medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Exploring Associations with Severe Hypoxemic Respiratory Failure in COVID-19 Patients upon Admission: A Model for Severe Hypoxemic Respiratory Failure in 329 Unvaccinated, Hospitalized COVID-19 Patients.

An Exploratory Model for Respiratory Failure in Admitted SARS-CoV-2 Patients in Galveston, Texas, USA.

John W. Davis*[1, 2, 3] and Beilin Wang,[1]* Ewa Tomczak,[1] Chia-Chi Fu,[2] Wissam Harmouch,[2] David R. Reynoso,[1] Philip Keiser,[1] Miguel M. Cabada[1, 4]+

1. Division of Infectious Diseases, Department of Internal Medicine, The University of Texas Medical Branch, Galveston, Texas.
2. School of Medicine, The University of Texas Medical Branch, Galveston, Texas
3. Graduate School of Biomedical Sciences, The University of Texas Medical Branch, Galveston, Texas.
4. "Alexander von Humboldt Tropical Medicine Institute, Universidad Peruana Cayetano Heredia, Lima, Peru

*Wang B and Davis JW contributed equally to the study.

Address Correspondence to:

jowdavis@utmb.edu

Word count: 3316

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

ABSTRACT

Objective

The severe acute respiratory syndrome-Coronavirus-2 (SARS-CoV-2) has caused a pandemic claiming more than 4 million lives worldwide. Overwhelming Coronavirus-Disease-2019 (COVID-19) respiratory failure placed tremendous demands on healthcare systems increasing the death toll. Cost-effective prognostic tools to characterize COVID-19 patients' likely to progress to severe hypoxemic respiratory failure are still needed.

Design

We conducted a retrospective cohort study to develop a model utilizing demographic and clinical data collected in the first 12-hours admission to explore associations with severe hypoxemic respiratory failure in unvaccinated and hospitalized COVID-19 patients.

Setting

University based healthcare system including 6 hospitals located in the Galveston, Brazoria and Harris counties of Texas.

Participants

Adult patients diagnosed with COVID-19 and admitted to one of six hospitals between March 19th and June 31st, 2020.

Primary outcome

The primary outcome was defined as reaching a WHO ordinal scale between 6-9 at any time during admission, which corresponded to severe hypoxemic respiratory failure requiring high-flow oxygen supplementation or mechanical ventilation.

Results

We included 329 participants in the model cohort and 62 (18.8%) met the primary outcome. Our multivariable regression model found that lactate dehydrogenase (OR 3.38 (95% CI 2.04-5.59)), qSOFA score (OR: 2.24 (95% CI 1.22-4.12)), neutrophil to lymphocyte ratio (OR:1.08 (95% CI 1.02-1.14)), age (OR: 1.04 (95% CI 1.02-1.07)), BMI (OR: 1.08 (95% CI1.03-1.13)), oxygen saturation or admission SpO2 (OR: 0.91 (95% CI 0.83-0.99)), and admission date (OR: 0.99 (95% CI 0.98-0.99)). The final model showed an area under curve (AUC) of 0.85. The sensitivity analysis and point of influence analysis did not reveal inconsistencies.

Conclusions

Our study demonstrated that a combination of accessible demographic and clinical information provide a powerful predictive tool to identify subjects with CoVID-19 likely to progress to severe hypoxemic respiratory failure.

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

STRENGTHS

- Our study utilized objective and measurable demographic and clinical information regularly available in healthcare settings even among patients unable to communicate.
- Our primary outcome corresponds to WHO ordinal score which would allow compare our results to other studies and in other settings.
- Our model could serve as an effective point of service tool during early admission to assist in clinical management and allocation of resources to unvaccinated patients.

LIMITATIONS

- Our study is a retrospective study of unvaccinated COVID19 patients, and validation of our prediction model in the rest of our study population is still needed.
- In addition, testing our model in a more recent cohort after emergence of new SARS-CoV-2 variants will be needed to assess its robustness.

KEYWORDS:

SARS-CoV-2, COVID19, hypoxemic respiratory failure, WHO ordinal scores, prognostic model, predictive model

ABBREVIATIONS:

Severe acute respiratory syndrome-Coronoavius-2 (SARS-CoV-2), Coronaviurs-Disease-19 (COVID-19), World Health Organization (WHO), Center of Disease Control and Prevention (CDC), University of Texas Medical Branch (UTMB), Texas Department of Criminal Justice (TDCJ), Glawgow coma scale (GCS), Quick Sequential Organ Failure Assessment (qSOFA), Neutrophil to lymphocyte ratio (NLR), Lactate Dehydrogenase (LDH), C-reactive protein (CRP), Body mass index (BMI), Interquartile range (IQR), Standard deviation (SD), Stepwise Akaike Information Criteria (AIC), Goodness of fit (GOF), Receiver operating characteristics curve (ROC), Area under the curve (AUC)

