

# Machine learning–aided algorithm design for prediction of severity from clinical, demographic, biochemical and immunological parameters: Our COVID-19 experience from the pandemic

Suchitra Kumari<sup>1</sup>, Swagata Tripathy<sup>2</sup>, Saurav Nayak<sup>1</sup>,  
Aishvarya S. Rajasimman<sup>3</sup>

<sup>1</sup>Department of Biochemistry, All India Institute of Medical Sciences, Bhubaneswar, Odisha, India, <sup>2</sup>Department of Anesthesiology, All India Institute of Medical Sciences, Bhubaneswar, Odisha, India, <sup>3</sup>Department of Radiodiagnosis, All India Institute of Medical Sciences, Bhubaneswar, Odisha, India

## ABSTRACT

**Background:** The severity of laboratory and imaging finding was found to be inconsistent with clinical symptoms in COVID-19 patients, thereby increasing casualties. As compared to conventional biomarkers, machine learning algorithms can learn nonlinear and complex interactions and thus improve prediction accuracy. This study aimed at evaluating role of biochemical and immunological parameters–based machine learning algorithms for severity indexing in COVID-19. **Methods:** Laboratory biochemical results of 5715 COVID-19 patients were mined from electronic records including 509 admitted in COVID-19 ICU. Random Forest Classifier (RFC), Support Vector Machine (SVM), Naive Bayesian Classifier (NBC) and K-Nearest Neighbours (KNN) classifier models were used. Lasso regression helped in identifying the most influential parameter. A decision tree was made for subdivided data set, based on randomization. **Results:** Accuracy of SVM was highest with 94.18% and RFC with 94.04%. SVM had highest PPV (1.00), and NBC had highest NPV (0.95). QUEST modelling ignored age, urea and total protein, and only C-reactive protein and lactate dehydrogenase were considered to be a part of decision-tree algorithm. The overall percentage of correct classification was 78.31% in the overall algorithm with a sensitivity of 87.95% and an AUC of 0.747. **Conclusion:** C-reactive protein and lactate dehydrogenase being routinely performed tests in clinical laboratories in peripheral setups, this algorithm could be an effective predictive tool. SVM and RFC models showed significant accuracy in predicting COVID-19 severity and could be useful for future pandemics.

**Keywords:** Algorithm, biochemical parameter, COVID-19 severity, machine learning

## Introduction

SARS-CoV-2 infection seems to have a broad clinical spectrum, encompassing asymptomatic infection, mild upper respiratory

tract illness and severe viral pneumonia with respiratory failure and even death.<sup>[1]</sup> However, in most cases of COVID-19, clinical manifestations are typically mild, and these patients may have reported no dyspnoea, no significant increase in respiratory rate and no respiratory distress. Even when shortness of breath, loss of appetite, confusion, persistent pain or pressure in the chest and a high temperature and other symptoms of severity were seen in some patients, the severity of laboratory and imaging

**Address for correspondence:** Dr. Suchitra Kumari,  
Department of Biochemistry, All India Institute of Medical  
Sciences, Bhubaneswar, Sijua, Patrapada, Odisha, PIN - 751 019, India.  
E-mail: suchitrakumari76@gmail.com

Received: 30-10-2023

Revised: 11-12-2023

Accepted: 12-12-2023

Published: 24-05-2024

### Access this article online

#### Quick Response Code:



**Website:**  
<http://journals.lww.com/JFMPC>

**DOI:**  
10.4103/jfmpe.jfmpe\_1752\_23

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow\_reprints@wolterskluwer.com

**How to cite this article:** Kumari S, Tripathy S, Nayak S, Rajasimman AS. Machine learning–aided algorithm design for prediction of severity from clinical, demographic, biochemical and immunological parameters: Our COVID-19 experience from the pandemic. J Family Med Prim Care 2024;13:1937-43.

findings was found to be inconsistent with clinical symptoms.<sup>[2,3]</sup> Elevated cytokine profile has been the immediate go-to testing in cases of determining severity. Factors such as interleukins, interferon gamma, fibroblast growth factor and such related immunological parameters that characterise a cytokine storm have been associated with severity of COVID-19.<sup>[4-6]</sup> Recent studies have also determined the effect of routine biochemical investigations such as urea, creatinine, liver enzymes, lactate dehydrogenase (LDH), ferritin and C-reactive protein (CRP) to be associated with severity induced because of COVID-19.<sup>[7-9]</sup>

