Predicting Criticality in COVID-19 Patients

Roger A. Hallman*
Roger.Hallman.TH@dartmouth.edu
Thayer School of Engineering,
Dartmouth College
Hanover, New Hampshire, USA &
Naval Information Warfare Center
Pacific
San Diego, California, USA

Anjali Chikkula†
anjali.k.chikkula.20@dartmouth.edu
Evercore Inc.
Houston, Texas, USA

Temiloluwa Prioleau
tprioleau@dartmouth.edu
Dartmouth College
Hanover, New Hampshire, USA

ABSTRACT

The COVID-19 pandemic has infected millions of people around the world, spreading rapidly and causing a flood of patients that risk overwhelming clinical facilities. Whether in urban or rural areas, hospitals have limited resources and personnel to treat critical infections in intensive care units, which must be allocated effectively. To assist clinical staff in deciding which patients are in the greatest need of critical care, we develop a predictive model based on a publicly-available data set that is rich in clinical markers. We perform statistical analysis to determine which clinical markers strongly correlate with hospital admission, semi-intensive care, and intensive care for COVID-19 patients. We create a predictive model that will assist clinical personnel in determining COVID-19 patient prognosis. Additionally, we take a step towards a global framework for COVID-19 prognosis prediction by incorporating statistical data for geographically and ethnically diverse COVID–19 patient sets into our own model. To the best of our knowledge, this is the first model which does not exclusively utilize local data.

CCS CONCEPTS

• **Applied computing** → **Health informatics**; • **Information systems** → **Decision support systems**; *Data analytics.*

KEYWORDS

COVID-19, Statistical Analysis, Predictive Modeling, Decision Support

ACM Reference Format:
Roger A. Hallman, Anjali Chikkula, and Temiloluwa Prioleau. 2020. Predicting Criticality in COVID-19 Patients. In *Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (BCB '20), September 21–24, 2020, Virtual Event, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3388440.3412463

*Roger A. Hallman is supported by the United States Department of Defense SMART Scholarship for Service Program, funded by USD/R&E (The Under Secretary of Defense-Research and Engineering), National Defense Education Program (NDEP) / BA-1, Basic Research.

†Work performed while a student at Dartmouth College, Hanover, New Hampshire, USA.

BCB '20, September 21–24, 2020, Virtual Event, USA

https://doi.org/10.1145/3388440.3412463

1 INTRODUCTION

The Coronavirus Disease 2019 (COVID-19) pandemic has infected millions of people and led to hundreds of thousands of fatalities around the world [2]. Since its January 2020 arrival in the United States [21], many large American cities have become epicenters of the disease. Many hospitals, which already have limited critical care resources, run the risk of being overwhelmed by COVID-19 cases requiring hospitalization whether in major urban centers or more rural areas [3]. Moreover, widespread, accurate testing for COVID-19 infection has been slow to scale to an adequate capacity for a country as large as the United States and there are multiple documented discrepancies in access to testing [14].

Using data sets from early COVID-19 hospitalizations, ML models can be constructed to predict hospitalization and progression to critical illness. Given that hospitals have limited intensive care capacity, predictive modeling offers clinicians the capability to efficiently route patients, as well as allocate resources and personnel. To assist clinical personnel with patient routing and resource allocation during treatment, we develop a ML-based predictive model that provides an easily explainable, rule-based framework to determine which level of care a COVID-19 patient requires.

Our main contribution in this work is the development of a ML-based model for predicting criticality in patients admitted to a clinical facility after testing positive for COVID-19. While the majority of existing prognosis prediction models focus on predicting mortality risk, our model determines whether a COVID-19 patient should be (*i*) sent home to self-quarentine, (*ii*) admitted to the hospital for monitoring, (*iii*) a semi-intensive care unit (SICU), or (*iv*) a regular ICU. As our goal is to create a decision support tool that will be useful in real world clinical settings, we strive to meet the transperency criteria set forth in the TRIPOD Statement [12]: our model is built from publicly-available data [4] rich in clinical biomarkers and the entirety of our code is made publicly available [11]. This enables us to develop a fully transparent and verifiable model which demonstrates a high correlation between several clinical laboratory results and a prediction of criticality. While some of our results match those achieved by other studies, we believe that this congruence between predictive models builds a consensus around which clinical features are truly predictive of critical COVID-19 infections. To this end, we aggregate our results with those of other predictive models to create a robust, consolidated framework for effective COVID-19 patient routing. As far as we are

aware, this work is the first to integrate data from geographically and ethnically diverse COVID-19 patient cohorts.

2 BACKGROUND AND RELATED WORK

COVID-19 is an infection of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), first reported in connection with bushmeat served at a seafood market in Wuhan, Hubei, China [32]. Common symptoms [32] include fever (82.2%), cough (61.7%), fatigue (44.0%), dyspnea (41%), and anorexia (41.0%), though approximately 18% of COVID-19 infections are asymptomatic. The virus spreads rapidly through close contact and by droplets, and there is little evidence of airborne transmission. It's incubation period averages 3-9 days and carriers are contagious approximately 2.5 days before symptoms present; there are multiple studies confirming that asymptomatic carriers are still contagious. Moreover, nearly half of the known COVID-19 infections were transmitted before symptom presentation arose [32].

