BMJ Health & Care Informatics

Long short-term memory model identifies ARDS and in-hospital mortality in both non-COVID-19 and COVID-19 cohort

Jen-Ting Chen ,[1,2] Rahil Mehrizi,[3] Boudewijn Aasman,[4] Michelle Ng Gong,[2] Parsa Mirhaji[5]

To cite: Chen J-T, Mehrizi R, Aasman B, *et al*. Long short-term memory model identifies ARDS and in-hospital mortality in both non-COVID-19 and COVID-19 cohort. *BMJ Health Care Inform* 2023;**30**:e100782. doi:10.1136/bmjhci-2023-100782

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjhci-2023-100782).

Received 19 April 2023
Accepted 21 August 2023

[1]Department of Medicine, Division of Pulmonary and Critical Care Medicine, UCSF, San Francisco, California, USA
[2]Department of Medicine, Division of Critical Care Medicine, Montefiore Medical Center, Bronx, New York, USA
[3]Department of Medicine, Albert Einstein College of Medicine, Bronx, New York, USA
[4]Center for Health Data Innovations, Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, New York, USA
[5]Department of Genetics, Albert Einstein College of Medicine, Bronx, New York, USA

Correspondence to
Dr Jen-Ting Chen;
tina.chen@ucsf.edu

ABSTRACT

Objective To identify the risk of acute respiratory distress syndrome (ARDS) and in-hospital mortality using long short-term memory (LSTM) framework in a mechanically ventilated (MV) non-COVID-19 cohort and a COVID-19 cohort.

Methods We included MV ICU patients between 2017 and 2018 and reviewed patient records for ARDS and death. Using active learning, we enriched this cohort with MV patients from 2016 to 2019 (MV non-COVID-19, n=3905). We collected a second validation cohort of hospitalised patients with COVID-19 in 2020 (COVID+, n=5672). We trained an LSTM model using 132 structured features on the MV non-COVID-19 training cohort and validated on the MV non-COVID-19 validation and COVID-19 cohorts.

Results Applying LSTM (model score 0.9) on the MV non-COVID-19 validation cohort had a sensitivity of 86% and specificity of 57%. The model identified the risk of ARDS 10 hours before ARDS and 9.4 days before death. The sensitivity (70%) and specificity (84%) of the model on the COVID-19 cohort are lower than MV non-COVID-19 cohort. For the COVID-19 + cohort and MV COVID-19 + patients, the model identified the risk of in-hospital mortality 2.4 days and 1.54 days before death, respectively.

Discussion Our LSTM algorithm accurately and timely identified the risk of ARDS or death in MV non-COVID-19 and COVID+ patients. By alerting the risk of ARDS or death, we can improve the implementation of evidence-based ARDS management and facilitate goals-of-care discussions in high-risk patients.

Conclusion Using the LSTM algorithm in hospitalised patients identifies the risk of ARDS or death.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Acute respiratory distress syndrome (ARDS) is commonly under-recognised in clinical settings, which can lead to delays in evidence-based management.

WHAT THIS STUDY ADDS

⇒ A long short-term memory algorithm trained on mechanically ventilated patients can identify the risk of ARDS development or in-hospital mortality using structured electronic health record data without the use of chest X-ray analysis. SARS-CoV-2 infection has a noted high incidence of ARDS. The model, trained on mechanically ventilated non-COVID-19 patients, performed well on COVID-19 patients, with an evaluation of 1.82 patients needed to identify 1 patient at risk of ARDS or death in the hospital.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Being able to identify the risk of ARDS, regardless of COVID-19 status, early can improve compliance with evidence-based management and allow prognostication.

INTRODUCTION

Acute respiratory distress syndrome (ARDS) affects nearly a quarter of all acute respiratory failure patients requiring mechanical ventilation. It contributes to high morbidity and mortality of critically ill patients.[1] ARDS is consistently under-recognised, leading to delays in implementing evidence-based best practices, such as the use of lung-protective ventilation strategies.[2 3] The onset of the COVID-19 pandemic overwhelmed the healthcare system in the USA, and patients with severe to critical SARS-CoV-2 infections had a high incidence of ARDS and high mortality. This was especially true early in the pandemic, before the discovery of using early steroids and other immunosuppressants for treatment.[4 5] An electronic health record (EHR)-based decision support system that accurately identifies patients with ARDS can improve the management and escalation of these critically ill patients.[6] Different machine learning techniques, such as L2-logistic regression, artificial neural networks and XGBoost gradient boosted tree models, have leveraged EHR to identify or predict ARDS, yielding robust statistical discrimination as reported in studies.[7–9] In a distinct study, Zeiberg *et al* applied L2-regularised logistic regression to structured EHR data

sourced from a single-centre population within the initial 7 days of hospitalisation. A meticulous two-physician chart review established the gold standard diagnosis of ARDS. Despite the rarity of ARDS occurrences (2.5%) within the testing cohort of this investigation, the area under the receiver operating curve (AUROC) attained an impressive value of 0.81.[7] Other investigations centred on using the Medical Information Mart for the ICU databases.[10 11] These endeavours relied on diverse data sources such as free-text entries, diagnostic codes and radiographic reports for both the diagnosis and prediction of ARDS.[10 11]

We aimed to train a deep learning model using long short-term memory (LSTM) framework and active learning method using a historic dataset from a mechanically ventilated (MV) non-COVID-19 cohort to identify patients with risk of ARDS or in-hospital mortality. We validated the model on an MV non-COVID-19 cohort, a COVID+ cohort and a subgroup of MV COVID+ cohort.

MATERIALS AND METHODS

The study was conducted at Montefiore Medical Center, encompassing three hospital sites.

COHORT ASSEMBLY

MV non-COVID-19 cohorts

Non-COVID-19 cohort 1 was constructed between 1 January 2017 and 31 August 2018 (figure 1). We included MV adults in the ICU with ages greater than 18. Each patient's chart was reviewed for ARDS.

Ground truth labelling: ARDS gold-standard identification

We defined ARDS using the Berlin criteria: hypoxaemia (arterial oxygen tension (PaO2) to fractional inspired oxygen (FiO2) ratio (PFR)≤300 with positive pressure ventilation ≥5cmH_20), bilateral infiltrates on chest radiographs by independent review and a presence of ARDS risk factors (sepsis, shock, pancreatitis, aspiration, pneumonia, drug overdose and trauma/burn) not solely due to heart failure.[12] We used the first date and time of PFR≤300 with confirmed bilateral infiltrates within 24 hours as the time of ARDS presentation (ToP of ARDS).

Active learning

We used the 'active learning' technique to provide additional adult MV patients from July 2016 to December 2016 and September 2018 to December 2019 (AL-cohort).[13] A preliminary recurrent neural network was developed using the LSTM model and trained with the original non-COVID-19 cohort 1. Next, we applied the preliminary model to the AL-cohort. We used pool-based sampling and uncertainty techniques to identify records from AL-cohort to be reviewed and labelled by clinicians.[13] The uncertainty technique includes patients whose scores are very close to the cut-off, which means the model is least confident about them. We chose a cut-off of 0.80 and selected all records with a score between 0.75 and 0.85. We created the MV non-COVID-19 cohort 2 using the top 1% of the highest, lowest 1% and medium scores of the AL cohort. This allowed us to enrich MV non-COVID-19 cohort 2 with patients with ARDS or those who died in the hospital.

COVID-19 validation cohort

We included all hospitalised adult patients with and without mechanical ventilation with a positive SARS-Cov-2 transcription-mediated amplification assay from 1 March 2020 to 17 April 2020 in the COVID-19 cohort.