The guidelines for the diagnosis and treatment of novel coronavirus pneumonia include specific criteria for severe COVID-19, such as respiratory rate, haemoglobin oxygen saturation (SaO<sub>2</sub>) and oxygenation index (PaO<sub>2</sub>/FiO<sub>2</sub>).<sup>[10]</sup> These criteria, however, are susceptible to subjective and objective factors, which may result in a slower diagnostic time and the possibility of misdiagnosing severe COVID-19.<sup>[11]</sup> With the prevailing hospital infrastructure and constraint of limited ICU facilities, it is rational to search for a predictive marker that can accurately diagnose severe COVID-19 patients. Prediction models estimate the risk of infection that has been developed and help in assisting medical staff, timely triaging of patients, especially in set-ups with limited healthcare resources.<sup>[12,13]</sup> Primary health care plays a pivotal role in reflecting the country's response to a pandemic. Primary care physicians identify and triage suspected COVID-19 cases requiring ICU services, thereby facilitating an early diagnosis and reducing the morbidity and mortality. Provision of a digital predictive risk assessment tool at primary healthcare setups will reduce the workload of primary care providers and family physicians and will improve the patient care.

Conventional prediction models use linear regression and are flexible in prediction accuracy. With the advancement in data science, machine learning (ML) algorithms have gained more popularity because they can learn the nonlinear and complex interactions between the dependent variables and thus can improve the prediction accuracy drastically in clinical researches. Earlier studies have found the usefulness of ML models to predict scenarios ranging from the risk of developing cardiovascular disease to stratifying early onset neonatal sepsis.<sup>[14,15]</sup> This study was conducted using Random Forest Classifier (RFC), Support Vector Machine (SVM), Naive Bayesian Classifier (NBC) and K-Nearest Neighbours (KNN) classifier models to predict the severity of COVID-19 infection in ICU admission that will aid in the processing of an efficient algorithm for primary healthcare practitioners.

## Material and Methods

*Study Setup:* This study was conducted in the Department of Biochemistry of All India Institute of Medical Sciences (AIIMS), Bhubaneswar. It was a retrospective comparative study of the detailed demographic and clinical data of COVID-19-affected in-patients. This study included laboratory data of patients

diagnosed with COVID-19 and admitted to an in-patient department (IPD) and ICU of AIIMS, Bhubaneswar, between July 2020 and May 2022. Most outpatients and inpatients of the institute were residents of eastern Indian states, Odisha and West Bengal.

*Biochemical Parameters:* The serum high-sensitivity C-reactive protein, ferritin, IL-6, lactate dehydrogenase, urea, creatinine, uric acid, sodium, potassium, chloride and liver function tests (total bilirubin, direct bilirubin, AST, ALT, ALP, total protein and albumin) were the parameters performed in the clinical biochemistry laboratory of AIIMS, Bhubaneswar, for the COVID-19 patients. Biochemical parameters were estimated by fully automated chemistry analyser (Beckman Coulter 5800; Beckman Coulter Inc., Brea, California) using system compatible packs. Serum ferritin and IL6 were measured by Siemens Advia XP Chemiluminescence Immunoassay analyser (Siemens Healthcare Diagnostics, Tarrytown, New York). The study was approved by the Ethics Committee of AIIMS, Bhubaneswar (T/IM-NF/Biochem/21/162 dated 28<sup>th</sup> Feb 2022).

*Data Acquisition and Processing:* The data were duly extracted from the stored electronic records. All the patients who had all the required data for the aforementioned tests without any missing data were included in the study. All the participants with COVID-19 were divided into two groups: non-intensive care unit (non-ICU, those admitted to COVID-19 wards) and ICU (those admitted to ICU). Biochemical test results with missing serum hs-CRP, ferritin, IL-6, LDH and liver enzymes and results beyond the linearity limit of the specific assays used were excluded. If the same patient had undergone biochemical tests multiple times during the study period, only the first result was included in the data set.

*Developing Machine Learning Model:* After the review of the laboratory data, 5715 COVID-19 patients' biochemical test results were selected for the analysis by ML models. Four ML models were used for classification, Random Forest Classifier (RFC), Support Vector Machine (SVM), Naive Bayesian Classifier (NBC) and K-Nearest Neighbours (KNN). The method used for comparison was by estimating accuracy, recall (sensitivity), recall (specificity), positive predictive value (PPV) and negative predictive value (NPV). Of the total records in the data set, 80% were used as the training set, and the rest 20% were used as the testing set. Lasso regression was used to identify the most influential parameters affecting the predictive potential of the model. A decision tree was set up for a sub-divided data set containing approximately equal numbers of records from each group based on randomization. This decision tree was based on the data of lasso regression and was optimized for influential variables only. The graphical abstract is added as supplementary Figure 1.

*Statistical Analysis:* Statistical analysis was performed using IBM SPSS v26.0, XLSTAT v2022.1 and JASP v0.16. In all analysis, a P value of less than 0.05 was considered significant.