There are currently two primary methods of diagnosing COVID-19: (i) reverse transcriptase – polymerase chain reaction (RT-PCR) tests [13] and (ii) computed tomography (CT) scans [40]. RT-PCR tests, the preferred method for diagnosis, involves collecting samples from the patient for a lengthy testing process that may not return timely results [33]. RT-PCR tests have been scarce during the COVID-19 pandemic. CT scans, a highly precise x-ray procedure, show potential for more COVID-19 diagnoses. CT chest scans of COVID-19 patients show bilateral ground-glass opacification or consolidation, depending on the stage of the infection. However, the American College of Radiology recommends against the use of CT scanners for COVID-19 diagnosis [1]. There are several proxy diagnostic frameworks, as well as multiple treatment protocols (e.g., [19, 25]).

Though there have been early experimental reports of pharmaceutical therapies (e.g., [5]), there are not currently any approved vaccines or anti-viral treatments for COVID-19.Moreover, a challenge to producing an effective vaccine is that many coronavirus patients do not develop long-term immune responses [26, 29]. Barring the development of an effective vaccine, current research predicts at least two years of COVID-19 waves [22, 28]. Therefore, we believe that patient routing for clinical symptom management will remain a crucial component for treating COVID-19 infection and lowering mortality rates.

2.1 Related Work

COVID-19 is an emerging disease and there are many prototyped tools that will diagnose COVID-19 as well as predicting hospital admission, severity, mortality risk, and hospital stay. Wynants, et al. [35], identify 31 prediction models described in 27 studies: 3 models predicted hospital admission for pneumonia and other events (i.e., proxy diagnoses for COVID-19), 18 models detected COVID-19 infection (13 of which are based on CT scans), and 10 prognostication models to predict mortality or progression of disease severity. They determined that all of the studies suffered from a high risk of overfitting, as well as bias due to non-representative selection of control patients and exclusion of other patients. Moreover, many models were built using data sets with missing data due to the fact that attending physicians do not order a standardized set of laboratory tests.

The majority of work that does not rely on CT scans has come from Chinese researchers [9, 17, 23, 31, 36–38]. Researchers have also built a handful of diagnostic [7] or prognostic [30] models using data from elsewhere in the world. Most of these models focus on predicting which patients are at a high mortality risk, thus inferring which patients will require more drastic interventions. Our predictive model is generally different in that we do not focus on mortality and predict patient prognoses with a finer granularity.

All of the models referenced provide limited descriptive statistics for all features that correlate with prognoses in their models, specifically sample size, median, and interquartile range. We use statistical computation methods described in [24, 34] to estimate mean (μ) and standard deviation (σ) for each feature. Some models give a degree of granularity, separating patients into cohorts (e.g., patients with a mild infection and patients with a serious infection), so we pool the derived μ's and σ's for each feature using calculations described in [6]. Table 1 shows the calculated μ and σ for selected features in each model.

3 MATERIALS AND METHODS

3.1 Data Resources

The Hospital Israelita Albert Einstein (HIAE), located in São Paulo, Brazil has released a data set [4] that provides a variety of information for 5,644 patients between 28 March–03 April, 2020. This information includes 108 features ranging from categorical data to numerical biomedical readings which have been standardized to a mean of 0 an standard deviation of 1. Of the 5,644 patients in the dataset, 558 tested positive for COVID-19; patient routing based on prognosis is given in Table 2.

3.2 Methodologies Employed

While our data set is rich in patient feature information, there are also extraneous and empty features. As such, we first perform a series of cleaning and data manipulation steps to prepare for model construction. Our process for model construction is shown in Figure 1.

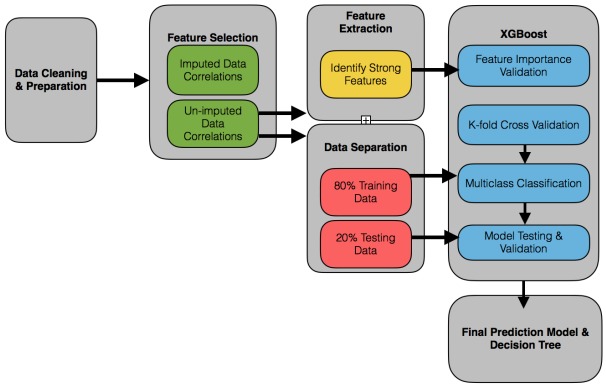

Figure 1: Flowchart of data cleaning processes and approach to model construction.

Table 1: Estimated Descriptive Statistics (μ, σ) for COVID-19 Diagnostic and Prognosis Prediction Models.

Model	n	hs-CRP (mg/L)	LDH (U/L)	ALT (IU/L)	Neutrophil Count (×1, 000/mcL)	Base Excess (mmol/L)
Gong, et al. [17]	189	(15.06, 18.09)	(199.39, 77.98)	(23.18, 14.67)	Not Recorded	Not Recorded
Petrilli, et al. [30]	1,582	(116.16, 63.91)	Not Recorded	(36.30, 21.98)	Not Recorded	Not Recorded
Lu, et al. [23]	75	(27.45, 20.71)	Not Recorded	Not Recorded	(3.81, 2.34)	Not Recorded
Xie, et al. [36]	444	(68.18, 66.42)	(387.16, 204.38)	(29.80, 20.82)	Not Recorded	Not Recorded
Yan, et al. [37]	375	(42.79, 71.52)	(356.92, 295.63)	Not Recorded	(6.66, 6.11)	Not Recorded
Batista, et al. [7]	102	(38.30, 53.90)	Not Recorded	Not Recorded	Not Recorded	Not Recorded

Table 2: COVID-19 Patient routing based on prognosis.