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

INTRODUCTION

The severe acute respiratory syndrome-Coronavirus- 2 (SARS-CoV-2) is a novel coronavirus discovered in 2019. It is the etiologic agent for the largest viral pandemic of the 21st century thus far, followed by H1N1 Influenza A emerged in 2009-2010 [1, 2]. The virus emerged from the Wuhan province of China in December 2019, where it caused a local outbreak of Coronavirus-Disease-19 (COVID-19) characterized by bilateral pneumonia associated with hypoxemic respiratory failure [1]. During the early pandemic, a case series from the Wuhan province showed that 81% of COVID-19 cases were mild, 14% progressed to severe disease, and 5% developed critical illness defined as respiratory failure, septic shock and/or multiple organ dysfunction [3]. Studies evaluating the risk of progression among infected subjects admitted to the hospital have used different outcomes to define severe diseases. These included criteria from the American Thoracic Society's on severity of community acquired pneumonia [4], the Berlin definition of acute respiratory distress syndrome [5], death or mechanical ventilation [6, 7], and/or the World Health Organization (WHO) ordinal scale [8, 9]. The WHO ordinal scale to classify the clinical status of patients with COVID-19 has been widely adopted in randomized control trials such as ACTT-1 and ACTT-2 [10-12]. Harmonization of the measures used to evaluate the severity COVID-19 across different studies could ease the comparison of study results and application of evidence-based interventions.

COVID19 associated hospitalizations caused an overwhelming demand on the healthcare system of the United States. Shortage in ventilators and personal protection equipment posed significant challenges in management of cases in US hospitals early

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

in the pandemic [13]. During 2020, CDC estimated 345,000 deaths attributed to COVID19 [14]. Mask mandates and COVID-19 vaccination were associated with decreased new cases and deaths, especially amongst elderly populations [15-17]. However, highly transmissible variants, such as Delta, have displaced wild type SARS-Cov-2 and become the predominant lineage in the US during summer 2021 [18]. Unvaccinated COVID-19 individuals remain up to 25 times more likely to be hospitalized or dead compared to vaccinated individuals. Rising hospitalizations and deaths among unvaccinated individuals are driving a new pandemic surge posing again a significant burden to the health system [19]. Effective tools to predict severe respiratory failure remain crucial for clinicians to plan interventions and allocate resources for admitted COVID19 patients.

SARS-CoV-2 severity and mortality are associated with individual characteristics and risk factors. Prediction tools to identify patients likely to progress to severe diseases have been proposed using variables such as age, comorbid conditions, and laboratory tests [4, 5, 7-9, 20-24]. However, the heterogeneity in the definitions of severe illness and the limited availability of certain laboratory tests, especially in low-resource settings, have decreased the generalizability of these tools. Laboratory tests such as serum IL-6 or procalcitonin may not be accessible in small medical centers. Similarly, information on comorbidities may not be available in patients unable to provide a history. Simple, objective, and accessible tools to predict progression to severe COVID 19 are still needed to guide clinicians during case surges and dwindling of resources.

We conducted a retrospective cohort study in the University of Texas Medical Branch Health System to develop an exploratory model for severe hypoxemic respiratory failure

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

in unvaccinated, hospitalized COVID-19 patients. We defined severe disease as requiring high-flow oxygen or mechanical ventilation as suggested by the ordinal scale scores proposed by WHO [25]. We used readily available information within 12 hours of admission such as demographics, vital signs and laboratory values to construct a sensitive and specific model.

METHODS

Study Design

We hypothesized that a combination of objective clinical and laboratory findings on admission can identify subjects with higher risk of progression to severe respiratory failure due to COVID-19 in our hospitals. To test this hypothesis, we performed a retrospective, multi-site cohort study on adult patients admitted for COVID-19 to the University of Texas Medical Branch (UTMB) Health System. Our health system includes 6 hospitals located in the Galveston, Brazoria, and Harris counties of Texas. These hospitals are distributed in the Galveston, League City, Clear Lake, and Angleton campuses and have trauma center level designations ranging between levels I and IV. We retrieved the medical record numbers of all patients ≥ 18 years old admitted to hospitals in any of the four campuses with a positiveSARS-CoV-2 molecular test between March 19th and June 31st of 2020. We used the WHO ordinal scale of disease severity for COVID-19 to define our outcomes [25]. This is an eleven-category ordinal scale ranging from a value of zero for patients with no virological evidence of infection to 10 for patients who died due to COVID-19. Our primary outcome was defined as reaching a WHO ordinal scale between 6-9 during admission corresponding to severe respiratory failure requiring oxygen supplementation using high-flow nasal cannula or

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

mechanical ventilation. Patients presenting with a WHO ordinal scale < 6 that were discharged at the time of review of their medical record were enrolled. Patients who met ordinal scale 6-9 scores on the first vital signs obtained on admission were excluded. The maximum ordinal scale score met during admission was consider the subjects' ordinal scale. The study protocol was approved by UTMB Institutional Review Board (20-0126) and the Texas Department of Criminal Justice (TDCJ) Institutional Review Board (#819-RM20).