## Results

Our study had a total of 5715 patients of which 509 patients were the ones who required admission in ICU. The 4 ML models used 80% of the data as the training set and 20% as the validation set. The accuracy of SVM was the highest with 94.18% followed by RFC, which had a value of 94.04%. SVM also had the highest PPV (1.00), whereas naive Bayesian classification had the highest NPV (0.95). The model performance data have been tabulated in Table 1.

The PPV was determined to be low because of the class imbalance caused by the large disparity in the number of patients in the groups. Therefore, a sub-division of the non-ICU group was created with equal number of patients as in the ICU group by randomization. In this sub-division, Random Forest had the highest accuracy, PPV and NPV. The summary of the model performance is tabulated in Table 2.

A lasso regression was performed to identify the parameters that affect the predictive power the highest. The most influential parameters with non-zero coefficients were selected [Figure 1]. Models were created using only these parameters (age, CRP, urea, total protein and LDH) in a Random Forest classification as that was found to be the most superior ML algorithm in this case. The model parameters have been tabulated in Table 3. The receiver operating characteristic curve (ROC) curve for the RFC is shown in Figure 2.

The entire data set was reduced to a randomized sub-divided data set of 1009 records (ICU = 509 and non-ICU = 500).

**Table 1: Comparison of Machine Learning models' performance**

	RFC	SVM	NBC	KNN
Accuracy	94.04%	94.18%	87.84%	93.84%
For the positive class, that is group specific for "1" <sup>a</sup>				
Precision (PPV)	0.45	1.00	0.14	0.31
Recall (sensitivity)	0.04	0.02	0.20	0.03
For the negative class, that is group specific for "0" <sup>a</sup>				
Precision (NPV)	0.94	0.94	0.95	0.94
Recall (specificity)	1.00	1.00	0.92	1.00

1<sup>a</sup>: Patients admitted in ICU, 0<sup>a</sup>: Patients not admitted in ICU

**Table 2: Model performance of equally distributed sub-division of patients**

	RFC	SVM	NBC	KNN
Accuracy	75.19%	72.13%	57.63%	66.79%
For the positive class, that is group specific for "1" <sup>a</sup>				
Precision (PPV)	0.76	0.73	0.70	0.69
Recall (sensitivity)	0.75	0.71	0.29	0.63
For the negative class, that is group specific for "0" <sup>a</sup>				
Precision (NPV)	0.75	0.71	0.54	0.65
Recall (specificity)	0.75	0.74	0.88	0.71

1<sup>a</sup>: Patients admitted in ICU, 0<sup>a</sup>: Patients not admitted in ICU

A decision-tree analysis was performed, so that an effective classification algorithm can be generated. The QUEST modelling ignored age, urea and total protein, and only CRP and LDH were considered to be part of the algorithm. The decision tree for the training set and testing set has been shown in Figures 3 and 4, respectively.

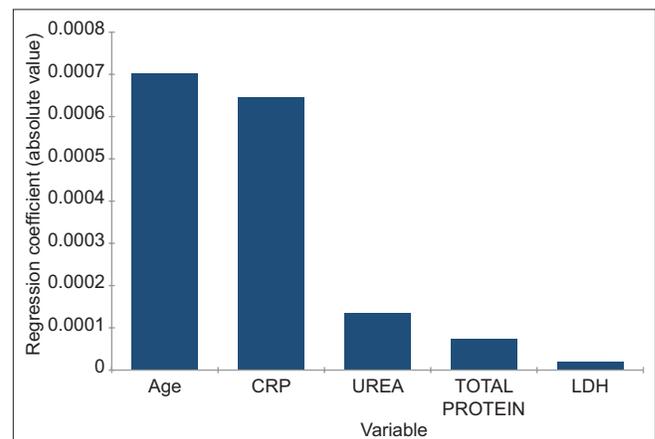
The overall percentage of correct classification was 78.31% in the overall algorithm (73.46% in the training set and 65.66% in the testing set). The algorithm can be simplified to the following:

1. If CRP >140.8 mg/dL—classify as severe (thus needing ICU admission);
2. Else, If CRP < or = 140.8, then check for LDH. If LDH > 542.10 U/L—classify as severe;
3. Classify all remaining patients as not requiring ICU admission; to be kept under observation.

This simplified algorithm increases the sensitivity, and NPV thus can be a better screening tool for early severity detection in case of patients. The overall classification strength has been tabulated in Table 4.