Prognosis/Patient Routing	Count
Mild/Not Admitted	506
Mild/Admitted	36
Severe/Semi-Intensive Care Unit (SICU)	8
Critical/Intensive Care Unit (ICU)	8

3.2.1 Data Cleaning. To construct a criticality prediction model, our data set is first filtered to only include patients who tested positive for SARS-CoV-2 RT-PCR test, reducing the patient sample size from 5,644 to 558 observations. Amongst the feature observations, our data set provides patience admittance information to three levels of care: regular, semi-intensive, and intensive-care units. From these classifications, a criticality measure is developed to represent non-admittance (0), admittance to a regular ward (1), admittance to the semi-intensive unit (2), and admittance to the intensive care unit (3). These 4 classes serve as a proxy for criticality. Through the data cleaning process, missing data becomes evident. Only 6 features are complete and no patient has a full set of feature data. Thus, simply discarding incomplete data is not an option, and missing data must be addressed. One approach involves using means and modes data to complete each feature. Our decision to use the means and modes of the data sets is motivated by a desire to ensure data remains as representative of the sample population as possible. Simply employing global or even country-level population means and modes exposes the model to bias from external factors. Using imputed data in the feature extraction phase, it yields one strongly correlated feature (criticality itself); this is not useful, so we proceed with the unimputed data, employing a model that can handle the missing data issue.

3.2.2 Feature Extraction. Feature extraction is conducted by first evaluating the correlation between criticality and the patient biomarkers. The Pearson correlation coefficient involves obtaining a Least-Squares fit for the data, and yields a value between -1 and 1, with 1 and -1 representing perfect positive or negative relationships, respectively, and 0 indicating no relationship. We are interested in features that are highly correlated with criticality. We construct a ranked list of absolute correlation values for each feature, allowing for easy identification of the strongest features. A correlation list is constructed using unimputed data. This process identifies nine out of the 108 total features with an absolute correlation over 0.4.

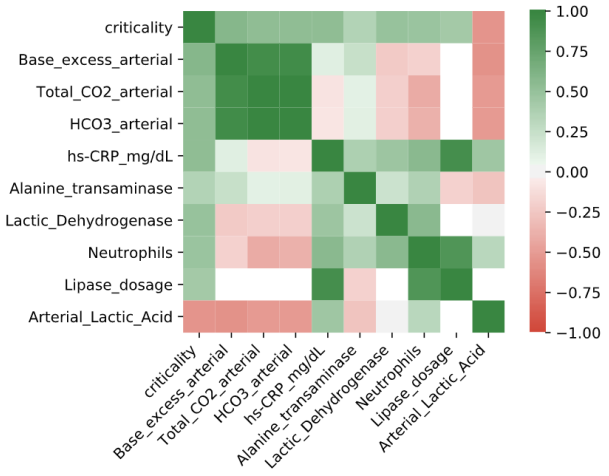

Figure 2: A heat map correlating the most significant clinical biomarkers with criticality.

3.2.3 Feature Selection. Selected features are found from the unimputed data set's correlation list; each feature has an absolute correlation over 0.4 and is presented in Figure 2. The figure illustrates not only the range of correlations between the features and our observation of interest, criticality, but also, the relationship between each feature. This further allows us to reduce the number of features employed, simplifying the model and reducing the risk of overfitting our data set. These features are then plotted against criticality to ensure correlation is indeed strong and not a result of small data sets. Features with few data sets are then removed from the strong selected features list. The remaining strong features and the criticality measure are then inputted into the XGBoost algorithm.

3.2.4 Applying XGBoost. XGBoost [10] is an optimized distributed gradient boosting library that is both efficient and flexible, able to handle large and small data sets alike. This model provides many benefits to the study, namely intelligent missing data handling and parallel tree boosting, which allows for easy visualization of the decision tree employed by the model. These two key features of the algorithm make it the best suited approach when dealing with limited and often incomplete patient data while being interpretable to medical professionals. Additionally, XGBoost has a strong feature selection mechanism and has proven effective by other similarly-conducted COVID-19 studies (e.g., [38]).

Boosting involves training a series of weak models and compensating based on the weaknesses of each preceding series in order to improve model prediction. Unlike most other modeling approaches, boosting relies on several models rather than a single model to create a stronger model ensemble. For this study, a multi-class classification selected as the weak model. XGBoost optimizes the standard boosting algorithm approach model by using gradient boosting. Rather than adjusting weights at each data point in the model, gradient boosting aims to minimize error, or the difference between the prediction value and the true value.

