	0.001 (3)	0.017 (16)
Systolic blood pressure (mmHg)	0.063 (229)	0.023 (22)
Diastolic blood pressure (mmHg)	0.063 (229)	0.023 (22)
Mean arterial pressure (mmHg)	0.079 (287)	0.018 (17)
Glasgow Coma Scale	0.27 (977)	0.016 (15)
Red blood cells (millions/μL)	0.012 (44)	0.01 (9)
White blood cells (thousands/μL)	0.006 (22)	0.009 (8)
Platelets (thousands/μL)	0.014 (49)	0.01 (9)
Hematocrit (%)	0.006 (23)	0.01 (9)
Red blood cell dist. width (%)	0.057 (207)	0.012 (11)
Mean corpuscular volume (fL)	0.025 (91)	0.011 (10)
Mean corpuscular hemoglobin/ MCH (pg)	0.074 (269)	0.011 (10)
MCH concentration (g/dL)	0.025 (91)	0.01 (9)
Neutrophils (%)	0.24 (869)	0.152 (142)
Lymphocytes (%)	0.167 (603)	0.15 (141)
Monocytes (%)	0.177 (641)	0.152 (142)
Eosinophils (%)	0.209 (755)	0.152 (142)
Basophils (%)	0.256 (925)	0.152 (142)
Band cells (%)	0.752 (2720)	0.454 (425)
Sodium (mmol/L)	0.004 (13)	0.01 (9)
Potassium (mmol/L)	0.008 (29)	0.009 (8)
Chloride (mmol/L)	0.009 (33)	0.009 (8)
Bicarbonate (mmol/L)	0.057 (207)	0.009 (8)
Blood urea nitrogen (mg/dL)	0.004 (14)	0.01 (9)
Creatinine (mg/dL)	0.007 (27)	0.01 (9)
Glucose (mg/dL)	0.006 (23)	0.011 (10)
Aspartate aminotransferase (units/L)	0.174 (628)	0.218 (204)
Alanine aminotransferase (units/L)	0.177 (640)	0.219 (205)
Alkaline phosphatase (units/L)	0.184 (665)	0.223 (209)
C-reactive protein (mg/L)	0.946 (3420)	0.916 (858)
Direct bilirubin (mg/L)	0.808 (2923)	0.82 (768)
Total bilirubin (mg/L)	0.185 (670)	0.227 (213)
Total protein (g/dL)	0.184 (664)	0.841 (788)
Calcium (mg/dL)	0.021 (77)	0.027 (25)
Albumin (g/dL)	0.16 (577)	0.279 (261)
Troponin (ng/mL)	0.591 (2138)	0.505 (473)
Prothrombin time (sec)	0.396 (1431)	0.035 (33)
Partial thromboplastin time (sec)	0.545 (1973)	0.038 (36)
pH	0.244 (883)	0.104 (97)
Partial pressure of oxygen (mmHg)	0.223 (807)	0.134 (126)
Arterial oxygen saturation (mmHg)	0.015 (54)	0.726 (680)
Deceased	0.0 (0)	0.0 (0)
Vasopressors administered	0.0 (0)	0.0 (0)
Ventilator used	0.0 (0)	0.0 (0)