## Discussion

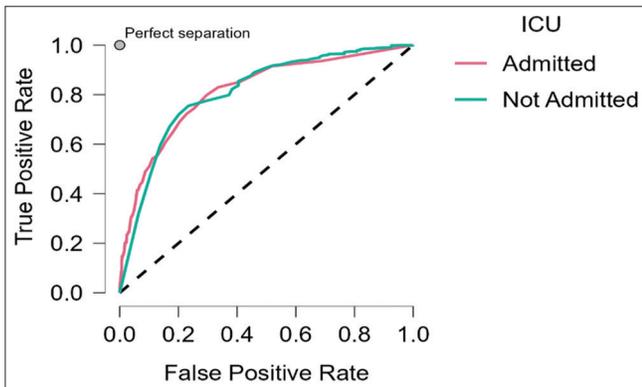
Machine learning prediction models combining several features to estimate the risk of COVID-19 infection can mitigate the burden on healthcare system by facilitating timely triage practices in limited healthcare resources. Several researchers have attempted to use ML to predict the severity of COVID-19. Few studies demonstrated the role of artificial intelligence in detecting clusters of cases and predicting where this virus would affect in the future, as well as predicting the evolution of illness.<sup>[16,17]</sup> This study was aimed at understanding the role of biochemical and immunological parameters in routine laboratory use as a marker for severity of COVID-19. It is one of the few studies pioneering the use of ML algorithm for severity indexing in patients affected by the pandemic. Our study showed that there is significant predictive efficacy of the various ML algorithms that have performed with accuracy more than 90%. A similar



**Figure 1: Regression coefficient–based most influential parameters**

study in American population identified risk factors pertaining to clinical parameters, rather than laboratory-based parameters, were significantly able to classify patients based on severity.<sup>[18]</sup> Mariam Laatifi *et al.*<sup>[19]</sup> used Uniform Manifold Approximation

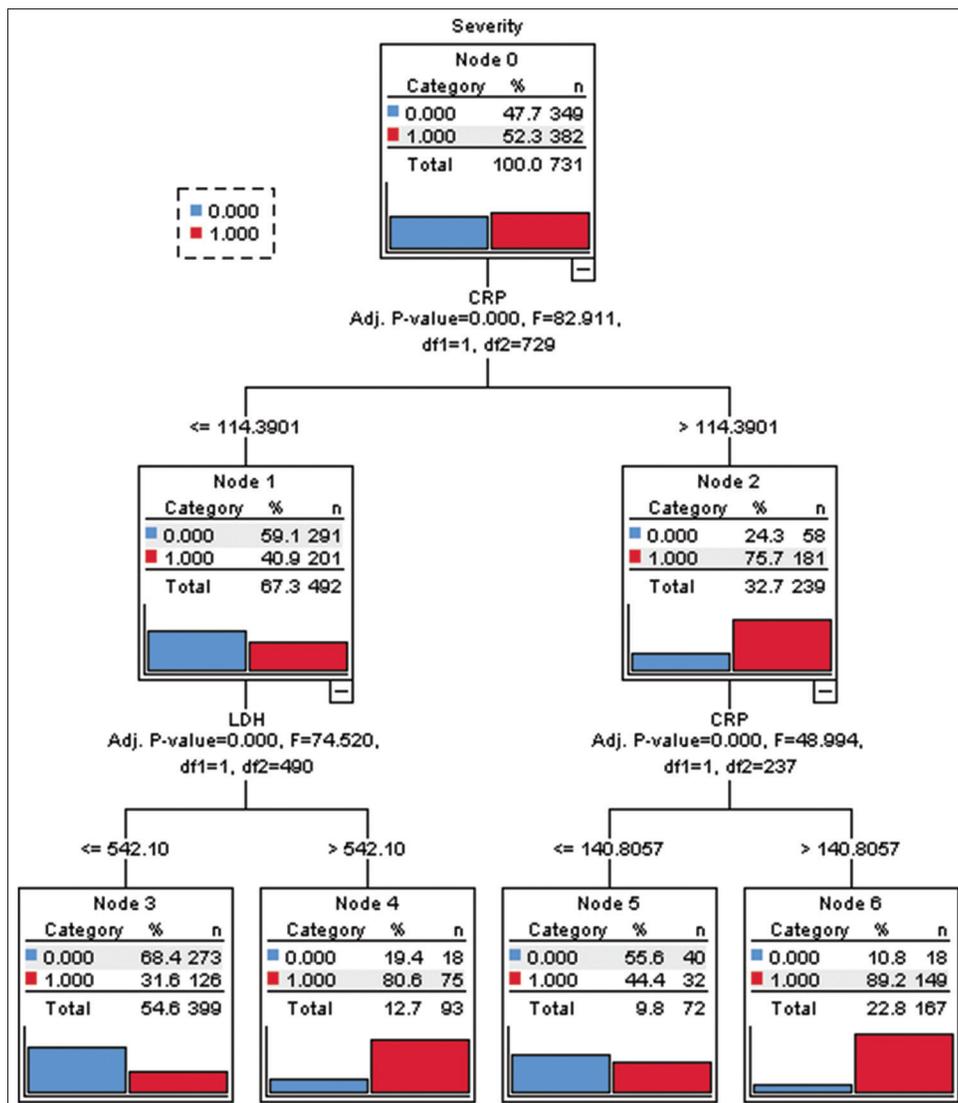
and Projection (UMAP) approach with 100% accuracy, specificity and sensitivity in prognostic prediction using different ML classifiers such as XG Boost, Ada Boost, Random Forest and Extra Trees. Moulai K *et al.*<sup>[20]</sup> studied on ML predictors of COVID-19 mortality and concluded that Random Forest (RF) ML model enables in predicting the COVID-19 mortality with a reasonable level of accuracy. Yazeed Zoabi *et al.*<sup>[21]</sup> suggested that predictions generated using a gradient-boosting machine model built with decision-tree base-learners 20 had better performance when clinical and laboratory parameters were combined for