To begin our predictive model construction, the cleaned unimputed data is randomly split into two sets: training data (70%) and testing data (30%). Due to data sparsity issues amongst higher criticality classes (*SICU* and *ICU*), we need a greater split of data for testing. XGBoost is then run on the training data using a Multi-Class Softmax Classifier. This classifier relegates data to classes using a softmax function, a form of the logistic regression that normalizes input values such that the probability distribution sums to 1. We use this probability distribution to reconstruct interval values of each feature and their corresponding criticality levels.

Feature selection is verified using XGBoost's feature importance algorithm, which ranks features based on their importance in the decision making process. The model is then validated using the testing data set. In order to cross-validate the results, we employ K-fold validation with 10 folds. K-fold cross validation is a procedure that is often used to estimate the skill of a model on new data.The results from this process are presented and discussed in the following section.

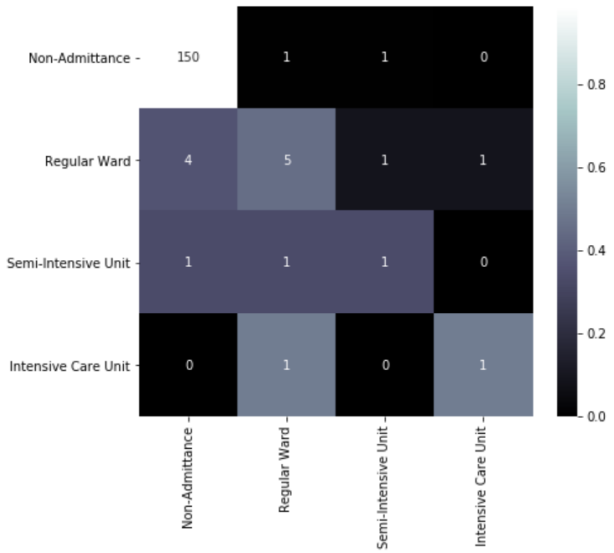

Figure 3: Confusion matrix based on our testing data set.

4 RESULTS AND DISCUSSION

4.1 A Predictive Model Built on Fully Transparent Data

Our prognosis model assumes that a patient has already tested positive for COVID-19 by one of the diagnosis methods described in Section 2. Our model (Table 3) predicts which clinical routing is most appropriate to the patient based on five features that correlate strongly with critical infections. Mild COVID-19 infections where the patient is not admitted to the hospital are characterized by low levels of hs-CRP and ALT, as well as a relatively normal neutrophil count. ICU admission is appropriate for the most critical infections, which are characterized by higher hs-CRP and low base excess levels.

Predicting which COVID-19 patients should be admitted to the ICU and which patients should be sent home is straight forward; predicting whether patients should be admitted to the hospital's general ward or a SICU is more nuanced. Where our model defines one set each of criteria for ICU admittance and no hospital admittance; admittance to the general ward and SICU each have three sets of criteria. Mild COVID-19 infections where general hospital admission is predicted may be characterizes by (*i*) low hs-CRP levels and a higher neutrophil count, (*ii*) moderate hs-CRP and ALT levels, as well as a normal neutrophil count, and (*iii*) moderate hs-CRP, lower LDH levels, and a reduced neutrophil count. Severe infections, for which SICU admittance is appropriate, are characterized by (*i*) moderate hs-CRP, lower LDH levels, and a reduced neutrophil count, (*ii*) moderate hs-CRP and LDH levels, and (*iii*) higher hs-CRP and base excess levels.

We chose a 70%/30% train-test split to ensure that our model was trained on all cases. The overwhelming majority of our test set was given a mild prognosis and, rather than admittance to the hospital, were instructed to self-quarantine at home. On the other hand, only 5 patients were admitted to either SICU or ICU. Specifically, our model classified Mild/Not Admitted for 152 members of the test set while 11 members received a prognosis of Mild/Admitted. Our model classified 3 patients as Severe/SICU and 2 patients as Critical/ICU; the model does not predict mortality. Our model's performance is presented in Table 4.

In order to validate the findings, we run a stratified 10-fold cross validation approach to the model. In the end, we find a mean error of 0.0824, which is lower than the error of a single testing and training fold. This technique is essential in enforcing class distribution with multiple rounds of training and testing and can partially resolve the issue of class imbalance. Overall, the low error also suggests that the model is not overfitting the data, and thus results should be replicable using a different data set.

Of the 152 patients in the test set classified as Mild/Not Admitted, one patient had a ground truth prognosis of Mild/Admitted and another patient had a ground truth of Severe/SICU. Our model predicted 11 test set patient prognoses as Mild/Admitted, when only 5 of those patients were actually admitted to the hospital; 4 patients were not admitted while the SICU and ICU received one each of these patients. Our model classified 3 test set patients as Severe/SICU; however, only one of those patients was admitted to the SICU; 1 patient was admitted to the regular hospital ward and the other was not admitted. Of the 2 patients classified as Critical/ICU, 1 patient was actually admitted to the regular hospital ward. Our confusion matrix is in Figure 3.

Table 3: COVID-19 Prognosis Classifier and Patient Routing Model.