Training and validation cohort splitting

MV non-COVID-19 cohorts 1 and 2 were combined as the MV non-COVID-19 cohort. We randomly selected 80% of patients for training (MV non-COVID training cohort) and validation to learn model parameters and find optimal hyperparameters. The trained model was validated on the remaining 20% of the non-COVID-19 cohort (MV non-COVID-19 validation cohort), the COVID-19 cohort and the MV COVID-19 cohort separately (figure 1).

EHR DATA COLLECTION AND PROCESSING

Clinical data were collected through automated abstraction of EHR data. Raw EHR data for each admission were abstracted, sampled and validated (online supplemental table 2).

Sampling

Raw longitudinal EHR data were sampled every hour. Sampling was necessary since the different variables were recorded at different timestamps with different frequencies to aggregate the longitudinal data into hourly snapshots. If the data were recorded multiple times within 1 hour, we computed the minimum and maximum based on all recorded measurements. If it was not recorded at all within the 1-hour time frame, we considered it as 'missing'. For data that were recorded exactly once during an hour, the minimum and maximum would be the same.

Data validation

Data validation was performed by range checking (online supplemental table 2). If the recorded measure was outside the valid range, we discarded it and treated it as a missing value.

Missing data

The missing data were handled by 'forward imputing', where the most recent value fills the missing value. If there were no data available for imputation, we used normal values. We used the lower bound of the normal range as the minimum and the upper bound as the maximum value for those timestamps. A feature vector of size 132 represents each timestamp.

MODEL TRAINING

LSTM network is a paradigm of recurrent neural networks that can capture the temporal information

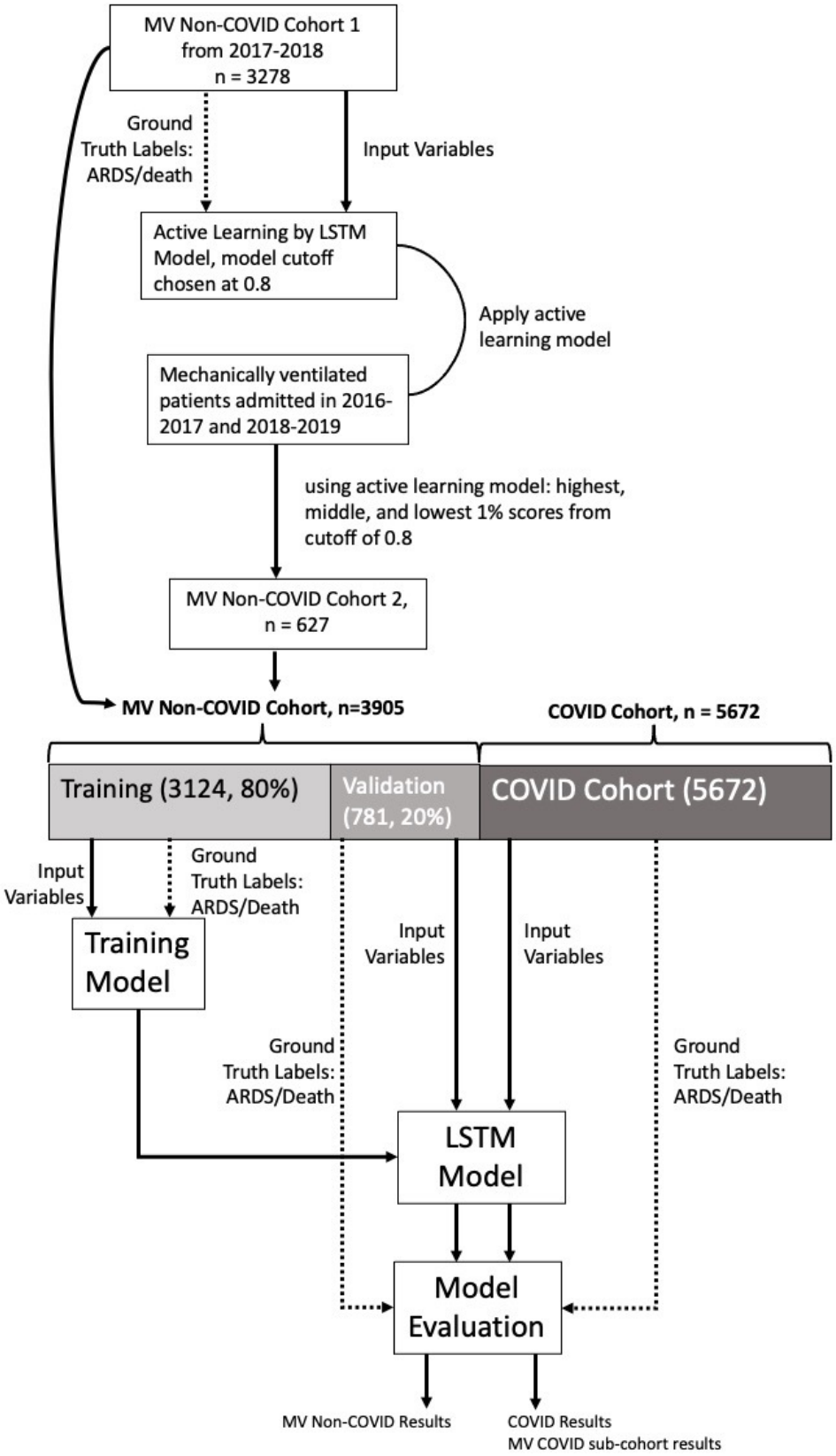


Figure 1 Cohort assembly and model training. ARDS, acute respiratory distress syndrome; LSTM, long short-term memory; MV, mechanically ventilated.

of sequential data.[14] We used the EHR data, including the previous 12 hours, as the network inputs to train a model that can generate a predictive score for every patient at every hour. The network consisted of an LSTM unit with 10 filters, followed by a drop-out layer with 50% probability of keeping.[15] The network ended

Table 1 Cohorts characteristics

Variables	MV non-COVID-19 cohort	Training	Validation		
		MV non-COVID-19 (training) cohort	Non-COVID-19 (validation) cohort	COVID-19 cohort	MV COVID-19 subcohort
n	3905	3124	781	5672	803
Age, year, mean±SD	65.0±14.7	65.0±14.8	65.3±14.4	60.80±17.2	62.1±13.9
Gender					
Male, n (%)	1741 (44.6)	1437 (46)	328 (42)	2665 (47)	319 (40)
Female, n (%)	2164 (55.4)	1686 (54)	452 (58)	3006 (53)	484 (60)
Race or ethnicity					
White, n (%)	1015 (26)	749 (24)	249 (32)	623 (11)	177 (22)
Black, n (%)	1718 (44)	1405 (45)	320 (41)	2495 (44)	345 (43)
Other, n (%)	1171 (30)	968 (31)	210 (27)	2552 (45)	281 (35)
ARDS determination					
PaO2/FiO2 ratio ≤300, n (%)	3211 (82.2)	2579 (82.6)	632 (80.9)	617 (10.9)	617 (77)
CXR interpretation					
Yes (consistent with ARDS), n (%)	1333 (34.1)	35.4 (35.4)	260 (33.3)	565 (10)	565 (82)
Indeterminant, n (%)	313 (8.0)	7.1 (7.1)	60 (7.7)	18 (.3)	18 (2.2)
No (not consistent with ARDS), n (%)	2259 (57.8)	57.6 (57.6)	461 (59)	34 (.6)	34 (4.2)
Risk factors					
Aspiration, n (%)	407 (10.4)	10.3 (10.3)	86 (11)		
Shock, n (%)	1520 (38.9)	39.2 (39.2)	299 (38.3)		
Pneumonia, n (%)	1530 (39.2)	39.8 (39.8)	288 (36.9)	5672 (100)	803 (100)
Sepsis, n (%)	1885 (48.3)	48.8 (48.8)	362 (46.4)		
Pancreatitis, n (%)	42 (1.1)	1.1 (1.1)	9 (1.2)		
Burn, n (%)	3 (0.1)	3 (0.1)	0 (0)		
Overdose, n (%)	98 (2.5)	2.5 (2.5)	21 (2.7)		
Transfusion, n (%)	1191 (30.5)	30.7 (30.7)	232 (29.7)		
Congestive heart failure, n (%)	914 (23.4)	23.6 (23.6)	178 (22.8)		
Presence of at least one risk factor, n (%)	2739 (70.1)	70.6 (70.6)	362 (46.4)	5672 (100)	803 (100)
Clinical outcomes					
Mechanically ventilated, n (%)	3905	3124	781	803 (14.2)	803
ARDS, n (%)	1646 (42.2)	1326 (42.4)	320 (41.0)	583 (10.3)	583 (72.6)
In-hospital mortality, n (%)	1033 (26.5)	848 (27.1)	185 (23.7)	907 (16.0)	418 (52.1)
ARDS or in-hospital mortality, n (%)	2044 (52.3)	1655 (53.0)	389 (49.8)	1235 (21.9)	746 (92.9)