**Figure 2:** ROC curve for Random Forest classification using selective parameters

**Table 3: Random Forest classification model of only selected parameters based on lasso regression**

Evaluation Metrics	Non-ICU	ICU	Overall
Accuracy	93.6%	93.6%	93.6%
Precision	0.939	0.615	0.918
Recall	0.996	0.085	0.936
AUC	0.815	0.823	0.819



**Figure 3:** Decision tree of training data set

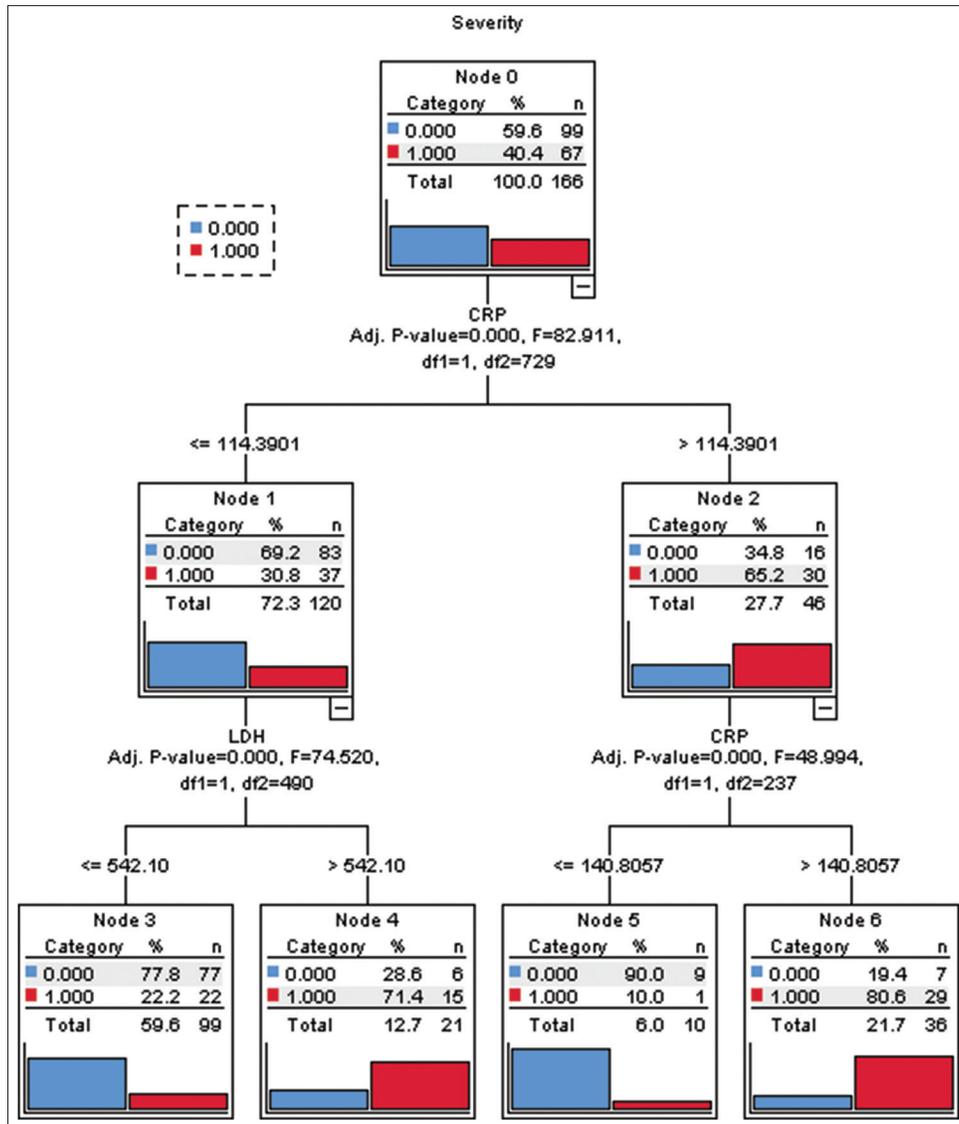


Figure 4: Decision tree of testing data set

Table 4: Classification strength of decision-tree algorithm in training and testing set and simplified algorithm