Mild/Not Admitted	Mild/Admitted	Severe/SICU	Critical/ICU
hs-CRP < -0.46 & Neutrophil count < 0.95 & ALT < -0.37	hs-CRP < -0.46 & Neutrophil count ≥ 0.95	$-0.46 \leq$ hs-CRP < 0.07 & LDH ≥ -0.82	hs-CRP ≥ 0.07 & Base Excess < 0.27
	$-0.46 \leq$ hs-CRP < 0.07 & Neutrophil count < 0.95 & ALT ≥ -0.37	$-0.46 \leq$ hs-CRP < 0.07 & LDH < -0.82 & Neutrophil count < -0.11	
	$-0.46 \leq$ hs-CRP < 0.07 & LDH < -0.82 & Neutrophil count ≥ -0.11	hs-CRP ≥ 0.07 & Base Excess ≥ 0.27	

Table 4: Model performance on testing data

	Precision	Recall	F1-score	Support
Mild/Not Admitted	0.97	0.99	0.98	152
Mild/Admitted	0.62	0.45	0.53	11
Severe/SICU	0.33	0.33	0.33	3
Critical/ICU	0.50	0.50	0.50	2
Accuracy			0.93	168
Macro Average	0.61	0.57	0.58	168
Weighted Average	0.93	0.93	0.93	168

4.2 Discussion

Our analysis shows 9 clinical biomarkers with a significant correlation to COVID-19 patient criticality (Figure 2). Using the XGBoost Library's Feature Importance capability, we are able to further pair this down to 6 features (Figure 4): hs-CRP (mg/dL), Neutrophil count, ALT (IU/L), LDH (U/L), Base excess (arterial blood gas analysis), and Total CO2 (arterial blood gas analysis). Of these features, hs-CRP and LDH have been shown to be heavily correlated with criticality in other studies [17, 38]. hs-CRP is an inflammatory biomarker [18] that is important in predicting which cases will be require SICU or ICU admittance. A normal measurement for hs-CRP is 1 mg/dL and Petrilli, et. al. [30], found that early elevation of hs-CRP is one of the strongest predictors for mechanical ventilation or mortality. The normal range for LDH in adults is 100-190 (U/L) and elevated LDH is indicative of damaged of diseased tissues [16]. ALT is an enzyme found mostly in the liver; the normal range for ALT is less than 40 IU/L and elevated ALT (i.e., >1,000 IU/L) is a sign of liver damage [20]. The normal level of neutrophils, a type of white blood cell, is 1,500-8,000 neutrophils/mcL ; a low neutrophil count may make a person susceptible to an increased risk of infection [15] while a high neutrophil count is indicative of an ongoing immune response. Base excess is a measure of the acid-base balance in the blood with a standard reading of 0 mmol/L and a normal range of ±3; elevated levels may indicate respiratory or metabolic acidosis [8]. The models that we surveyed do not typically consider base excess in their set of features, and the reason that our model utilizes it is likely due to the irregularity of laboratory tests ordered for patients that contributed to our training data. Nonetheless, there is literature that records base excess levels in COVID-19 patients [39]. Furthermore, there are published treatment recommendations that provide guidelines for low base excess levels [27].

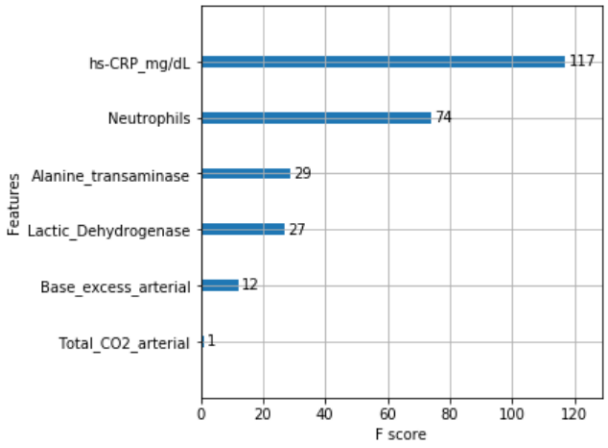

Figure 4: Feature importance by F-score.

4.2.1 Limitations. Data unavailability and feature incompleteness are key limitations, both in this and other related studies. Given that COVID-19 remains an evolving situation, many countries and health institutions are slow to release data. Available data sets tend to be small and limited in scope. Thus these smaller sets lead to an issue of class under-representation. The low instances of *SICU* and *ICU* cases in the set (both only have 8 observations while the *Not-Admitted* has 506 observations) undermines the model's ability to sufficiently learn each class. In other words, XGBoost does not have enough cases of *SICU* and *ICU* patients to properly learn and later predict these cases. A milder case of under-representation is observed with the *Mild/Admitted* class, which has a 0.62 precision rate and only 36 observations. Due to class under-representation, we witness a lower precision and recall rate for these three classes. In Figure 3, our model's confusion matrix is well-trained to diagnose Non-Admittance cases, but as criticality increases it becomes less accurate. Due to the overrepresentation of these Non-Admittance classes, there is a slight bias towards lower criticality levels. This can be observed in the tendency to diagnose a lower criticality level as true criticality increases.

