



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Contents lists available at ScienceDirect

## International Journal of Medical Informatics

journal homepage: [www.elsevier.com/locate/ijmedinf](http://www.elsevier.com/locate/ijmedinf)

# Prognosing the risk of COVID-19 death through a machine learning-based routine blood panel: A retrospective study in Brazil

Daniella Castro Araújo<sup>a,b,\*</sup>, Adriano Alonso Veloso<sup>b</sup>, Karina Braga Gomes Borges<sup>c</sup>,  
Maria das Graças Carvalho<sup>c</sup>

<sup>a</sup> Huna, São Paulo, SP, Brazil

<sup>b</sup> Departamento de Ciência da Computação, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

<sup>c</sup> Faculdade de Farmácia, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

## ARTICLE INFO

## Keywords:

COVID-19  
Machine learning  
Prognosis  
Blood biomarkers  
Artificial intelligence  
Imbalance

## ABSTRACT

**Background::** Despite an extensive network of primary care availability, Brazil has suffered profoundly during the COVID-19 pandemic, experiencing the greatest sanitary collapse in its history. Thus, it is important to understand phenotype risk factors for SARS-CoV-2 infection severity in the Brazilian population in order to provide novel insights into the pathogenesis of the disease.

**Objective::** This study proposes to predict the risk of COVID-19 death through machine learning, using blood biomarkers data from patients admitted to two large hospitals in Brazil.

**Methods::** We retrospectively collected blood biomarkers data in a 24-h time window from 6,979 patients with COVID-19 confirmed by positive RT-PCR admitted to two large hospitals in Brazil, of whom 291 (4.2%) died and 6,688 (95.8%) were discharged. We then developed a large-scale exploration of risk models to predict the probability of COVID-19 severity, finally choosing the best performing model regarding the average AUROC. To improve generalizability, for each model five different testing scenarios were conducted, including two external validations.

**Results::** We developed a machine learning-based panel composed of parameters extracted from the complete blood count (lymphocytes, MCV, platelets and RDW), in addition to C-Reactive Protein, which yielded an average AUROC of  $0.91 \pm 0.01$  to predict death by COVID-19 confirmed by positive RT-PCR within a 24-h window.

**Conclusion::** Our study suggests that routine laboratory variables could be useful to identify COVID-19 patients under higher risk of death using machine learning. Further studies are needed for validating the model in other populations and contexts, since the natural history of SARS-CoV-2 infection and its consequences on the hematopoietic system and other organs is still quite recent.

## 1. Introduction

Researchers have made unprecedented rapid progress in understanding the occurrence, progression, and treatment of the COVID-19 disease, but one of the most intriguing questions remains unanswered: why do some people die while others present mild symptoms? Although previous studies have reported that older age, male gender, smoking, and other conditions such as hypertension, diabetes mellitus, obesity, and chronic lung disease are risk factors for severe illness or death [1–3], these factors alone do not explain all variability in disease severity observed among individuals.

Recently, the COVID-19 Host Genetics Initiative has reported 13 genome-wide significant loci that are associated with SARS-CoV-2 infection or severe manifestations of COVID-19, several of them corresponding to previously documented associations to lung or autoimmune and inflammatory diseases [4]. Besides host genetics, it is well known that environmental, clinical, and social factors are also important to the disease severity [5].

Despite an extensive network of primary care availability, Brazil has suffered profoundly during the (SARS-CoV-2) pandemic, experiencing the greatest sanitary collapse in its history [6]. Home to just over 2.7 percent of the world's population, Brazil accounts for almost 12% of

\* Corresponding author at: Huna, São Paulo, SP, Brazil.

E-mail address: [dani@huna-ai.com](mailto:dani@huna-ai.com) (D.C. Araújo).

<https://doi.org/10.1016/j.ijmedinf.2022.104835>

Received 8 June 2022; Received in revised form 17 July 2022; Accepted 19 July 2022

Available online 27 July 2022

1386-5056/© 2022 Elsevier B.V. All rights reserved.

recorded fatalities. With 30.6 million cases and 664 thousand deaths (as of 6 May 2022), the country has the second-highest total of deaths in the world, behind only the United States [7]. In this context, it is highly relevant to study the patterns associated with the mortality rate due to COVID-19 in Brazil.