	Training Set	Testing Set	Simplified Algorithm
Sensitivity	66.45	78.90	87.95
Specificity	86.15	77.19	61.47
PPV	89.68	86.87	69.49
NPV	58.64	65.67	83.64
Overall Model AUC	0.744 (0.711–0.777, P<0.001)		0.747 (0.714–0.780, P<0.001)

predicting COVID-19 severity. Previous studies with ML models in cases of COVID 19 are usually prone to the problem of not being able to be translated into healthcare system because they often provide mixed results.<sup>[22]</sup> However, the advantage our study has over these concerns is that, compared with the earlier studies that were analysed in the early stages of the pandemic, our study was done at a later stage.<sup>[23-26]</sup> This provided us with

an ample insight into the clinical scenario, and data mining of records was thus worked efficiently to provide as much relevant data as required to the ML models.

Most studies relied on laboratory-based parameters for predicting severity in COVID-19. IL-6 was the most prominently studied parameter.<sup>[27,28]</sup> However, it has also been associated significantly with CRP, ALT, AST, urea, eGFR and serum ferritin.<sup>[24,28-30]</sup> In our study, all these parameters were studied, however, to minimize the use of resources and for swift diagnosis with cost-effectiveness, which can be used even in a primary healthcare set-up; a decision-making regression analysis was used to reduce it down to fewer parameters. Finally, the decision tree used two primary parameters, i.e., CRP and LDH for severity prediction in such patients. This is in line with previous studies in Egypt and India, which also have had severity analysis, and CRP and LDH were significantly important.<sup>[7,31,32]</sup> Cugnata *et al.*<sup>[33]</sup> analysed

Italian data for COVID-19 severity, which was one of the worst affected countries using multiple integrated statistical analysis methodologies and model and concluded with similar results as this study.

Our study had certain limitations because the data available for ICU and non-ICU patient were skewed. Therefore, before algorithm processing, the samples had to be randomized to bring numerical parity to avoid unequal variations. This study included the analytical tests that were being performed routinely during the time of COVID 19 pandemic, which might have narrowed down the spectrum. Notwithstanding the limitations, this study provides the information pertaining to three global waves of pandemic from Eastern India using ML and artificial intelligence models to predict severity in COVID-19 patients, which can be readily translated into primary healthcare set-ups, based on the decision-making algorithm generated. Timely assessment of severity with limited resources at peripheral healthcare set-ups with simple routinely available biochemical investigations will strengthen the quality patient care practices of primary care physicians.

## Conclusions

This study of large COVID-19 data from a tertiary care hospital provided a robust decision-making algorithm with minimal tests to be performed and analysed. This algorithm can thus be easily used by primary care physicians in deciding the ICU admission, thus lessening burden on the peripheral healthcare system. As CRP and LDH are laboratory tests performed routinely in most of rural clinical laboratories, this algorithm can be applied for future for severity assessment as a tool of pandemic preparedness.

## Key points

- COVID-19 symptomatic patients have been reported to rapidly progress to unexpected complications needing intensive care unit admission and ventilator support.
- The laboratory parameters were not consistent with clinical outcomes and were not useful predictors of COVID-19 severity.
- Machine learning (ML) algorithms have gained popularity because they can learn nonlinear and complex interactions and thus can improve the prediction accuracy.
- The results of this study represent 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> wave of COVID-19, i.e., 5715 COVID-19 patient biochemical and inflammatory parameters routinely used to manage the pandemic.
- Accuracy of Support Vector Machine (SVM) was highest with 94.18% and Random Forest Classifier (RFC) with 94.04%. SVM had highest PPV (1.00), and Naive Bayesian Classifier (NBC) had highest NPV (0.95) as observed in 5795 COVID-19 patients.
- C-reactive protein- and lactate dehydrogenase (LDH)-based decision-tree algorithm was best in classifying ICU admission in COVID-19 and, being routinely performed in all rural

laboratories, could be an effective primary care predictive tool.

## Ethical approval

This study was approved by Institutional Ethical Committee of AIIMS, Bhubaneswar (T/IM-NF/Biochem/21/162 dated 28<sup>th</sup> Feb 2022).

## Acknowledgments

We are thankful to our colleagues and technical staff of the Clinical Biochemistry Laboratory, AIIMS, Bhubaneswar.

## Financial support and sponsorship

Nil.