Failure to properly classify would prove costly in the case of severe and critical patients. Patients who present high criticality

would not be identified, which could affect the level of treatment they receive. There are several means of addressing imbalanced data such as adjusting performance metrics. We employ this technique by using precision and recall as valuable measures over accuracy when assessing the strength of the model. The confusion matrix also clearly demonstrates the value of the model and its slight bias towards lower criticality. We could use a Synthetic Minority Over-sampling Technique, or SMOTE, to create more synthetic data on which to run our model. However, SMOTE does not increase accuracy, rather it increases recall at the cost of precision. SMOTE is also better suited for classes that are rare. From global data, we know higher levels of criticality are not rare. COVID-19 has high mortality rates, thus these severe classes are likely prevalent; if we can access a larger set, we may account for these cases more accurately. Overall, we find the best way to address the issue of imbalanced data is through greater data collection. Pooling patient clinical information and outcomes across a country or region would yield a larger and more representative sample to train the model on. As countries and hospitals begin to release more anonymous patient clinical data, a more robust predictive model may be developed.

Furthermore, as mentioned earlier in the study, many healthcare providers order a variety of tests, not always ordering the same test set for every patient. As such, we see many clinical features are left incomplete, resulting in fewer complete patient information sets. While XGBoost is adept in handling missing information, we believe a more complete data set would yield a stronger predictive model.

Despite these limitations, we still find hs-CRP and LDH as features that are heavily correlated with criticality, an insight that is validated by other studies. One of the key objectives of this study, aside from constructing a predictor, is to test the generalizability and repeatability of other related works. The identification of the same clinical features across these studies, each constructed from completely independent data sets, strengthens our findings and suggests pooling data may improve model accuracy. Overall, we confirm the results of other studies and remain confident in this approach to predicting criticality in COVID-19 patients.

5 A CONSOLIDATED FRAMEWORK FOR COVID-19 PATIENT PROGNOSIS PREDICTION AND ROUTING

Table 5: Aggregated Global Descriptive Statistics $(\bar{\mu}, \bar{\sigma})$ of Features for Predicting Criticality in COVID-19 Patients.

Feature	n	$(\bar{\mu}, \bar{\sigma})$
hs-CRP (mg/L)	2,767	(86.34, 62.29)
LDH (U/L)	1,008	(340.70, 228.20)
ALT (IU/L)	2,215	(33.88, 21.22)
Neutrophil Count ($\times 1,000$/mcL)	450	(6.19, 5.66)
Base Excess (mmol/L)	58	(1.30, 1.53)

COVID-19 prognosis models have been proposed based on small studies from individual hospitals, and many of these from China. Individual models, built on very localized populations, may not be generalizable. Due to privacy regulations, published models do not provide raw data sets, but may provide descriptive statistics, (e.g., sample size, median, and interquartile ranges). We were able to estimate means and standard deviation of highly correlated features for each model (Table 1). We pool [6] by feature across all models to create aggregated global descriptive statistics by feature, with an aggregated global mean estimate ($\bar{\mu}$) and an aggregated global standard deviation estimate ($\bar{\sigma}$) (Table 5).

Our model (Table 3) was trained on a data set that was standardized to ($\mu = 0, \sigma = 1$) for all features as a measure to protect patient privacy, but negates its applicability beyond the HIAE. Moreover, hospital administration will not release raw data nor provide descriptive statistics [4]. As our goal in this work is to create a predictive model which will provide a useful decision support capability to clinicians, we feed our aggregated global descriptive statistics into our model (Table 6). To the best of our knowledge, this is the first attempt to integrate geographically and demographically diverse data into a predictive prognosis model for COVID-19.

6 CONCLUSIONS AND FUTURE WORK

We have presented a fully transparent prognosis prediction model for COVID-19 patients. Unlike many of the other predictive models that have been produced around the world, we use a publicly available data set [4] to build our model and have made the entirety of our code publicly available [11]. After presenting and analyzing our results and model, we aggregate our model with a number of predictive models from other researchers to create a consolidated framework for COVID-19 prognosis and patient routing. There are no recommended anti-viral therapies or approved vaccines for COVID-19 at this time, therefore we believe that clinical symptom management will remain the primary treatment for the disease. Clinicians have limited available resources, therefore predictive models which support clinical resource and personnel allocation provide a critical capability that may lead to improved patient survival.

The major limitation of our current model is that we trained it on a single publicly available data set from a single hospital. We would very much like to retrain our model with a larger data set that includes clinical biomarkers from geographically and demographically diverse COVID-19 patients. This can easily be accomplished as more hospitals release COVID-19 patient data sets that have been processed to comply with patient privacy regulations. Additionally, our model independently prioritized the same clinical biomarkers as other similar studies. This gives us greater confidence in combining data sets across countries and healthcare institutions to develop a more robust model. As country or institution-specific factors do not appear to bias feature selection, there is great opportunity to expand criticality modelling as more data sets become available.

REFERENCES

[1] Acr recommendations for the use of chest radiography and computed tomography (ct) for suspected covid-19 infection. Tech. rep., American College of Radiology, March 22, 2020.
[2] Coronavirus resource center. https://coronavirus.jhu.edu/, 2020.
[3] Covid act now. https://covidactnow.org/, 2020.
[4] Diagnosis of covid-19 and its clinical spectrum: Ai and data science supporting clinical decisions (from 28th mar to 3rd apr). https://www.kaggle.com/einsteindata4u/covid19#dataset.xlsx, 2020.