ARDS, acute respiratory distress syndrome; CXR, chest X-ray; FiO2, fractional inspired oxygen; MV, mechanically ventilated; PaO2, arterial oxygen tension.

with a linear layer and a Sigmoid activation function to output a score from 0 to 1, which is interpreted as the probability of developing ARDS or in-hospital mortality.

MODEL EVALUATION

We applied the model on the MV non-COVID-19 validation cohort and COVID-19 cohort hourly to produce the score for that timestamp which is an indication of the probability of ARDS development or death. For each cohort, we calculated the AUROC. We also calculated the sensitivity, specificity, positive predictive value (PPV), negative predictive value, and F1 score at different risk thresholds (cutoffs). We use the highest F1 score to generate a confusion matrix for selecting a score cut-off.

Table 2 Model diagnostics

TREAT-ECARDS model diagnostics	MV non-COVID-19 cohort	COVID-19 cohort	MV COVID-19 subcohort
Sensitivity	0.86	0.7	0.92
Specificity	0.57	0.84	0.23
Positive predictive value	0.66	0.55	0.94
Negative predictive value	0.8	0.91	0.17
Receiver operating curve	0.78	0.83	0.7
F1 score	0.75	0.61	0.93
No needed to evaluate	1.52	1.82	1.06

MV, mechanically ventilated.

The warning time is the first time the score exceeds the predefined cut-off. We continued running the test until the score exceeded the cut-off or discharge time. We evaluated model timeliness based on ARDS and death, ARDS and not death, no ARDS and death, no ARDS and not death and compared the actual ToP ARDS time/death time with the warning time.

FEATURE IMPORTANCE

Feature importance identifies a subset of features that are the most relevant for the accuracy of the model. We used local interpretable model-agnostic explanations (LIME),[16] to determine the importance of each variable to the accuracy of the model. The feature importance value was determined for 200 randomly sampled patients in each cohort using LIME, then calculated the average across all samples.

RESULTS

Cohort description

MV non-COVID-19 cohort 1 included 3278 patients (online supplemental table 1 and figure 1). MV Non-COVID-19 cohort 2 was derived from the active learning, consisting of 627 patients (online supplemental table 1). We combined MV Non-COVID-19 cohorts 1 and 2 to create the MV non-COVID-19 Cohort (n=3905, table 1). COVID-19 cohort included 5672 patients (table 1). Online supplemental table 3 shows the descriptive statistics of all variable fields in the MV non-COVID and COVID-19 cohorts.

MODEL DIAGNOSTICS

MV non-COVID-19 validation cohort

Based on the highest F1 score, we chose a model score cut-off at 0.90. The model diagnostics are presented in table 2, figure 2. The model warned of patient risk at a median of 10 hours (IQR −75 to 4) before ARDS and −225 hours or 9 days (IQR −461 to 101 hours) before death in the hospital (table 3). In ARDS survivors, the majority of the patients had ARDS risk identified before intubation and before ARDS diagnosis (table 3). For ARDS non-survivors, the model warned at 1 hour (IQR −38 to 9) before intubation, −20 hours (IQR −115 to 0.3) before ARDS and at −314 hours (IQR −589 to −128 hours) before death (table 3).

COVID-19 cohort and MV COVID-19 subcohort

Using the same cut-off of 0.9, we applied the model to COVID-19 and MV COVID-19 subcohorts. The model diagnostics are presented in table 2 and figure 2. When the model was applied to the COVID-19 cohort, the PPV was lower and more patients needed to be screened compared with the MV non-COVID-19 validation cohort. Whereas in the MV COVID -19 subcohort patients had a high prevalence of ARDS and in-hospital mortality, the PPV and number needed to evaluate were much lower than in the MV non-COVID-19 Validation Cohort.

In the COVID-19 cohort, the model warned the patient was likely to have ARDS or in-hospital mortality 3 hours after intubation and at ToP ARDS (table 3). Among the non-survivors, the model warned 2.4 days before in-hospital mortality (IQR 4.7–0.83) in COVID-19 patients, and 1.54 days before in-hospital mortality (IQR 3.6–0.46) in MV COVID-19 patients (table 3).

FEATURE IMPORTANCE

For both the MV non-COVID-19 and COVID-19 cohorts, we randomly selected 200 encounters from each cohort and performed LIME (online supplemental figure 1). The top contributors are similar in the MV non-COVID-19 and COVID-19 cohorts. The most important variable to the model was lactate level in discriminating the clinical outcome. The model consistently used lactate, age, cryoprecipitate transfusion, dopamine, bicarbonate level and epinephrine as important input variables (online supplemental figure 1).

DISCUSSION

From a cohort of pre-COVID-19 pandemic patients on mechanical ventilation, we developed and validated an LSTM model to identify patients at risk for ARDS or in-hospital mortality. This model was successfully integrated into EHR and identified patients at risk for ARDS