Patient or Public Involvement: None

Data Collection

We collected data directly from the Epic (Verona, Wisconsin) electronic medical records. The data was transcribed into a questionnaire created in the REDCap (Nashville, Tennessee) data capture system. Data coders were trained using a dummy dataset before using medical records. All coders were trained until they could obtain 100% accuracy on dummy datasets before proceeding to data collection. Eighty-nine randomly selected charts underwent evaluation by the principal investigators and the data extraction personnel. These evaluations were compared to calculate the inter-rater reliability using Kappa statistics. When the personnel had a Kappa < 0.8, they were re-trained, and discrepancies were discussed with the principal investigators. Evaluations were repeated until a Kappa >0.8 was reached. The data extraction personnel collected data on demographics, clinical history and course, vital signs, peak oxygen requirement, and laboratory results. Demographic data included date-of-birth, sex, body mass index (BMI) and admission campus. The clinical data collected included discharge status, discharge date, date of demise (if applicable), do not intubate (DNI) status, first set of

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

vital signs, presenting symptoms, symptom duration, comorbidities (diabetes mellitus, cardiovascular, pulmonary, and kidney disease), oxygen saturation, oxygen supplementation modality, and oxygen flow requirements. Cardiovascular diseases included hypertension, coronary artery disease, heart failure, history of cardiac resynchronization device, or left ventricular assisting device. Lung diseases included asthma, chronic obstructive airway disease, pulmonary hypertension, obstructive sleep apnea, interstitial lung disease, use of home oxygen, or use of home continuous positive airway pressure (CPAP) or bilevel positive airway pressure (BiPAP). Chronic kidney disease included any stage 1-5 or end stage renal disease requiring renal replacement therapy. Subjects with an admission note documenting "disorientation" or "inability to follow commands" were deemed to have a Glasgow coma scale (GCS) < 15. The systolic blood pressure, respiratory rate, and GCS were used to calculate the Quick Sequential Organ Failure Assessment (qSOFA). The maximum oxygen requirement at any given day after admission was used as the peak oxygen requirement, and the subject was deemed to have met the primary endpoint if the peak oxygen requirements was high flow nasal canula or mechanical ventilation. Data on admission laboratory results include absolute neutrophil and lymphocyte counts; serum lactate dehydrogenase (LDH), D-dimer, C-reactive protein (CRP), procalcitonin, and troponin I. Only the first laboratory tests obtained within 12 hours of admission were recorded.

Statistical Analysis

The REDCap dataset was downloaded to a database on SAS (Version 9.4, Cary NC) and R (Version 4.0.2). Frequencies, means with standard deviations (±SD), and medians with interquartile ranges (IQR) were calculated to describe the distribution of

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

the variables. Pearson correlation was conducted to assess collinearity between covariates. A Pearson correlation coefficient ≥ 0.80 was considered highly collinear, whereas lower values were assessed by both strength of correlation and biologic plausibility. High collinearity between two variables did not occur in our dataset. Mean imputation was used to replace BMI when the value was missing. Multiple imputation was not performed because data were not missing at random relative to the primary outcome. To evaluate the effect of utilizing the sample mean to replace BMI missing values, the analysis was also performed excluding those cases.

A multivariable logistic regression analysis was conducted to model variables with the highest predictive value for severe COVID-19. Variables for the model were selected based on review of the literature on COVID-19 and clinical relevance (e.g, objectivity and availability). The composite variables age/BMI and age/sex were created and utilized in the models because of existing evidence of interactions between those variables individually.[1] Stratified analysis was performed for interaction terms ultimately included in the model. Stepwise Akaike Information Criteria (AIC) reduction was utilized to optimize the model, reducing residual deviance while prioritizing model simplicity. Cook's Distance method was utilized for assessing points of influence, where a Cook's D ≥ 1 was considered highly influential. The Hoslem-Lemeshow goodness of fit (GOF) test statistic was utilized to evaluate the match between the predicted and observed risk of progression to WHO ordinal score 6 to 9. A receiver operating characteristics curve (ROC) and area under the curve (AUC) analysis was performed to assess overall model fidelity. Misclassification tables at various cut-points were created to determine optimal

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

threshold score to use the model as a screening tool, with a target sensitivity of at least 90%.

RESULTS

We identified 930 verified subjects admitted to the UTMB Health System with a positive SARS-CoV-2 test during March 19th – June 31st, 2020. The first 352 consecutive charts were reviewed to develop the predictive models. The demographics and clinical characteristics of the cohort prior to exclusion are shown in Supplemental Figure 1. 18 subjects were excluded because they met WHO ordinal scale scores between 6-9 on the first vital signs measured, and 5 were excluded due to numerous missing values (Figure 1).

[Figure 1]

329 subjects were included in the cohort and 62 (18.8%) met the primary endpoint. The TDCJ population accounted for 27.6% of cohort population but there were no significant differences in the proportion of subjects meeting the primary endpoint according to inmate status ($p = 0.459$, data not shown). Subjects reaching the ordinal scale 6-9 were significantly older than subjects who did not (Table 1). More male subjects met the primary end point but the difference between groups was not statistically significant (Table 1). The top three comorbidities for subjects with ordinal scales < 6 were cardiovascular 51.5%, diabetes mellitus 32.4%, and pulmonary 21.0%. For subjects with ordinal scale 6-9, the top three conditions were cardiovascular 67.3%, diabetes mellitus 38.5%, and liver disease 15.4% (Table 1). Twenty-four subjects died during admission (7.3%) and 20 of them met criteria for ordinal score 6-9. Seven percent

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

(6.7%) of subjects with ordinal score <6 and 9.7% of subjects with ordinal scale 6-9 had DNI order. Comfort care was implemented in 1.9% of those with ordinal score <6 and in 16.1% of those with ordinal score 6-9. The characteristics of subjects with ordinal scales 6-9 across all campuses are shown (Table 1).