Table 6: A Global Predictive COVID-19 Patient Prognosis and Routing Model

Mild/Not Admitted	Mild/Admitted	Severe/SICU	Critical/ICU
hs-CRP < 57.69 & Neutrophil count < 11.57 & ALT < 26.03	hs-CRP < 57.69 & Neutrophil count ≥ 11.57	$57.69 \leq$ hs-CRP < 90.70 & LDH ≥ 153.58	hs-CRP ≥ 90.70 & Base Excess < 1.71
	$57.69 \leq$ hs-CRP < 90.70 & Neutrophil count < 11.57 & ALT ≥ 26.03	$57.69 \leq$ hs-CRP < 90.70 & LDH < 153.58 & Neutrophil count < -0.11	
	$57.69 \leq$ hs-CRP < 90.70 & LDH < 153.58 & Neutrophil count ≥ 5.57	hs-CRP ≥ 90.70 & Base Excess ≥ 1.71	

[5] Arshad, S., Kilgore, P., Chaudhry, Z. S., Jacobsen, G., Wang, D. D., Huitsing, K., Brar, I., Alangaden, G. J., Ramesh, M. S., McKinnon, J. E., et al. Treatment with hydroxychloroquine, azithromycin, and combination in patients hospitalized with covid-19. *International Journal of Infectious Diseases* (2020).

[6] Arsham, H. Statistical thinking for managerial decisions: Pooling the sampling estimates for mean, variance, and standard deviation, 2015. https://home.ubalt.edu/ntsbarsh/business-stat/opre504.htm#rTwoIndTest.

[7] Batista, A. F. d. M., Miraglia, J. L., Donato, T. H. R., and Chiavegatto Filho, A. D. P. Covid-19 diagnosis prediction in emergency care patients: a machine learning approach. *medRxiv* (2020).

[8] Burns, G. P. Arterial blood gases made easy. *Clinical medicine (London, England) 14,1* (2014).

[9] Caramelo, F., Ferreira, N., and Oliveiros, B. Estimation of risk factors for covid-19 mortality - preliminary results. *medRxiv* (2020).

[10] Chen, T., and Guestrin, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (2016), pp. 785–794.

[11] Chikkula, A. https://github.com/akchikkula111/cs8920S/tree/gh-pages, 2020.

[12] Collins, G. S., Reitsma, J. B., Altman, D. G., and Moons, K. G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): The tripod statement. *Annals of Internal Medicine 162*, 1 (2015), 55–63. PMID: 25560714.

[13] Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M. L., et al. Detection of 2019 novel coronavirus (2019-ncov) by real-time rt-pcr. *Eurosurveillance 25*, 3 (2020), 2000045.

[14] Dixon, J. Why am i, as a geriatric medicine fellow with symptoms, unable to get tested for covid-19 while politicians, oil executives, and nba players are? *Journal of the American Geriatrics Society* (2020).

[15] Fetterman, A., and Haldeman-Englert, C. Health encyclopedia – white blood cell count. Tech. rep., University of Rochester Medical Center, 2020.

[16] Fraser, M., and Haldeman-Englert, C. Health encyclopedia – lactic acid dehydrogenase (blood). Tech. rep., University of Rochester Medical Center, 2020.

[17] Gong, J., Ou, J., Qiu, X., Jie, Y., Chen, Y., Yuan, L., Cao, J., Tan, M., Xu, W., Zheng, F., Shi, Y., and Hu, B. A tool to early predict severe 2019-novel coronavirus pneumonia (covid-19) : A multicenter study using the risk nomogram in wuhan and guangdong, china. *medRxiv* (2020).

[18] Greco, F., and Sather, R. Health encyclopedia – c-reactive protein (blood). Tech. rep., University of Rochester Medical Center, 2020.

[19] Gulick, R. M., Lane, H. C., Masur, H., Pau, A. K., Aberg, J., Adimora, A., Baker, J., Bedimo, R., Cantrill, S., Collier, A. C., Coopersmith, C., Daar, E., Davis, S. L., Dzierba, A. L., Evans, L., Gandhi, R., Glidden, D., Grund, B., Hardy, E. J., Hughes, B. L., Johnson, S., Keller, M. J., Kim, A., Lennox, J. L., Levy, M. M., Martin, G., Naggie, S., Simpson, S. Q., Swindells, S., Tebas, P., Tien, P., Waghmare, A. A., and Wilson, K. C. Coronavirus disease 2019 (covid-19) treatment guidelines. Tech. rep., National Institutes of Health, 2020.

[20] Haldeman-Englert, C., Trevino, H., and Turley, K. Health encyclopedia – alt. Tech. rep., University of Rochester Medical Center, 2020.

[21] Holshue, M. L., DeBolt, C., Lindquist, S., Lofy, K. H., Wiesman, J., Bruce, H., Spitters, C., Ericson, K., Wilkerson, S., Tural, A., et al. First case of 2019 novel coronavirus in the united states. *New England Journal of Medicine* (2020).

[22] Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H., and Lipsitch, M. Projecting the transmission dynamics of sars-cov-2 through the postpandemic period. *Science* (2020).