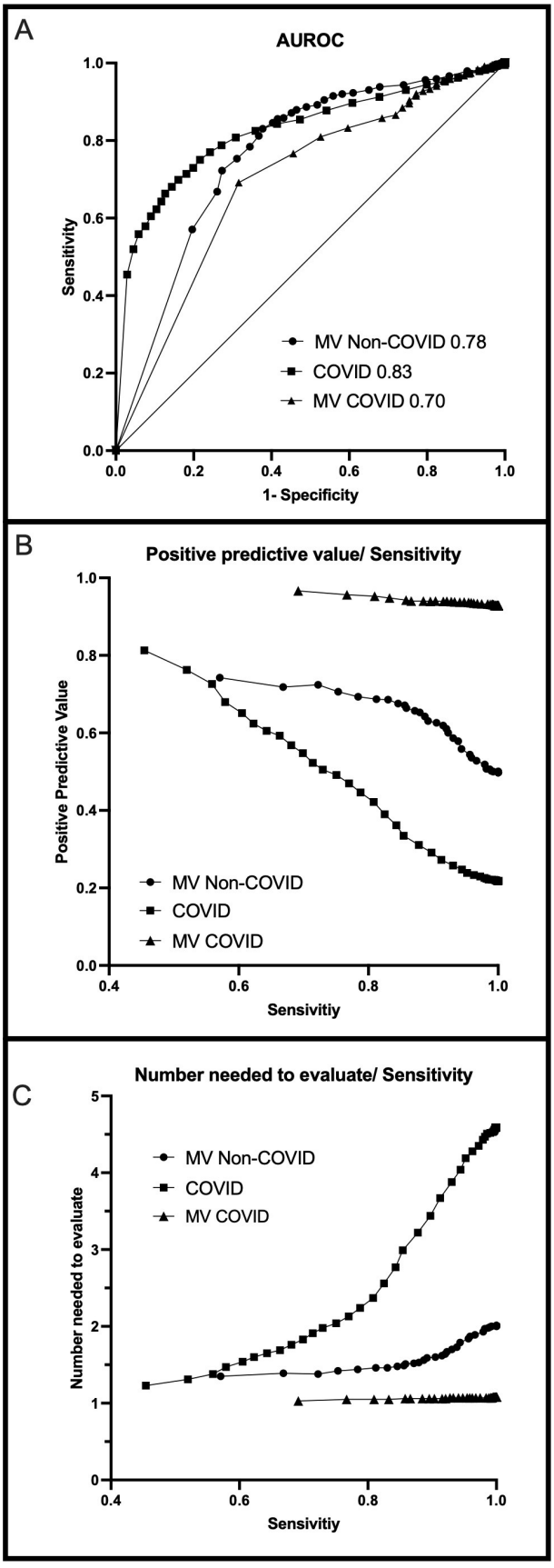


Figure 2 Model diagnostics, AUROC, PPV with sensitivity and NNE with sensitivity. AUROC, area under the receiver operating curve; MV, mechanically ventilated; NNE, number needed to evaluate; PPV, positive predictive value.

or in-hospital mortality in all adults hospitalised with and without COVID-19 infection, regardless of mechanical ventilation status. The model was also able to warn well before the events of ARDS or death in both the MV non-COVID-19 and COVID-19 cohorts. The timeliness of the model allows clinicians to modify management and implement evidence-based practices promptly.

This is the first utilisation of an LSTM network for identifying the risk of ARDS and in-hospital mortality. The LSTM is a recurrent neural network that uses feedback layers to capture temporal aspects such as sequences and trends. This approach is well suited for this study because past events and the progression of patient status are often valuable to determine the probability of ARDS or death. As in the reality of managing critically ill patients, physiological observations at each time point are taken into account. Their change and progression or regression inform the decisions at the subsequent processing of this information. This is well suited for dynamically changing situations to monitor and identify patients progressing to ARDS or in-hospital mortality. LSTM models have been used to predict heart failure, transfusion needs in the ICU, and mortality in the neonatal ICU, all with better predictive utility than traditional logistic regression models.[17–19] We chose to include ARDS diagnosis and in-hospital mortality as our patient-centred outcomes of interest instead of ARDS or in-hospital mortality alone, as in previous ARDS prediction studies.[6 7 20] Identifying the risk of ARDS or in-hospital mortality has shown real clinical implications when managing patients, mitigating the ambiguity that sometimes can exist in ARDS clinical diagnosis based on shifting diagnostic criteria.[7 8 20–22]

This cohort is one of the largest validated ARDS gold standards developed by manual chart review and active learning from a single centre. We did not rely on ICD-10 diagnosis codes or radiology reports to identify ARDS. Instead, we followed the Berlin criteria using PFR, independent review of chest X-ray for the presence of bilateral infiltrates and risk factors of ARDS in the patients' chart. Our model performed similarly to previously reported models using other machine learning methods, ranging from 0.71 to 0.90.[7 9–11 21] We forgo chest X-ray interpretation as input variables, as in Zeiberg *et al*.[7] Other large-scale ARDS identification studies which used natural language processing of radiology reports and diagnostic codes in clinical settings would delay ARDS recognition and rely heavily on clinician decisions.[9 11] Using chest radiographs for the diagnosis of ARDS has its limitations, as studies show high interobserver variabilities despite training.[12 23] In addition, radiology report turn-around times can range from 15 min to 26 hours, depending on the study location, availability of staff and hospital resources.[24 25] This reliance on chest radiograph interpretations may delay ARDS diagnosis.

Despite the different clinical characteristics of the study cohorts, being MV patients non-COVID-19 versus non-MV COVID-19 patients, important features in risk identification were broadly consistent between the cohorts using

Table 3 Timeliness of model

Cohort, (n)	Correctly identifies, n (%)	Time from intubation, median (IQR), hours	Before intubation, n (%)	After intubation, n (%)	Time from ARDS label, median (IQR), hours	Before ARDS, n (%)	After ARDS, n (%)	Time from death median (IQR), hours
MV non-COVID-19 cohort								
ARDS, (204)	166 (81.4)	0 (–12.8 to 26.0)	87 (52.4)	79 (47.6)	0 (–43.8 to 12.0)	115 (69.3)	51 (30.7)	
Death, (69)	60 (87)							225.5 (–461.3 to –101.3)
ARDS and death, (116)	108 (93.1)	-1 (–38.8 to 9.3)	68 (63.0)	40 (37.0)	20 (–115.5 to 0.3)	81 (75.0)	27 (25.0)	314 (–588.5 to –127.8)
ARDS or death, (389)	274 (70.4)	1 (–17.8 to 15.0)	155 (56.6)	119 (43.4)	10.0 (–75.5 to 4.0)	196 (71.5)	78 (28.5)	225.5 (–461.3 to –101.3)
No ARDS or death, (392)	223 (56.9)							
COVID-19 cohort								
ARDS, (328)	318 (97)	3 (–8.8 to 11.0)	136 (42.8)	182 (57.2)	0 (–16 to 9.0)	141 (44.3)	128 (40.3)	
Death, (652)	308 (47.2)							58 (–112 to –20)
ARDS and death, (255)	237 (92.9)	4 (–1 to 18)	86 (36.3)	156 (65.8)	0 (–12 to 10)	125 (52.7)	112 (47.3)	112 (–211.3 to –52)
ARDS or death, (1235)	555 (44.9)	3 (–3.5 to 13.0)	222 (40.0)	333 (60.0)	0 (–14.0, 10.0)	266 (47.9)	240 (43.2)	58 (–112 to –20)
No ARDS or death, (4437)	3724 (83.9)							
MV COVID-19 subcohort								
ARDS, (328)	318 (97)	3 (–8.8 to 11.0)	136 (42.8)	182 (57.2)	0 (–16.0 to 9.0)	141 (44.3)	128 (40.3)	
Death, (163)	128 (78.5)							37 (–87 to –11)
ARDS and death, (255)	237 (92.9)	4 (–1, 18)	86 (36.3)	156 (65.8)	0 (–12 to 10)	125 (52.7)	112 (47.3)	112 (–211. 3 to –52)
ARDS or death, (746)	555 (74.4)	3 (–3.5 to 13.0)	222 (40.0)	333 (60.0)	0 (–14.0 to 10.0)	266 (47.9)	240 (43.2)	37 (–87 to –11)
No ARDS or death, (57)	13 (22.8)							

ARDS, acute respiratory distress syndrome; MV, mechanically ventilated.

lactate, age, cryoprecipitate transfusion, dopamine, bicarbonate level and epinephrine as important input variables. LIME can directly associate model features to increased or decreased risk of ARDS or death in an individual, on a patient-by-patient-level.[26 27] We randomly sampled 200 patients in each cohort and obtained an average of the absolute LIME values to understand what features were generally used. This does not provide a clinical explanation and rationale for why features may relate to higher or lower scores. Instead, it sheds light on important features that the model needs as its input data to predict a score accurately, whether additive or subtractive, to the risk. Norepinephrine was the most commonly used vasopressor for both cohorts; intriguingly, it did not contribute to the model consideration. The model rarely used vasopressors such as dopamine and epinephrine to discriminate the outcome of ARDS and/or in-hospital mortality. Oxygen support devices were also not deemed important on average; we postulate that our gold standard labelling required mechanical ventilation for ARDS identification, making oxygen support devices less important in the discrimination.