Previous studies have used blood biomarkers data to develop prognostic models for COVID-19 death using Machine Learning (ML), frequently reaching over 0.90 Area under the ROC Curve (AUROC). These studies suggest that ML techniques are able to unlock the predictive power of non-linear relationships between blood biomarkers [8]. Therefore, in addition to the most common assessment methods used to monitor the progress of pulmonary disease, such as X-rays and CT-scan images, blood tests could also be used as indicators of the COVID-19 severity [9,10]. One of the first studies to address the problem was the retrospective proposed by Yan et al. (2020) [11]. They used a database of 375 patients from a hospital in Wuhan, China, to develop a ML model using lactic dehydrogenase, lymphocyte and high-sensitivity C-reactive protein (hs-C-reactive protein (CRP)) as features, and further validated the model using a database of 110 patients [11]. In the recent survey conducted by Carobene et al. (2022) [12], they found 34 ML studies that aim to predict the risk of intensive care admission, mechanical ventilation and/or death due to COVID-19 using blood biomarkers, demographic characteristics, comorbidities and vital signs as features. Four of these studies used only laboratory data as features [13] [14,15] [9], of which none have been externally validated. Booth et al. (2020) have built a panel composed of five lab biomarkers (c-reactive protein, blood urea nitrogen, serum calcium, serum albumin, and lactic acid) of 398 patients from Texas, USA, for the prediction of death by COVID-19. They achieved 93% AUROC using a support vector machine model [13]. From a database of 196 hospitalized patients from Wuhan, China, Luo et al. (2021) proposed a multi-criteria decision-making algorithm that achieved 93% AUROC using only CBC biomarkers and age to predict severity risk due to COVID-19 [14]. Qomariyah et al. (2021) analysed a dataset composed of 1,000 patients from Indonesia to build a model for predicting the risk of COVID-19 death, using eleven blood biomarkers [9]. Analysing routine blood analytes, Darapaneni et al. (2021) presented a lasso regression that obtained 87% f1-score for predicting the risk of ICU hospitalization, using a database of 5,644 individuals from a Brazilian hospital [15].

Besides the latter, other prognostic studies were conducted using Brazilian databases. Analysing a Brazilian dataset composed of 1,945 patients from a Brazilian hospital, Aktar et al. (2021) developed ML models to predict the risk of ICU admission, trained with routine blood biomarkers and other parameters, such as comorbidities and blood gas analysis [16]. Fernandes et al. (2021) developed a ML model to predict the risk of ICU admission, use of mechanical ventilation and/or death analysing a database of 1,040 patients from a hospital in Brazil, using the routine biomarkers lymphocytes, CRP and ferritin, together with Intensive Care Unit (ICU) scores [17]. Famiglini et al. (2022) developed ML models for the prediction of ICU patient admission using only Complete Blood Count (CBC) data from Italy [18], which has been further externally tested in databases from multiple countries [19], but had the worst performance rates in the Brazilian ones. All of these studies aimed to predict the risk of ICU admission, and some of them used parameters other than routine blood biomarkers. Therefore, to our knowledge, there are no previous studies that have developed ML risk models to predict the risk of death by COVID-19 using only routine and non-expensive blood biomarkers data in Brazil.

Here, we propose a ML model for death prediction caused by COVID-19 disease that was obtained by combining different routine laboratory blood biomarkers data. We have considered only blood tests collected in a 24-h time window before or after the first positive Reverse Transcription-Polymerase Chain Reaction (RT-PCR) diagnosis. We used datasets from two private Brazilian hospitals - Hospital Sírio Libanês (HSL) and Hospital Beneficência Portuguesa (HBP) [20]. Validation was performed both intra and inter-hospitals, and five different measures

were reported: AUROCs for training and cross-validation testing at each hospital, AUROC for training and cross-validation testing at both hospitals together, and AUROCs for training in one hospital and external testing in the other.

Besides being the first study to use only routine laboratory variables to predict the risk of COVID-19 death for patients in Brazil, another significant contribution of this study is that the proposed model was externally validated, limiting overfitting and underspecification risks. Furthermore, our total database size is the biggest one in Brazil with a focus on this issue. It is worth emphasizing that the identification of mortality predictors can be valuable for clinical risk stratification, which will undoubtedly contribute to optimizing monitoring and therapy of patients with COVID-19.