[23] Lu, J., Hu, S., Fan, R., Liu, Z., Yin, X., Wang, Q., Lv, Q., Cai, Z., Li, H., Hu, Y., et al. Acp risk grade: a simple mortality index for patients with confirmed or suspected severe acute respiratory syndrome coronavirus 2 disease (covid-19) during the early stage of outbreak in wuhan, china. *Preprints with The Lancet* (2020).

[24] Luo, D., Wan, X., Liu, J., and Tong, T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Statistical methods in medical research 27*, 6 (2018), 1785–1805.

[25] Marik, P. E. Evms critical care covid-19 management protocol. Tech. rep., Eastern Virginia Medical School, April 20, 2020.

[26] Mo, H., Zeng, G., Ren, X., Li, H., Ke, C., Tan, Y., Cai, C., Lai, K., Chen, R., Chan-Yeung, M., et al. Longitudinal profile of antibodies against sars-coronavirus in sars patients and their clinical significance. *Respirology 11*, 1 (2006), 49–53.

[27] Mojoli, F., Mongodi, S., Orlando, A., Arisi, E., Pozzi, M., Civardi, L., Tavazzi, G., Baldanti, F., Bruno, R., and Iotti, G. A. Our recommendations for acute management of covid-19, 2020.

[28] Moore, K. A., Lipsitch, M., Barry, J. M., and Osterholm, M. T. Covid-19: The cidrap viewpoint. Tech. rep., Center for Infectious Disease Research and Policy (CIDRAP), University of Minnesota, April 30, 2020.

[29] Payne, D. C., Iblan, I., Rha, B., Alqasrawi, S., Haddadin, A., Al Nsour, M., Alsanouri, T., Ali, S. S., Harcourt, J., Miao, C., et al. Persistence of antibodies against middle east respiratory syndrome coronavirus. *Emerging infectious diseases 22*, 10 (2016), 1824.

[30] Petrilli, C. M., Jones, S. A., Yang, J., Rajagopalan, H., O'Donnell, L. F., Chernyak, Y., Tobin, K., Cerfolio, R. J., Francois, F., and Horwitz, L. I. Factors associated with hospitalization and critical illness among 4,103 patients with covid-19 disease in new york city.

[31] Shi, Y., Yu, X., Zhao, H., Wang, H., Zhao, R., and Sheng, J. Host susceptibility to severe covid-19 and establishment of a host risk score: findings of 487 cases outside wuhan. *Critical Care 24*, 1 (2020), 1–4.

[32] Siordia, J. A. Epidemiology and clinical features of covid-19: A review of current literature. *Journal of Clinical Virology 127* (2020), 104357.

[33] Strickler, L., and Kaplan, A. Private labs do 85 percent of u.s. covid-19 tests but still struggle with backlogs, shortages. https://www.nbcnews.com/health/health-news/private-labs-do-85-percent-u-s-covid-19-tests-n1177866, April 8, 2020.

[34] Wan, X., Wang, W., Liu, J., and Tong, T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC medical research methodology 14*, 1 (2014), 135.

[35] Wynants, L., Van Calster, B., Bonten, M. M. J., Collins, G. S., Debray, T. P. A., De Vos, M., Haller, M. C., Heinze, G., Moons, K. G. M., Riley, R. D., Schuit, E., Smits, L. J. M., Snell, K. I. E., Steyerberg, E. W., Wallisch, C., and van Smeden, M. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ 369* (2020).

[36] Xie, J., Hungerford, D., Chen, H., Abrams, S. T., Li, S., Wang, G., Wang, Y., Kang, H., Bonnett, L., Zheng, R., et al. Development and external validation of a prognostic multivariable model on admission for hospitalized patients with covid-19. *Preprints with The Lancet* (2020).

[37] Yan, L., Zhang, H.-T., Goncalves, J., Xiao, Y., Wang, M., Guo, Y., Sun, C., Tang, X., Jin, L., Zhang, M., Huang, X., Xiao, Y., Cao, H., Chen, Y., Ren, T., Wang, F., Xiao, Y., Huang, S., Tan, X., Huang, N., Jiao, B., Zhang, Y., Luo, A., Mombaerts, L., Jin, J., Cao, Z., Li, S., Xu, H., and Yuan, Y. A machine learning-based model for survival prediction in patients with severe covid-19 infection. *medRxiv* (2020).

[38] Yan, L., Zhang, H.-T., Xiao, Y., Wang, M., Sun, C., Liang, J., Li, S., Zhang, M., Guo, Y., Xiao, Y., Tang, X., Cao, H., Tan, X., Huang, N., Jiao, B., Luo, A., Cao, Z., Xu, H., and Yuan, Y. Prediction of criticality in patients with severe covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in wuhan. *medRxiv* (2020).

[39] Yang, P., Ding, Y., Xu, Z., Pu, R., Li, P., Yan, J., Liu, J., Meng, F., Huang, L., Shi, L., et al. Epidemiological and clinical features of covid-19 patients with and without pneumonia in beijing, china. *Medrxiv* (2020).

[40] Yoon, S. H., Lee, K. H., Kim, J. Y., Lee, Y. K., Ko, H., Kim, K. H., Park, C. M., and Kim, Y.-H. Chest radiographic and ct findings of the 2019 novel coronavirus disease (covid-19): analysis of nine patients treated in korea. *Korean journal of radiology 21*, 4 (2020), 494–500.