Table 1. Demographic and clinical characteristics of subjects included in the cohort

		WHO ordinal scale < 6 (N = 267)	WHO ordinal scale 6-9 (N = 62)	
		Mean (± SD)		p-value
	Age, years	55.9 (17.8)	62.9 (13.4)	<0.001
	qSOFA	0.288 (0.478)	0.581 (0.560)	<0.001
	Oxygen saturation, %	95.9 (3.38)	93.4 (4.56)	<0.001
	C-reactive protein, mg/dL	4.60 (7.78)	11.7 (10.1)	<0.001
	NLR***	5.69 (5.10)	9.92 (11.2)	0.005
	BMI	31.7 (7.44)	33.9 (8.93)	0.068
	D-dimer, □g/mL	2.10 (8.13)	3.35 (16.6)	0.564
Categorized LDH	LDH, U/L Normal	180 (67.4)	18 (29.0)	<0.001
	1-2 x ULN †	76 (28.5)	30 (48.4)	
	> 2x ULN †	11 (4.1)	14 (22.6)	
		N (%)		
Discharge status	Death	4 (1.5)	20 (32.3)	<0.001
	Alive	263 (98.5)	42 (67.7)	
Code status	Regular	244 (91.4)	46 (74.2)	<0.001
	DNI	18 (6.7)	6 (9.7)	
	Comfort care	5 (1.9)	10 (16.1)	
Sex	Male	154 (57.7)	41 (66.1)	0.282

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

	Female	113 (42.3)	21 (33.9)	
Campus	Galveston	150 (57.3)	23 (44.2)	-
	Angleton	20 (7.6)	4 (7.7)	
	League City	40 (15.3)	15 (28.8)	
	Clear Lake	47 (17.9)	8 (15.4)	
	Non UTMB transfer	4 (1.5)	2 (3.8)	
Comorbidities	Cardiovascular	161 (51.5)	35 (67.3)	-
	Diabetes Mellitus	85 (32.4)	20 (38.5)	
	Pulmonary	55 (21.0)	6 (11.5)	
	Renal	34 (13.0)	5 (9.6)	
	Liver	22 (8.4)	8 (15.4)	
	HIV	2 (0.8)	2 (3.8)	
	Malignancy on chemotherapy	3 (1.1)	1 (1.9)	
	Solid organ transplant	6 (2.3)	1 (1.9)	
	Dementia	11 (4.2)	2 (3.8)	
	Other	35 (13.4)	5 (9.6)	
	No known medical conditions	34 (13.0)	6 (11.5)	

† Upper limit of normal

** One subject (0.4%) with ordinal scale <6 lacked information on this variable

*** Neutrophil to lymphocyte ratio

The variables included in the initial regression model were admission date, age/sex, age/BMI, oxygen saturation, neutrophil to lymphocyte ratio (NLR), procalcitonin, D-dimer, LDH, CRP, troponin I, and qSOFA score (Supplemental Figure 2). The initial model was highly significant and identified several candidate predictive variables. The candidate clinical and laboratory variables age, BMI, oxygen saturation, qSOFA score, CRP, procalcitonin, NLR, D-dimer and LDH were incorporated into prognostic model. All subjects with elevated troponin I levels (6/6) were intubated which precluded the

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

evaluation of this variable as a predictor in the analysis. After stepwise AIC reduction, the final model included 7 variables: oxygen saturation, NLR, D-dimer, qSOFA, LDH, age, BMI, and admission date (Table 2).

Table 2. Final multivariable regression model after stepwise AIC reduction.

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Admission Date	0.989	0.981	0.998
Age	1.044	1.020	1.069
BMI	1.082	1.034	1.133
Admission SpO2	0.907	0.834	0.987
Neut./Lymph. Ratio	1.078	1.023	1.137
Admission D-Dimer	0.979	0.954	1.005
LDH	3.379	2.043	5.588
qSOFA Score	2.240	1.218	4.118

LDH was categorized as normal, 1x<X<2x upper limit of normal (ULN), and >2x ULN.

Running the model excluding subjects with missing BMI values (n=34) did not affect the general significance or goodness of fit of the model (Table 3).

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Table 3. Multiple regression model excluding cases with missing BMI

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Admission Date	0.989	0.981	0.998
Age	1.045	1.019	1.071
BMI	1.083	1.035	1.133
Admission SpO2	0.915	0.841	0.996
Neut./Lymph. Ratio	1.063	1.012	1.117
LDH	3.156	1.924	5.177
qSOFA Score	1.918	1.030	3.574

LDH was categorized as normal, 1x<X<2x ULN, and >2x ULN.