In clinical practice, ARDS is underdiagnosed, which leads to increased exposures in management that are detrimental to patients, such as high tidal volume ventilation and delayed implementation of evidence-based practices that are helpful.[2 3 28–31] We used continuous data at 1-hour intervals starting at hospital admission to identify the early risk of an adverse outcome. Indeed, in the non-COVID-19 cohort, we identified ARDS hours before intubation and at the time of ToP ARDS. The majority of patients (56.5%) had been identified before ARDS diagnosis in the MV non-COVID-19 cohort, and this remained the case in the COVID+ cohort (43%). Implemented and delivered as a clinical decision support system, the early recognition would allow clinicians to initiate treatment such as LTVV as early as possible, when it may more positively impact outcomes.[3]

Furthermore, the model identified the risk of in-hospital mortality 9 days in advance in the non-COVID-19 cohort and 2 days in advance in the COVID-19 cohort. This has significant implications for triaging patients during surge capacity. In the MV non-COVID-19 cohort, there was no concern for ventilator or ICU resource allocation. Early identification of risk for death would alert the clinician to implement aggressive management and allow the treating physician to consider early palliation intervention/conversation. In the setting of a high volume surge of respiratory illness, such as the onset of the COVID-19 pandemic, where the incidences of ARDS and death are high, identifying adverse outcomes days in advance could help the clinician in making necessary triage decisions for resource allocation.[32–34]

Our study has some limitations. First, our cohorts were constructed from a single centre in the Bronx, and the patients' characteristics may not be generalisable to other centres and populations. However, our medical centre consists of three hospitals ranging from community and academic to tertiary transplant centres, thus spanning a wide spectrum of disease severity. In addition, we validated the algorithm in the COVID-19 cohort regardless of the respiratory support type, demonstrating consistent model performance across different cohorts. Second, although we were able to determine feature importance using LIME on 200 samples from each cohort, we were unable to discern the actual direction of association with the risk of ARDS or death. We cannot discern if the individual variables increase or decrease the risk of ARDS or death, despite their importance to the overall model. However, the consistency in features used to determine risk between the validation cohorts is reassuring. Ultimately, the variables that we included in models are variables known to be clinically associated with ARDS or death; therefore, the direction of influence on risk assessment is less germane. The strength of our study lies in the predictive nature of this algorithm and the timeliness of its predictions. Using longitudinal data from admission allowed the LSTM model to learn from the progression of the patient's clinical status over time. This model also was flexible to have similar diagnostic performance in patients with different clinical characteristics.

In conclusion, our LSTM model identified risk for ARDS and in-hospital mortality on patients with or without COVID-19 regardless of mechanical ventilator support. The model identified patients early, which implies management changes can be implemented early.

Acknowledgements The authors thank Mr. Daniel Ceusters for his administrative assistance for this project.

Contributors J-TC: data collection and monitoring, data analysis and manuscript preparation. RM: data analysis and manuscript preparation. BBA: data collection and data analysis. MNG: idea and project generation, data analysis, manuscript review. PM: idea and project generation, data analysis, manuscript review. PM is the guarantor of the overall content.

Funding This study was supported by Agency for Healthcare Research and Quality (AHRQ), grant number R18HS026188 and National Insitute of Health/ National Center for Advancing Translational Science (NCATS) Einstein-Montefiore CTSA Grant Number 1UM1TR004400.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study was approved by Albert Einstein College of Medicine Internal Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study. All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iD

Jen-Ting Chen http://orcid.org/0000-0002-8530-2520

REFERENCES

1 Cartin-Ceba R, Kojicic M, Li G, *et al*. Epidemiology of critical care syndromes, organ failures, and life-support interventions in a suburban US community. *Chest* 2011;140:1447–55.
2 Bellani G, Laffey JG, Pham T, *et al*. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. *JAMA* 2016;315:788–800.
3 Needham DM, Yang T, Dinglas VD, *et al*. Timing of low tidal volume ventilation and intensive care unit mortality in acute respiratory distress syndrome. A prospective cohort study. *Am J Respir Crit Care Med* 2015;191:177–85.
4 Berlin DA, Gulick RM, Martinez FJ. Severe COVID-19. *N Engl J Med* 2020;383:2451–60.
5 Richardson S, Hirsch JS, Narasimhan M, *et al*. Presenting characteristics, Comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *JAMA* 2020;323:2052–9.
6 Wayne MT, Valley TS, Cooke CR, *et al*. "Electronic "Sniffer" systems to identify the acute respiratory distress syndrome". *Ann Am Thorac Soc* 2019;16:488–95.
7 Zeiberg D, Prahlad T, Nallamothu BK, *et al*. Machine learning for patient risk stratification for acute respiratory distress syndrome. *PLoS One* 2019;14:e0214465.
8 Wong A-KI, Cheung PC, Kamaleswaran R, *et al*. Machine learning methods to predict acute respiratory failure and acute respiratory distress syndrome. *Front Big Data* 2020;3:579774.
9 Le S, Pellegrini E, Green-Saxena A, *et al*. Supervised machine learning for the early prediction of acute respiratory distress syndrome (ARDS). *J Crit Care* 2020;60:96–102.
10 Taoum A, Mourad-Chehade F, Amoud H. Early-warning of ARDS using novelty detection and data fusion. *Comput Biol Med* 2018;102:191–9.
11 Apostolova E, Uppal A, Galarraga JE, *et al*. Towards reliable ARDS clinical decision support: ARDS patient analytics with free-text and structured EMR data. *AMIA Annu Symp Proc* 2019;2019:228–37.
12 Sjoding MW, Hofer TP, Co I, *et al*. Interobserver reliability of the Berlin ARDS definition and strategies to improve the reliability of ARDS diagnosis. *Chest* 2018;153:361–7.
13 Settles B. Active learning literature survey. Computer Sciences Technical Report 1648. University of Wisconsin–Madison; 2009.
14 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput* 1997;9:1735–80.
15 Srivastava N, Hinton G, Krizhevsky A, *et al*. Dropout: A simple way to prevent neural networks from Overfitting. *J Mach Learn Res* 2014;15:1929–58.
16 Ribeiro MT, Singh S, Guestrin C. Why should I trust you?": explaining the predictions of any Classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; San Francisco, California, USA: Association for Computing Machinery, 2016:1135–44
17 Shung D, Huang J, Castro E, *et al*. Neural network predicts need for red blood cell transfusion for patients with acute gastrointestinal bleeding admitted to the intensive care unit. *Sci Rep* 2021;11:8827.
18 Hagan R, Gillan CJ, Spence I, *et al*. Comparing regression and neural network techniques for personalized predictive Analytics to promote lung protective ventilation in intensive care units. *Comput Biol Med* 2020;126:104030.
19 Maheshwari S, Agarwal A, Shukla A, *et al*. A comprehensive evaluation for the prediction of mortality in intensive care units with LSTM networks: patients with cardiovascular disease. *Biomed Tech (Berl)* 2020;65:435–46.
20 Ding X-F, Li J-B, Liang H-Y, *et al*. Predictive model for acute respiratory distress syndrome events in ICU patients in China using machine learning Algorithms: a secondary analysis of a cohort study. *J Transl Med* 2019;17:326.
21 Fei Y, Gao K, Li WQ, *et al*. Prediction and evaluation of the severity of acute respiratory distress syndrome following severe acute Pancreatitis using an artificial neural network algorithm model. *HPB (Oxford)* 2019;21:891–7.
22 Sinha P, Delucchi KL, McAuley DF, *et al*. Development and validation of parsimonious Algorithms to classify acute respiratory distress syndrome phenotypes: a secondary analysis of randomised controlled trials. *Lancet Respir Med* 2020;8:247–57.
23 Goddard SL, Rubenfeld GD, Manoharan V, *et al*. The randomized educational acute respiratory distress syndrome diagnosis study: A trial to improve the radiographic diagnosis of acute respiratory distress syndrome. *Crit Care Med* 2018;46:743–8.
24 Towbin AJ, Iyer SB, Brown J, *et al*. Practice policy and quality initiatives: decreasing variability in turnaround time for radiographic studies from the emergency Department. *Radiographics* 2013;33:361–71.
25 Chan KT, Carroll T, Linnau KF, *et al*. Expectations among academic Clinicians of inpatient imaging turnaround time: does it correlate with satisfaction *Acad Radiol* 2015;22:1449–56.
26 Ribeiro MT, Singh S, Guestrin C. Why should I trust you?": explaining the predictions of any Classifier. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16); New York, NY, USA: Association for Computing Machinery, 2016
27 Elshawi R, Al-Mallah MH, Sakr S. On the Interpretability of machine learning-based model for predicting hypertension. *BMC Med Inform Decis Mak* 2019;19:146.
28 Weiss CH, Baker DW, Weiner S, *et al*. Low tidal volume ventilation use in acute respiratory distress syndrome. *Crit Care Med* 2016;44:1515–22.
29 Qadir N, Bartz RR, Cooter ML, *et al*. Variation in early management practices in moderate-to-severe ARDS in the United States: the severe ARDS: generating evidence study. *Chest* 2021;160:1304–15.
30 Duggal A, Rezoagli E, Pham T, *et al*. Patterns of use of Adjunctive therapies in patients with early moderate to severe ARDS: insights from the LUNG SAFE study. *Chest* 2020;157:1497–505.
31 Brower RG, Matthay MA, Morris A, *et al*. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med* 2000;342:1301–8.
32 Laventhal N, Basak R, Dell ML, *et al*. The ethics of creating a resource allocation strategy during the COVID-19 pandemic. *Pediatrics* 2020;146:e20201243.
33 Emanuel EJ, Persad G, Upshur R, *et al*. Fair allocation of scarce medical resources in the time of COVID-19. *N Engl J Med* 2020;382:2049–55.
34 Aliberti MJR, Szlejf C, Avelino-Silva VI, *et al*. COVID-19 is not over and age is not enough: using frailty for prognostication in hospitalized patients. *J Am Geriatr Soc* 2021;69:1116–27. 10.1111/jgs.17146 Available: https://onlinelibrary.wiley.com/toc/15325415/69/5