## Conflicts of interest

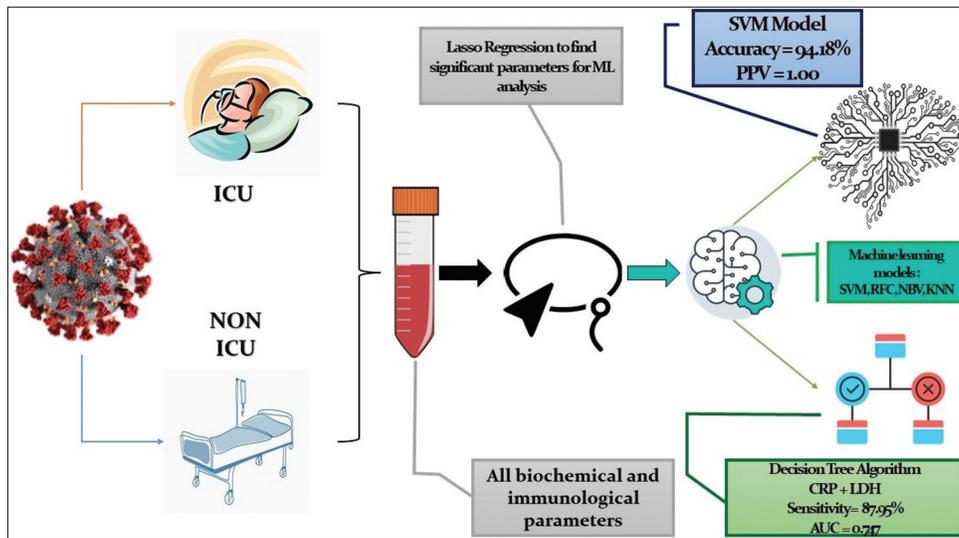
There are no conflicts of interest.

## References

1. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *Lancet* 2020;395:1054-62.
2. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *Int J Antimicrob Agents* 2020;55:105924.
3. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor recognition by the novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS coronavirus. *J Virol* 2020;94:e00127-20.
4. Zhang W, Zhao Y, Zhang F, Wang Q, Li T, Liu Z, *et al.* The use of anti-inflammatory drugs in the treatment of people with severe coronavirus disease 2019 (COVID-19): The Perspectives of clinical immunologists from China. *Clin Immunol* 2020;214:108393.
5. Conti P. Induction of pro-inflammatory cytokines (IL-1 and IL-6) and lung inflammation by COVID-19: Anti-inflammatory strategies. *J Biol Regul Homeost Agents* 2020;34:1.
6. Kritas SK, Ronconi G, Caraffa A, Gallenga CE, Ross R, Conti P. Mast cells contribute to coronavirus-induced inflammation: New anti-inflammatory strategy. *J Biol Regul Homeost Agents* 2020;34:9-14.
7. Kumari S, Nayak S, Tripathy S, Bhuniya S, Mangaraj M, Ramadass B, *et al.* Analysis of biochemical and inflammatory markers for predicting COVID-19 Severity: Insights from a tertiary healthcare institution of Eastern India. *Cureus* 2023;15:e33893. doi: 10.7759/cureus.33893.
8. Haljasmägi L, Salumets A, Rumm AP, Jürgenson M, Krassohhina E, Remm A, *et al.* Longitudinal proteomic profiling reveals increased early inflammation and sustained apoptosis proteins in severe COVID-19. *Sci Rep* 2020;10:20533.
9. Meizlish ML, Pine AB, Bishai JD, Goshua G, Nadelmann ER, Simonov M, *et al.* A neutrophil activation signature predicts critical illness and mortality in COVID-19. *Blood Adv* 2021;5:1164-77.