ROC/AUC analysis of the final model indicated an area under the curve of 0.85, indicating high efficacy of the overall model in predicting severe disease (Figure 2).

[Figure 2]

ROC/AUC analysis of the BMI sensitivity analysis was nearly identical to the final model (AUC = 0.85; Figure 3).

[Figure 3]

Modeling with a cohort excluding patients with active DNI did not produce differences in variables included in the model (Supplemental Figure 3).

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

DISCUSSION

Our study evaluated demographic and clinical variables measured within 12 hours of hospital admission in COVID19 patients as potential predictors of progression to severe respiratory failure. We used the WHO ordinal scale 6-9 to define patients with severe respiratory failure requiring significant life sustaining therapies. Our analysis demonstrated that a combination of routine accessible laboratory tests, vital signs, and demographic variables provided a powerful predictive model to assess risk of COVID-19 severe respiratory failure among patients admitted to our health system. This model used objective and measurable information available in acute care settings even if the patient is unable to communicate. The model could serve as an effective point of service tool during early admission to assist in clinical management and allocation of resources to unvaccinated persons.

LDH was highly predictive of severe disease in our model. Subjects with abnormal LDH on admission were 3.38 times more likely to progress to severe hypoxemic respiratory failure. This finding is supported by prior reports that LDH can predict severity of disease [4, 5]. However, our findings differ from those by Liang et al. showing increased LDH only increase the likelihood of severe disease by 0.2% [4]. A recent meta-analysis by Katzenschlager et al. evaluated the association between LDH levels and admission to the intensive care unit (12 studies) or death (23 studies) in patients with COVID 19. Although, LDH levels were statistically higher in those critically ill (pooled difference of medians: 140 U/L (95% CI 81-199) and those who died (pooled difference of medians: 189 U/L (95% CI 155-223)), the modest absolute increase in LDH levels was deemed clinically irrelevant by the authors [26]. The differences between these studies and ours

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

may be explained by the different outcome definitions used. In our system, the use of high flow nasal cannula was not necessarily associated with ICU admission but was included as part of the endpoint.

NLR has been associated with adverse outcomes in COVID19 patients [4, 7, 9]. Adverse outcomes observed in these studies also included death, which likely accounted for the absolute risk difference in our study [4, 9]. Loannou et al. reported that a ratio higher than 12.7 was associated with a 2.5-fold increase in the odds for mechanical ventilation in COVID19 patients [7]. In our cohort, a higher NLR was associated with modest increases in the odds for reaching ordinal scale 6-9. The contrast between our results and those of Loannou et al. may be related to the inclusion of less severe disease categories in our primary endpoint such as receiving high-flow nasal canula.

In our cohort, age and BMI were important predictors of COVID-19 respiratory failure. BMI was positively associated with progression to WHO ordinal score 6- 9. While the mean BMI imputation biases towards significance, multiple imputation is not appropriate when data are not missing randomly relative to primary outcome. However, excluding patients who did not have valid BMI data did not meaningfully change our model findings. In addition, the overall AUC did not change in comparison to the original model. Although, the age/BMI composite variable was retained by stepwise AIC reduction after excluding cases without valid BMI data, using a composite variable introduces unneeded complexity to the model. Thus, we ultimately decided to exclude the term to maintain simplicity. These data highlighted the association of BMI with severe COVID-19 and add to previous studies that support this association [7, 21, 27].

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

A proposed mechanism for this association is the increased labor of breathing in patients with high body mass index that impairs their capacity to adjust to changes in lung function leading to earlier non-invasive ventilation or mechanical ventilation [27]. We have chosen our predictors based on clinical practicality and mechanistic plausibility. While D-dimer elevations were not found to be a significant predictor for respiratory failure or death in some studies [4, 8], others found an association with adverse outcomes early in the pandemic [28, 29]. Pulmonary vasculature thrombosis was observed in autopsies of COVID-19 patients [30], and a recent study has suggested that heparin-based anticoagulation may protect noncritical COVID-19 patients from inpatient death [31]. In our model, D-dimer was not statistically associated with developing the primary endpoint in the multivariate analysis. However, removing D-dimer decreased our model performance. Given published literature and the mechanistic plausibility for severe diseases in those with elevated D-dimers we decided to keep it in the model.

A smaller proportion of subjects in our cohort met the WHO ordinal scale 6-9 than subjects in cohorts published earlier 2020. Only 19% of the subjects included in our modeling cohort met the primary endpoint compared to 22-26% reported by other authors [6, 8]. Our cohort was enrolled during a phase of rapidly evolving COVID-19 therapies and management approaches. With improvements in early interventions against virus replication and associated inflammation the number of patients requiring high flow oxygen or mechanical ventilation is expected to change. We also found a lower inpatient death rate compared to reports published around the same period [8]. All the subjects included in our cohort have been discharged at the time of data

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

collection. It is possible that a subpopulation of these subjects was re-admitted and expired after data collection was completed. Our study design limited data collection to the primary subject admission and may have missed mortality that occurred in subsequent encounters.