Supplemental Table 1

Variables	**Non-COVID Cohort 1 (2017-2018)**	**Non-COVID Cohort 2 (2016-2017 and 2018-2019)**
n	3278	627
Age, yr, mean +/-SD	65.84 ± 14.6	63.93 ± 14.9
Gender		
Male, n (%)	1453 (44.3)	288 (45.9)
Female, n (%)	1825 (55.7)	339 (54.1)
Race or ethnicity		
White, n (%)	820 (25)	144 (23)
Black, n (%)	1410 (43)	288 (46)
Other, n(%)	1048 (32)	195 (31)
ARDS Determination		
PaO2/FiO2 Ratio < = 300, n (%)	2586 (78.9)	625 (78.9)
CXR interpretation		
Yes (consistent with ARDS), n (%)	1327 (40.5)	231 (36.8)
Indeterminant, n (%)	225 (6.9)	88 (6.9)
No (not consistent with ARDS), n (%)	1951 (59.5)	308 (59.5)
Risk Factors		
Aspiration, n (%)	319 (9.7)	88 (14.0)
Shock, n (%)	1273 (38.9)	246 (39.2)
Pneumonia, n (%)	1276 (38.9)	254 (40.5)
Sepsis, n (%)	1540 (47.0)	345 (55.0)
Pancreatitis, n (%)	35 (1.1)	7 (1.1)
Burn, n (%)	3 (0.1)	0
Overdose, n (%)	79 (2.4)	19 (3.0)
Transfusion, n (%)	1075 (32.8)	115 (18.4)
Congestive Heart Failure, n (%)		
Presence of at least 1 risk factor, n (%)	2282 (69.6)	457 (72.9)
Clinical Outcomes		
Mechanically ventilated, n (%)		
ARDS, n (%)	1327 (40.8)	319 (50.9)
In-hospital Mortality, n (%)	786 (24.0)	247 (39.4)

Supplemental table 2:

Variable Field Name	**Field Type**	**validation**	**variable range (normal)**	**worse numbers (min)**	**worse numbers (max)**	**unidirectional**	**bidirectional**
Description	Field Type	data range	normal range for values	minimum cutoff for abnormal	maximum cutoff for abnormal	Worse value as only minimum or maximum	Worse value can be both minimal or maximum
Demographics							
Age	numeric	18-999					
gender	dichotomous	1 = Male 0= Female					
Race	categorical						
Ethnicity	categorical						
Vital Signs							
Respiratory Rate	numeric	0-60	14-24	<12	>24	no	yes
Temperature	numeric	0-110	97-99	<97	>99	no	yes
Heart rate/pulse	numeric	0-220	50-99	<50	>99	no	yes
Mean arterial blood pressure	numeric	0-200	80-90	<65		min	no
SpO2	numeric	0-100	92-100	<90		min	no
Systolic blood pressure	numeric	0-250	90-150	<90		min	no
Diastolic blood pressure	numeric	0-220	60-80	<60		min	no
Weight	numeric	0-999					
Height	numeric	0-999					
Ventilator settings							
Respiratory rate (Ventilator)	numeric	0-60	10-25	<10	>35	no	yes
Tidal volume (ventilator)	numeric	0-1000	around 350- 450 (unusual)				yes
Tidal volume expired (ventilator)	numeric	0-100	around 350-450 (unusual)				yes
Intubation (Mechanical Ventilation)	dichotomous	1=YES 0= NO					yes
Peak pressure (ventilator)	numeric	0-100	10-30		>40	max	no
Mean airway pressure	numeric	0-100	5-30		>20	max	no
PEEP (ventilator)	numeric	0-40	5-10		>10	max	no
Set FiO2 (ventilator)	numeric	21-100	21-40		>40	max	no
O2 support devices	Categorical						
Lab Values							
Sodium	numeric	0-999	135-150	<135	>145	no	yes
Potassium	numeric	0-20	3.5-5	<3.5	>5	no	yes
Chloride	numeric	0-999	70-140	<70	>115	no	yes
Carbon Dioxide	numeric	0-50	23-29	<23		no	yes