10. Qiu T, Liang S, Dabbous M, Wang Y, Han R, Toumi M. Chinese guidelines related to novel coronavirus pneumonia. *J Mark Access Health Policy* 2020;8(1):1818446. doi: 10.1080/20016689.2020.1818446.
11. Xiang J, Wen J, Yuan X, Xiong S, Zhou X, Liu C, *et al.* Potential biochemical markers to identify severe cases among COVID-19 patients. *Epidemiology* 2020. Available from: <http://medrxiv.org/lookup/doi/10.1101/2020.030.19.20034447>.
12. Jin JM, Bai P, He W, Wu F, Liu XF, Han DM, *et al.* Gender differences in patients with COVID-19: Focus on severity and mortality. *Front Public Health* 2020;8:152.
13. Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, *et al.* Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng* 2018;2:749–60.
14. Quesada JA, Lopez-Pineda A, Gil-Guillén VF, Durazo-Arvizu R, Orozco-Beltrán D, López-Domenech A, *et al.* Machine learning to predict cardiovascular risk. *Int J Clin Pract* 2019;73:e13389. doi: 10.1111/ijcp.13389.
15. Escobar GJ, Puopolo KM, Wi S, Turk BJ, Kuzniewicz MW, Walsh EM, *et al.* Stratification of risk of early-onset sepsis in newborns  $\geq 34$  weeks' gestation. *Pediatrics* 2014;133:30–6.
16. Vaishya R, Javaid M, Khan IH, Haleem A. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab Syndr* 2020;14:337–9.
17. Zhou K, Sun Y, Li L, Zang Z, Wang J, Li J, *et al.* Eleven routine clinical features predict COVID-19 severity uncovered by machine learning of longitudinal measurements. *Comput Struct Biotechnol J* 2021;19:3640–9.
18. Baker TB, Loh WY, Piasecki TM, Bolt DM, Smith SS, Slutske WS, *et al.* A machine learning analysis of correlates of mortality among patients hospitalized with COVID-19. *Sci Rep* 2023;13:4080.
19. Laatifi M, Douzi S, Bouklouze A, Ezzine H, Jaafari J, Zaid Y, *et al.* Machine learning approaches in Covid-19 severity risk prediction in Morocco. *J Big Data* 2022;9:5.
20. Moulaei K, Shanbehzadeh M, Mohammadi-Taghiabad Z. Comparing machine learning algorithms for predicting COVID-19 mortality. *BMC Med Inform Decis Mak* 2022;22:2.
21. Zoabi Y, Deri-Rozov S, Shomron N. Machine learning-based prediction of COVID-19 diagnosis based on symptoms. *NPJ Digit Med* 2021;4:3.
22. Kwok SWH, Wang G, Soheli F, Kashani KB, Zhu Y, Wang Z. An artificial intelligence approach for predicting death or organ failure after hospitalization for COVID-19: Development of a novel risk prediction tool and comparisons with ISARIC-4C, CURB-65, qSOFA, and MEWS scoring systems. *Respir Res* 2023;24:79.
23. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E. Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. *BMJ* 2020;369:m1328. doi: 10.1136/bmj.m1328. Update in: *BMJ* 2021;372:n236. Erratum in: *BMJ* 2020;369:m2204.
24. Bennett TD, Moffitt RA, Hajagos JG, Amor B, Anand A, Bissell MM. Clinical characterization and prediction of clinical severity of SARS-CoV-2 infection among us adults using data from the US National COVID cohort collaborative. *JAMA Netw Open* 2021;4:e2116901.
25. Wang L, Zhang Y, Wang D, Tong X, Liu T, Zhang S. Artificial Intelligence for COVID-19: A systematic review. *Front Med* 2021;8:704256.
26. Yadaw AS, Li Y chak, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of COVID-19 mortality: Development and validation of a clinical prediction model. *Lancet Digit Health* 2020;2:e516–25.
27. Sabaka P, Koščálová A, Straka I, Hodosy J, Lipták R, Kmotorková B. Role of interleukin 6 as a predictive factor for a severe course of Covid-19: Retrospective data analysis of patients from a long-term care facility during Covid-19 outbreak. *BMC Infect Dis* 2021;21:308.
28. Herold T, Jurinovic V, Arnreich C, Lipworth BJ, Hellmuth JC, von Bergwelt-Baildon M. Elevated levels of IL-6 and CRP predict the need for mechanical ventilation in COVID-19. *J Allergy Clin Immunol* 2020;146:128–36.e4.
29. Hachim MY, Hachim IY, Naeem KB, Hannawi H, Salmi IA, Hannawi S. D-dimer, Troponin, and urea level at presentation with COVID-19 can predict ICU admission: A single centered study. *Front Med* 2020;7:585003.
30. Hulkoti V, Acharya S, Kumar S, Talwar D, Khanna S, Annadatha A. Association of serum ferritin with COVID-19 in a cross-sectional study of 200 intensive care unit patients in a rural hospital: Is ferritin the forgotten biomarker of mortality in severe COVID-19? *J Fam Med Prim Care* 2022;11:2045.
31. Abd El-Aty SK, Ahmed SH, Abd El-Hamid RF, Zahran AM, Magdy DM, Bakr OH. Markers of inflammasome activation in Coronavirus disease 2019 (COVID-19) Egyptian patients: Impact on disease severity. *Egypt J Immunol* 2023;30:99–108.
32. Kumar V, Mahto M, Kumar S, Bansal A, Ranjan A, Ahmad S, *et al.* Variation in biochemical parameters in COVID-19 patients admitted at a tertiary care dedicated COVID hospital: A prospective study. *Ann Afr Med* 2023;22:176–82.
33. Cugnata F, Scarale MG, De Lorenzo R, Simonini M, Citterio L, Querini PR. Profiling Covid-19 patients with respect to level of severity: An integrated statistical approach. *Sci Rep* 2023;13:5498.

## Supplement Materials



Supplement Figure 1: Graphical abstract of the study