We included patients with DNI and comfort care orders in our cohort. Although, this is a group of subjects that would have not been able to reach all the ordinal scale scores in our endpoint, they would have been eligible for high flow oxygen and vasopressors. The stratified analysis that omitted subjects with DNI did not significantly change the prediction power of the model. The inclusion of this subpopulation in our cohort likely provided a conservative estimate of the odds of meeting ordinal scales 6-9.

Our study has several limitations to acknowledge. The troponin was not included in our model because all subjects with abnormal troponin met the primary outcome. Elevated troponin suggested myocardial injury which can be due to a direct effect from SARS-CoV-2 infection and/or a complication from sepsis and the inflammatory response described in COVID 19. The role of troponin as a predictor of COVID-19 associated mortality has been suggested in other studies [32, 33]. However, larger studies are necessary to evaluate their role in predicting severe COVID-19 respiratory failure. Our cohort was constructed prior to introduction of COVID-19 vaccination and therapeutic interventions such as dexamethasone or remdesivir [34]. The validation of our prediction model in the rest of our study population and in more recent cohorts after the emergence of new SARS-CoV-2 variants will be needed to assess its robustness. Our prediction model could contribute to the literature providing tools for clinicians, however

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

none of the variables in our model are specific to SARS-CoV-2 infection and further studies including other populations will be needed for validation.

CONCLUSION

This study provides a preliminary model for early identification of COVID19 patients at odds of progressing to severe COVID-19 within the first 12 hours of admission. This model will require further validation in larger datasets. Future studies will use this model as a tool for predicting severe COVID-19 disease in resource limited settings where effective vaccines and therapies are still unavailable.

Figure Legends

Figure 1: Flow Chart for Cohort Selection

Figure 2: ROC Curve, Final Model

Figure 3: ROC Curve, BMI Sensitivity Analysis

AUTHOR CONTRIBUTION

Fu C and Harmouch W both contributed to data collection and manuscript drafting. Reynoso DR and Keiser P both contributed to project development and data allocation. Cabada MC has been the senior author since project inception and providing editorial revision to the manuscript. Davis JW performed all statistical analysis and drafted the methods, results, and discussion sections in conjunction with Tomczak E and Wang B. Tomczak E was responsible for project conception, some data collection, and some manuscript drafting, and Wang B was responsible for the project conception, data collection execution and drafting all parts of manuscript.

ACKNOWLEDGEMENT

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

We thank Daniel Z. Bao, Ashley E Chen and Kyra Curtis from UTMB Medical School for their contribution to data collection. We thank Dr. Ion Mitrache, Director of Clinical Analyst at UTMB, for assistance in clinical data management.

COMPETING INTEREST

All authors have completed ICMJE forms and declare no competing interest in this study.

DISCLOSURE

The research contained in this document was coordinated in part by the Texas Department of Criminal Justice (Research Agreement #819-RM20). The contents of this document reflect the views of the author(s) and do not necessarily reflect the views or policies of the Texas Department of Criminal Justice.

FUNDING

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

DATA SHARING

SAS Code and de-identified data will be published at Dryad, an online repository.

ETHIC STATEMENT

The study protocol has obtained waiver of informed consent under the approvals of Institutional Review Board from UTMB (20-0126) and Texas Department of Criminal Justice (TDCJ) under protocol number #819-RM20.

REFERENCES

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

1. WHO. *COVID-19 clinical management. Living guidance.* 2021 January 25, 2021 [cited 2021 May 13th]; Second:[Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-clinical-2021-1.
2. WHO. *Past pandemics*. [cited 2021 June 11th]; Available from: https://www.euro.who.int/en/health-topics/communicable-diseases/influenza/pandemic-influenza/past-pandemics.
3. Wu, Z. and J.M. McGoogan, *Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72□314 Cases From the Chinese Center for Disease Control and Prevention.* Jama, 2020. **323**(13): p. 1239-1242.
4. Liang, W., et al., *Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19.* JAMA Intern Med, 2020. **180**(8): p. 1081-1089.
5. Poggiali, E., et al., *Lactate dehydrogenase and C-reactive protein as predictors of respiratory failure in CoVID-19 patients.* Clin Chim Acta, 2020. **509**: p. 135-138.
6. Cummings, M.J., et al., *Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study.* Lancet, 2020. **395**(10239): p. 1763-1770.
7. Ioannou, G.N., et al., *Risk Factors for Hospitalization, Mechanical Ventilation, or Death Among 10□131 US Veterans With SARS-CoV-2 Infection.* JAMA Netw Open, 2020. **3**(9): p. e2022310.
8. Garibaldi, B.T., et al., *Patient Trajectories Among Persons Hospitalized for COVID-19 : A Cohort Study.* Ann Intern Med, 2021. **174**(1): p. 33-41.
9. Wongvibulsin, S., et al., *Development of Severe COVID-19 Adaptive Risk Predictor (SCARP), a Calculator to Predict Severe Disease or Death in Hospitalized Patients With COVID-19.* Ann Intern Med, 2021. **174**(6): p. 777-785.
10. WHO. *WHO R&D Bluprint. novel Coronavirus. COVID-19 Therapeutic Trial Synoposis*. 2020 Februrary 18, 2020 [cited 2021 May 15th]; Available from: https://www.who.int/publications/i/item/covid-19-therapeutic-trial-synopsis
11. Beigel, J.H., et al., *Remdesivir for the Treatment of Covid-19 - Final Report.* N Engl J Med, 2020. **383**(19): p. 1813-1826.
12. Kalil, A.C., et al., *Baricitinib plus Remdesivir for Hospitalized Adults with Covid-19.* N Engl J Med, 2021. **384**(9): p. 795-807.
13. Ranney, M.L., V. Griffeth, and A.K. Jha, *Critical Supply Shortages - The Need for Ventilators and Personal Protective Equipment during the Covid-19 Pandemic.* N Engl J Med, 2020. **382**(18): p. e41.
14. Gundlapalli, A.V., et al., *Death Certificate-Based ICD-10 Diagnosis Codes for COVID-19 Mortality Surveillance - United States, January-December 2020.* MMWR Morb Mortal Wkly Rep, 2021. **70**(14): p. 523-527.
15. Guy, G.P., Jr., et al., *Association of State-Issued Mask Mandates and Allowing On-Premises Restaurant Dining with County-Level COVID-19 Case and Death Growth Rates - United States, March 1-December 31, 2020.* MMWR Morb Mortal Wkly Rep, 2021. **70**(10): p. 350-354.
16. Christie, A., et al., *Decreases in COVID-19 Cases, Emergency Department Visits, Hospital Admissions, and Deaths Among Older Adults Following the*