Bicarbonate	numeric	0-50	23-29	<23		min	no
Anion Gap	numeric	0-40	8-16		>16		yes
BUN	numeric	0-300	7-20		>20	max	no
Creatinine	numeric	0-20	0.2-1.2		>1.2	no	yes
Glucose	numeric	0-9999	80-120	<80	>120	no	yes
Calcium (Ca)	numeric	0-20	8.5-10.5	<7	>9	no	yes
Magnesium (Mg)	numeric	0-10	1.7-2.2	<1.7		min	no
Phosphate (Phos)	numeric	0-10	2.5-4.5	<2.5	>4.5	max	yes
HEMOGLOBIN	numeric	0-30	11-15	<10	>15	no	yes
HEMATOCRIT	numeric	0-100	20-40	<20	>40	no	yes
WBC	numeric	0-999	4-10	<4	>10	no	yes
Platelet Count	numeric	0-10000	150-450	<150	>450	no	yes
Neutrophil %	numeric	0-100	55-70			no	yes
Band %	numeric	0-100	0-4		>4	max	no
Alkaline phosphatase	numeric	> 0	20-140		>140	max	no
BILIRUBIN TOTAL	numeric	0-80	0.1-1.2		>1.2	max	no
Direct Bilirubin	numeric	0-80	0.1-1.2		>1.2	max	no
Total Protein	numeric	0-10	6-8.3	<6		max	no
ALBUMIN	numeric	0-10	2.5-5	<2.5		max	no
AST (aspartate aminotransferase)	numeric	0-10000	10-40		>40	max	no
Alanine Transferase (ALT)	numeric	0-10000	7-50		>50	max	no
Lipase	numeric	0-10000	0-160		>160		no
Amylase	numeric	0-10000	20-80		>140		no
PH VENOUS	numeric	6.0-9.0	7.35-7.45	<7.4	>7.55	no	yes
PH ARTERIAL	numeric	6.0-9.0	7.35-7.45	<7.4	>7.55	no	yes
PO2 ARTERIAL	numeric	0-999	60-80	<60		min	no
PO2 VENOUS	numeric	0-999	40-80	<40		min	no
PCO2 ARTERIAL	numeric	0-999	35-45	<35	>45	no	yes
PCO2 VENOUS	numeric	0-999	40-50	<40	>50	no	yes
Lactate	numeric	0-100	<2		>2	max	no
Influenza A	dichotomous	1=YES 0= NO					
Influenza A	dichotomous	1= positive, 0 = negative					
Influenza B	dichotomous	1=YES 0= NO					
Influenza B	dichotomous	1= positive, 0 = negative					
Respiratory Syncytial Virus (RSV)	dichotomous	1=YES 0= NO					
RSV result	dichotomous	1= positive, 0 = negative					
Legionella Antigen, Urine	dichotomous	1=YES 0= NO					

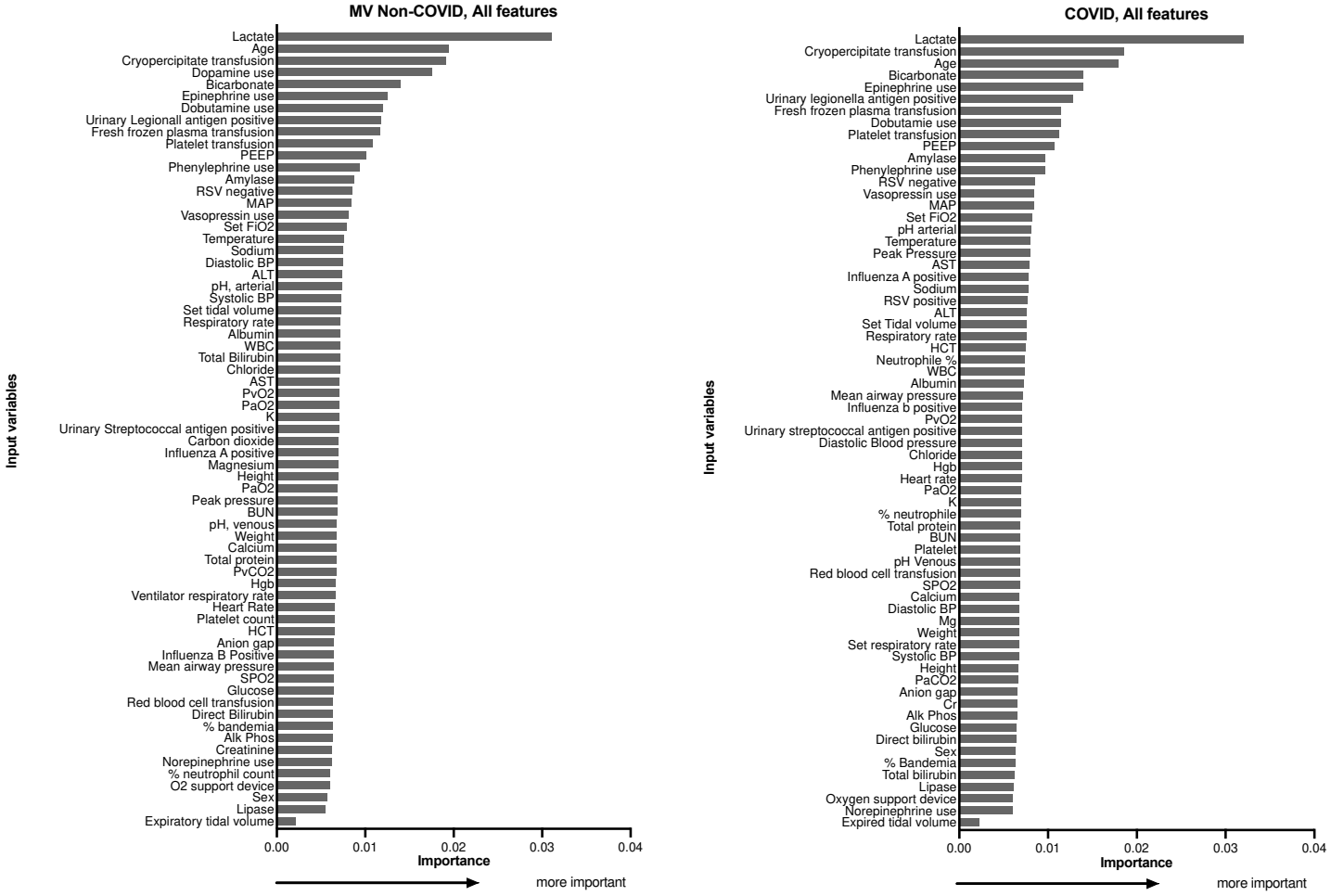
Legionella antigen test result	dichotomous	1= positive, 0 = negative					
Streptococcus pneumoniae Antigen, Urine	dichotomous	1=YES 0= NO					
Streptococcus pneumoniae urinary antigen test result	dichotomous	1= positive, 0 = negative					
Medications							
dobutamine use	dichotomous	1= YES 0= NO					
dobutamine	numeric	0-50	0-25		>0	max	no
dopamine use	dichotomous	1= YES 0= NO					
dopamine	numeric	0-100	0-25		>0	max	no
epinephrine use	dichotomous	1= YES 0= NO					
epinephrine	numeric	0-100	0-100		>0	max	no
norepinephrine use	dichotomous	1= YES 0= NO					
norepinephrine	numeric	0-100	0-100		>0	max	no
vasopressin use	dichotomous	1= YES 0= NO					
Vasopressin	numeric	0-20	0-12		>0	max	no
phenylephrine use	dichotomous	1= YES 0= NO					
phenylephrine	numeric	0-1000	0-500		>0	max	no
Transfusions							
PRBC	dichotomous	1= YES 0= NO			>0	yes	no
FFP	dichotomous	1= YES 0= NO			>0	yes	no
Platelet	dichotomous	1= YES 0= NO			>0	yes	no
Cryoprecipitate	dichotomous	1= YES 0= NO			>0	yes	no
Blood transfusion	dichotomous	1= YES 0= NO			>0	yes	no