## 2. Methods

### 2.1. Data preparation

Retrospective data from patients treated and/or admitted at the HSL from feb/2020 to jun/2020 and HBP from feb/2020 to mar/2021 provided by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) available at the COVID-19 Data Sharing/BR [20] were used. These datasets contain electronic health records from individuals who have laboratory confirmation of COVID-19 by RT-PCR assays, as well as their outcome (death or discharge). Firstly, we filtered the datasets considering only those individuals who had a positive RT-PCR result. Then, we filtered the blood tests data in a 24-h time window before or after the positive RT-PCR diagnosis, and dropped individuals who did not have at least a CBC result. We ended up with 4,374 individuals originating from the HSL, of whom 151 (3.4%) died; and 2,605 from the HBP, of whom 140 (5.3%) died, totaling 6,979 unique individuals, of whom 291 (4.2%) died. Fig. 1 shows a diagram of the datasets filtering.

The datasets initially contained more than 300 biomarkers each. After filtering the biomarkers that had results for at least 80% of the individuals for both datasets, we ended up with 23 biomarkers, plus sex. We also calculated seven blood-cell-derived indexes, totaling 30 variables and 208,074 measures (excluding missing values). To deal with the remaining missing values, we used an imputation approach based on K-Nearest Neighbors [21]. Supplementary material in Appendix 1, Table A1, shows the variables and number of measures for each database.

Because of the high imbalance regarding the target variable, given that our positive class corresponds to less than 6% of the databases, it seemed important to use a technique to adjust this issue. However, the recent studies of van den Goorbergh et al. (2022) [22] indicated that imbalance is not a problem in itself: imbalance correction methods may cause poor calibration and even worsen model performance in terms of the AUROC. Besides, Elor & Averbuch-Elor (2022) [23] study showed that balancing datasets with Synthetic Minority Over-Sampling Technique (SMOTE) re-sampling procedure may improve prediction performance for weak classifiers, but not for the state-of-the-art classifiers, such as lightGBM [24] and XGBoost [25]. Therefore, we chose to set an experiment to compare three different configurations of SMOTE to the as-is databases, that is, to training the model in the imbalanced databases. We used balanced metrics to evaluate the performance of the model in each one of these as-is/oversampling scenarios.

### 2.2. Models

We developed a large-scale exploration of risk models for the binary prediction of COVID-19 death, each one exploiting different interactions between the 30 features' raw values. For each model, we randomly selected up to 10 features, thus resulting in models with diverse predictive performance.

The models were built using a fast implementation of the LightGBM algorithm [24], which produces a complex model composed of hundreds of simple decision trees that are finally combined into a single model by