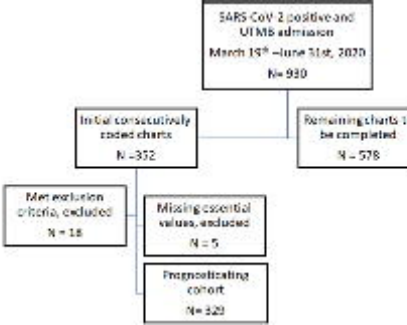
medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

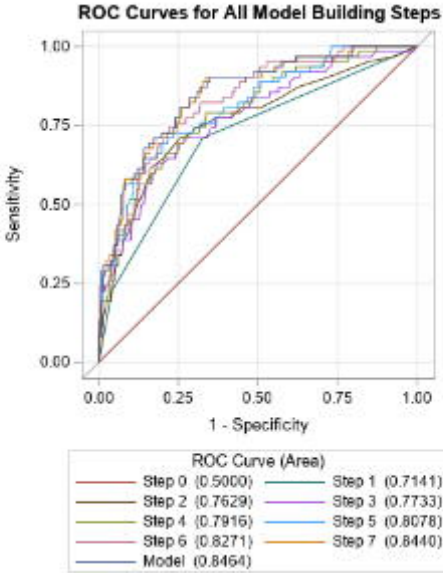
Introduction of COVID-19 Vaccine - United States, September 6, 2020-May 1, 2021. MMWR Morb Mortal Wkly Rep, 2021. **70**(23): p. 858-864.