Supplemental Table 3

	MV Non-COVID Cohort								COVID Cohort							
vital Signs	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Temperature	503168	98.53	2.13	32	98	98.4	99	109	183738	98.75	1.88	0.3	98	98.5	99.3	109.2
Heart Rate	1164782	88.25	18.6	0	75	87	100	220	282025	89.79	19.01	0	77	89	101	220
Respiratory Rate	1111556	20.64	6.04	0	17	20	24	60	263097	22.03	5.97	0	18	20	25	60
Systolic Blood Pressure	878855	123.7	22.58	0	108	122	138	250	232955	125.73	21.33	0	111	124	139	249
Diastolic Blood Pressure	878888	65.93	15.44	0	56	65	75	211	232958	68.81	13.86	0	59	68	78	201
Mean Arterial Blood Pressure	285389	79.15	15.88	0	69	77	87	200	54093	81.08	16.19	0	70	79	90	199
Pulse Oximetry (SPO2)	1158316	97.58	3.74	0	96	99	100	100	289133	95.98	4.25	0	94	97	99	100
Height	6123	167.89	10.77	116	160.02	167.64	175.26	167.89	4548	167.23	10.74	121.92	160.02	167.64	173.74	167.23
Weight	28652	83.71	24.4	22	66.9	80	95.98	83.71	8054	83.95	23.44	16.9	68.04	80.55	94.87	83.95
Ventilator Setting																
Vent Respiratory Rate	226974	22.69	6.21	0	18	22	26	60	25572	26.47	6.41	0	22	26	31	60
Set FiO2	641611	47.09	16.47	21	40	40	50	100	86027	57.15	21.99	21	40	50	70	100
PEEP	189652	6.4	2.99	0	5	5	7.5	40	21618	10.06	4.79	0	5	10	14	40
Mean Airway Pressure	183169	11.1	4.24	0	8	10	13	100	19746	15.58	5.8	0	11	15	20	100
Peak Pressure	189223	21.36	7.38	0	16	20	26	96	21101	26.42	8.81	0	20	26	32	96
Tidal Volume	166210	405.5	58.28	0	360	400	450	980	19839	387.82	55.21	0	350	400	420	800
Oxygen support device	383870	2.73	2.14	0	1	2	5	5	18663	1.96	1.94	0	0	1	3	5
expiratory tidal volume	1718	36.55	31.78	0	6	29	64	100	155	33.37	29.68	0	11	23	55	100
Laboratory tests																
Serum chemistry																
sodium	103001	139.85	9.38	1	137	140	144	194	36774	139.98	7.31	90	136	139	143	170
potassium	102581	4.2	0.79	1	3.8	4.2	4.6	14.5	35478	4.39	0.73	2	3.9	4.3	4.8	9
chloride	102750	100.9	8.64	1	97	101	105	151	36699	101.04	8.12	70	96	100	105	140
carbon dioxide	87810	23.32	5.27	1	20	23	26	50								
bicarbonate	8927	24.48	5.63	5	21	24	27	40	11992	25.58	5.67	5	22	25	29	40
blood urea nitrogen	102797	34.58	26.39	1	16	26	46	292	37025	41.85	36.87	5	15	28	57	300
creatinine	103026	1.96	1.85	0.1	0.8	1.2	2.46	19.88	37054	2.44	2.87	0.2	0.8	1.2	2.7	20
anion gap	102494	15.44	4.7	0	12	15	18	40	36464	16.16	4.77	1	13	15	19	40
calcium	102840	8.42	0.9	1	7.9	8.4	8.9	17.9	37023	8.53	0.78	3	8.1	8.5	9	15
magnesium	86171	2.09	0.37	0.4	1.9	2.1	2.3	10	25576	2.25	0.46	0.5	2	2.2	2.5	8.5
Total Protein	44435	5.64	1.07	0.7	4.9	5.6	6.3	10	20423	6.39	0.89	1	5.8	6.4	7	10
Albumin	44462	2.87	0.67	1	2.4	2.8	3.3	6.6	20673	3.18	0.64	2	2.7	3.2	3.6	5.7
Total bilirubin	44265	2.68	5.67	0	0.4	0.8	2	76.7	20366	0.75	1.28	0	0.3	0.5	0.7	30.6
direct bilirubin	44261	1.65	3.72	0.1	0.2	0.4	1.1	74.8	20366	0.47	1.04	0.2	0.2	0.2	0.4	20
alanine transferase	44271	130.27	457.63	1	16	30	69	10000	20555	66.37	221.76	6	20	33	59	7000
aspartate aminotransferase	44114	169.34	677.51	1	23	39	79	10000	20013	86.87	414.22	11	27	41	66	10000
alkaline phosphatase	44259	140.5	155.66	0	68	98	154	4995	20360	102.02	73.29	0	63	83	114	2100
glucose	102800	150.25	79.28	1	a	130	173	1932	36942	170.55	107.95	20	104	134	204	2000
Complete blood count																

White blood cell count	97575	12.07	7.55	0	7.6	10.7	14.8	321.5	32297	10.41	8.15	0	6.2	8.8	12.7	326.2
hemoglobin	97576	9.29	1.95	1.4	7.9	8.9	10.3	21.7	32297	10.98	2.52	1.8	8.9	11	12.9	21
hematocrit	97576	28.61	6.35	4	24.5	27.5	32	66.6	32297	34.75	7.82	4	28.7	35.2	40.5	70
platelet count	97577	220.1	133.28	1	124	199	289	1483	32297	260.47	125.61	0	173	240	328	1089
Neutrophile %	97513	65.28	26.02	0	60	74	83	100	32290	8.02	5.37	0	4.3	6.8	10.4	94.7
Band %	8876	5.63	9.12	0	0	2	7	78	4813	4.28	7.64	0	0	1	5	81
Blood Gas																
Arterial pH	58252	7.38	0.09	6.66	7.34	7.4	7.45	7.77	11994	7.35	0.11	6.8	7.29	7.37	7.43	7.73
Arterial PaO2	58341	138.78	83.29	1	83.2	116	162	602	11991	47.7	13.3	13	39	46	54	100
Arterial PaCO2	58382	41.03	10.91	1	34.5	39.6	45.4	132	11991	47.7	13.3	13	39	46	54	100
Venous pH	16825	7.34	0.1	6.34	7.29	7.35	7.4	7.86	5802	7.35	0.1	6.8	7.32	7.37	7.41	7.57
Venous PaO2	16844	49.65	37.5	1	33.4	40.1	52.3	613	5795	42.27	26.53	10	26	37	49	370
Venous PaCO2	16854	46.3	13.39	1	39	44.6	50.5	187	5801	45.4	10.8	12.8	38.7	44.2	50.1	100
Other Chemistry																
Lactate	8798	2.15	2.52	1	1	1.3	2	25	26314	1.94	1.58	1	1.2	1.5	2.1	25
amylase	1809	137.69	258.56	1	42	74	129	4952	469	159.98	368.07	20	51	74	123	4131
lipase	3074	127.75	384.01	5	20	36	81	8457	2198	116.36	489.58	20	24	41	73	8950
Microbiology/Infectious lab																
Urinary Streptococcus pneumoniae Antigen	1295	1	0	1	1	1	1	1	1511	1	0	1	1	1	1	1
Urinary Legionella Antigen	272	1	0	1	1	1	1	1	1558	1	0	1	1	1	1	1
Influenza A	2014	1	0	1	1	1	1	1	402	1	0	1	1	1	1	1
Influenza B	2125	1	0	1	1	1	1	1	1798	1	0	1	1	1	1	1
Respiratory syncytial virus	1415	1	0	1	1	1	1	1								
Medications and Transfusion	· For transfusion and medication, we use binary values, which show if the transfusion or medication is given to the patients or not. The below numbers show how many percent of encounters are given each of the transfusions and medications:															
Packed Red Blood Cells	0.4								0.083							
Platelet	0.14								0.008							
Fresh Frozen Plasma	0.12								0.01							
Cryoprecipitate	0.06								0.0003							
Norepinephrine	0.63								0.153							
Epinephrine	0.15								0.009							
phenylephrine	0.233								0.004							
vasopressin	0.33								0.017							
Dopamine	0.06								0.005							
Dobutamine	0.13								0.0003							

Supplemental Figure 1: Feature importance, LIME explanation of important features in identifying ARDS or in-hospital mortality



Expiratory tidal volume in both cohorts were largely missing. Laboratory electronic health record labels for measurement of serum bicarbonate had changed from “carbon dioxide” to “bicarbonate” were changed over the years. In 2020, serum bicarbonate was only available as “bicarbonate.” RSV was not available in the COVID cohort and both influenza b and influenza a tests were seldom preformed in the COVID cohort