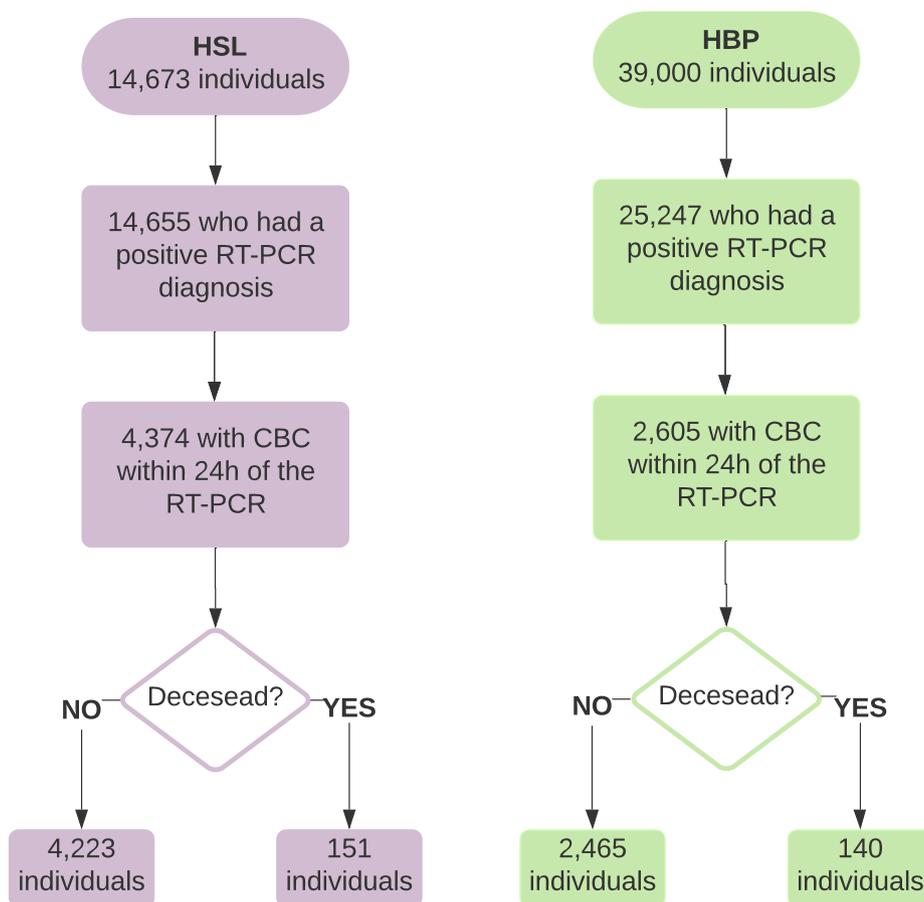


Fig. 1. Diagram of the number of individuals for both datasets.

a process known as boosting [26]. The predictive performance of each model is presented as a standard area under the ROC curve (AUROC) measurement [27]. For each model, five different testing scenarios were conducted, being three internal (1–3) and two external (4–5): (1) training and testing with HSL dataset, using 5-fold stratified-cross-validation; (2) training and testing with HBP dataset, using 5-fold stratified-cross-validation; (3) training and testing with HBP and HSL datasets together, using 5-fold stratified-cross-validation; (4) training with HSL dataset and external testing with HBP and (5) training with HBP dataset and external testing with HSL. In the first three scenarios, at each run, 4 folds were used as a training set and the 5th fold was used as a test set. To avoid data leakage, the data splitting was performed before any pre-processing and model construction steps [28], namely (1) missing value imputation, (2) re-sampling, (3) feature selection and (4) hyperparameter optimization.

We used the external robustness diagram proposed by Cabitza et al. [19] to display the external validation results, which integrates two different sets of metrics: the minimum dataset cardinality [29], that shows if the external database has sufficient sample size; and the multivariate similarity [30] of the external dataset, with respect to the training set. In addition, three quality dimensions are shown in light of the similarity: model discrimination power (AUROC); model utility (Net Benefit); and model calibration (Brier Score).

Similarly to FBeta-measure, which employs a parameter to control the trade-off between precision and sensitivity [31], we proposed to use a **Weighted Balanced Accuracy (WBA)**, which is defined in (1). Basically, we may improve sensitivity and specificity by optimizing WBA while performing a grid search on the model hyperparameters. Table 1 summarizes the search grid and the selected parameters.

$$WBA = \frac{2 * Sensitivity + Specificity}{3}$$

Fig. 2 presents the feature selection steps of the proposed methodology.

For every scenario, we reported seven different metrics, evaluated on the test sets, namely: AUROC, sensitivity, specificity, accuracy, balanced accuracy, WBA and the Brier score (as a measure of calibration).

We examined data of 4,374 individuals admitted in HSL, of whom 151 (3.4%) died; and of 2,605 individuals admitted in HBP, of whom 140 (5.4%) died.

Table 2 shows the baseline age and sex characteristics of the two groups for both hospitals. Major comorbidities, ethnicity and socioeconomic status were not available.

After following the methodology represented in Fig. 2, thousands of models with randomly selected features were developed. The one with the best performance in relation to the average AUROC estimate was obtained by plotting the rate of correctly classified positives among all positive predictions (i.e., the true-positive rate) as a function of incorrect positives among all negatives (i.e., the false-positive rate), at varying thresholds. Because the output of the model is a probability (i.e., the risk

Table 1  
Hyperparameter search space of the best model.

Hyperparameter	Search grid	Selected Parameter
max_bin	10, 50, 100, 200	100
max_depth	5, 10, 15	15
learning_rate	0.05, 0.1, 0.2	0.1
num_leaves	5, 10, 20	5
min_data	5, 10, 15	50