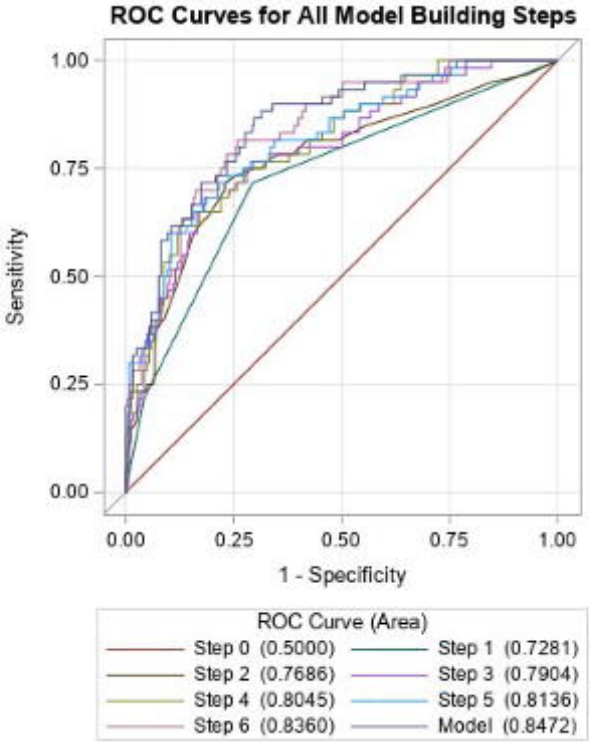
17. Joo, H., et al., *Decline in COVID-19 Hospitalization Growth Rates Associated with Statewide Mask Mandates - 10 States, March-October 2020.* MMWR Morb Mortal Wkly Rep, 2021. **70**(6): p. 212-216.
18. CDC. *COVID Data Tracker Weekly Review 8/5/2021*. 2021 [cited 2021 8/5]; Available from: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html.
19. Officials, W.H.C.-R.T.a.P.H., *Press Briefing by White House COVID-19 Response Team and Public Health Officials*. Aug 2, 2021: The White House, United States.
20. Wynants, L., et al., *Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal.* Bmj, 2020. **369**: p. m1328.
21. Simonnet, A., et al., *High Prevalence of Obesity in Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) Requiring Invasive Mechanical Ventilation.* Obesity (Silver Spring), 2020. **28**(7): p. 1195-1199.
22. Tjendra, Y., et al., *Predicting Disease Severity and Outcome in COVID-19 Patients: A Review of Multiple Biomarkers.* Arch Pathol Lab Med, 2020. **144**(12): p. 1465-1474.
23. Wu, C., et al., *Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China.* JAMA Intern Med, 2020. **180**(7): p. 934-943.
24. Ji, D., et al., *Prediction for Progression Risk in Patients With COVID-19 Pneumonia: The CALL Score.* Clin Infect Dis, 2020. **71**(6): p. 1393-1399.
25. *A minimal common outcome measure set for COVID-19 clinical research.* Lancet Infect Dis, 2020. **20**(8): p. e192-e197.
26. Katzenschlager, S., et al., *Can we predict the severe course of COVID-19 - a systematic review and meta-analysis of indicators of clinical outcome?* PLoS One, 2021. **16**(7): p. e0255154.
27. Anderson, M.R., et al., *Body Mass Index and Risk for Intubation or Death in SARS-CoV-2 Infection : A Retrospective Cohort Study.* Ann Intern Med, 2020. **173**(10): p. 782-790.
28. Liao, D., et al., *Haematological characteristics and risk factors in the classification and prognosis evaluation of COVID-19: a retrospective cohort study.* Lancet Haematol, 2020. **7**(9): p. e671-e678.
29. Zhou, F., et al., *Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study.* Lancet, 2020. **395**(10229): p. 1054-1062.
30. Ackermann, M., et al., *Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19.* N Engl J Med, 2020. **383**(2): p. 120-128.
31. Chowdhury, J.F., L.K. Moores, and J.M. Connors, *Anticoagulation in Hospitalized Patients with Covid-19.* N Engl J Med, 2020. **383**(17): p. 1675-1678.
32. García de Guadiana-Romualdo, L., et al., *Cardiac troponin and COVID-19 severity: Results from BIOCOVID study.* Eur J Clin Invest, 2021. **51**(6): p. e13532.

medRxiv preprint doi: https://doi.org/10.1101/2021.11.05.21265970; this version posted November 9, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

33. Wibowo, A., et al., *Prognostic performance of troponin in COVID-19: A diagnostic meta-analysis and meta-regression.* Int J Infect Dis, 2021. **105**: p. 312-318.
34. *Coronavirus Disease 2019 (COVID-19) Treatment Guidelines*. Aug 17th 2021 August 17, 2021 [cited 2021 August 20, 2021]; Available from: https://www.covid19treatmentguidelines.nih.gov/about-the-guidelines/whats-new/.


SARS-CoV-2 positive and UTMB admission
March 19th –June 31st, 2020
N= 910
Initial consecutively coded charts
N = 332
Remaining charts to be completed
N = 578
Met exclusion criteria, excluded
N = 18
Missing essential values, excluded
N = 5
Prognosticating cohort
N= 309


ROC Curves for All Model Building Steps
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00
ROC Curve (Area)
Step 0 (0.5000)
Step 1 (0.7141)
Step 2 (0.7629)
Step 3 (0.7733)
Step 4 (0.7916)
Step 5 (0.8078)
Step 6 (0.8271)
Step 7 (0.8440)
Model (0.8464)


ROC Curves for All Model Building Steps
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00
ROC Curve (Area)
Step 0 (0.5000)
Step 1 (0.7281)
Step 2 (0.7686)
Step 3 (0.7904)
Step 4 (0.8045)
Step 5 (0.8136)
Step 6 (0.8360)
Model (0.8472)

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	229.0	99.3147	5.3166	0.0211
Admission Date	1	-0.0102	0.00451	5.1249	0.0236
Biological Sex	1	0.5077	1.7130	0.1177	0.7315
Age	1	-0.00041	0.0597	0.0000	0.9946
Biological Sex*Age	1	-0.0130	0.0257	0.2586	0.6124
BMI	1	-0.0486	0.1084	0.1936	0.6599
BMI*Age	1	0.00203	0.00195	1.0398	0.2193
Admission SpO2	1	0.0963	0.0439	4.8183	0.0282
Neut./Lymph. Ratio	1	0.0762	0.0288	6.9776	0.0083
Procalcitonin	1	0.2200	0.5481	0.1612	0.6881
Admission D-Dimer	1	-0.0204	0.0137	2.2070	0.1374
LDH	1	1.1513	0.2995	14.7743	0.0001
CRP	1	0.2342	0.4337	0.2917	0.5891
qSOFA Score	1	0.6052	0.3196	7.6889	0.0056

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Age	1.044	1.020	1.069
BMI	1.082	1.034	1.133
Admission SpO2	0.907	0.834	0.987
Admission Date	0.989	0.981	0.998
Neut./Lymph. Ratio	1.078	1.023	1.137
Admission D-Dimer	0.979	0.954	1.005
LDH	3.379	2.043	5.588
qSOFA Score	2.240	1.218	4.118