## 2 DATABASES

**HSL:** 4,374 individuals

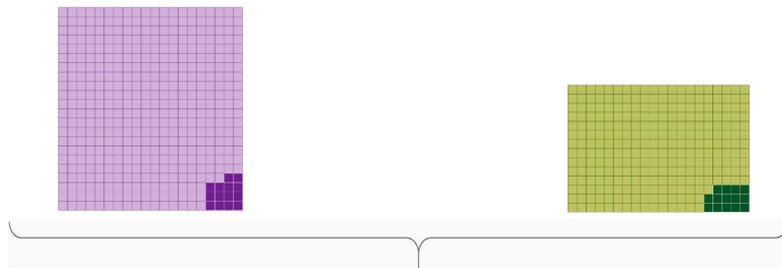
**HBP:** 2,605 individuals

4,223 who were discharged

151 who have died

2,465 who were discharged

140 who have died

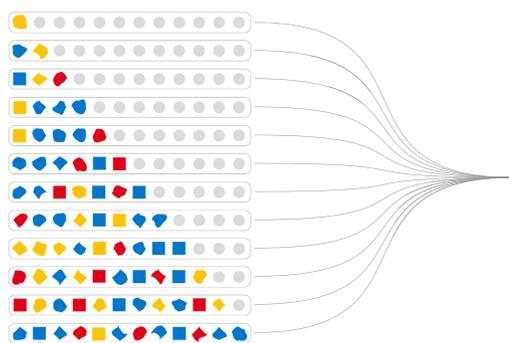


For each exam, there are **31 FEATURES**

**Fig. 2.** Diagram presenting the datasets and the feature selection steps of the proposed methodology. First, we went into a large exploration creating models from the random combination of up to 10 features (out of the 30 possible features). Then, for each one of these models, five different testing scenarios were conducted, being three internal (1–3) and two external (4–5): (1) training and testing with HSL dataset, using 5-fold stratified-cross-validation; (2) training and testing with HBP dataset, using 5-fold stratified-cross-validation; (3) training and testing with HBP and HSL datasets together, using 5-fold stratified-cross-validation; (4) training with HSL dataset and external testing with HBP and (5) training with HBP dataset and external testing with HSL. The mean AUROC and standard deviation (STD) of these five scenarios for each one of the models were estimated, and chosen the best-performing one (that is, the one with the highest mean AUROC with the lowest STD).

### PHASE 1: EXPLORATION

Random generation of panels that contain from 1 to 10 features

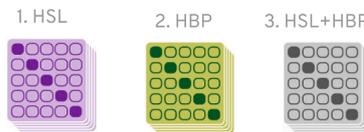


Some of these features are not filled for all individuals

### PHASE 2: VALIDATION

We performed model validation in 5 scenarios

Intra databases, using 5-fold cross validation:



Inter databases:



We estimated the mean AUROC for the five scenarios for each one of the models, leading us to the best performing model

**Table 2**  
Descriptive statistics of Discharged vs Death groups in both hospitals.

Hospital	Variable	Discharged Average ± SD	Death Average ± SD	Missing Values (%)
HSL	No. of individuals	4,223	151	-
	Age, years	52.4 ± 16.6	74.9 ± 12.6	2.2
	Sex, % female	44%	34%	0.0
HBP	No. of individuals	2,465	140	-
	Age, years	N/A <sup>1</sup>	N/A <sup>1</sup>	N/A
	Sex, % female	49%	39%	0.0

Due to privacy concerns, HBP chose to omit most of the individuals' age.

factors for COVID-19 death), each threshold is a value ranging from 0 to 1.

The best performing model had an AUROC of  $0.91 \pm 0.01$  and was composed of a panel containing five plasma variables, CRP, limphocytes, medium corpuscular volume (MCV), platelets, and red cell distribution width (RDW), as shown in [Table 3](#).

[Table 4](#) compares the model performance using different balancing techniques: SMOTE with three different configurations for the desired ratio between positive and negative samples ( $r$ ): 0.1, 0.5 and 1.0, maintaining the default nearest neighbors number of 5; and the as-is

database, with no oversampling. Measures AUROC, Brier Score and WBA after threshold optimization are reported in terms of the mean ± std for the five testing scenarios. As expected, because we use LightGBM, as the representative of the state-of-the-art [23], the general performance of the model is not improved with SMOTE. Actually, when we use higher ratios of 0.5 and 1.0 (meaning that we oversample the positive class until it has the total number of samples of respectively half and the same as the negative class), the Brier Score increases. On the other hand, when we use a ratio of 0.1 the performance metrics remain practically the same. Therefore, all the reported further results will consider the as-

**Table 3**  
Descriptive statistics of the five features that compose the best model.

Hospital	Variable	Discharged	Death	Missing values (%)	P value
		Average $\pm$ SD (min–max)	Average $\pm$ SD (min–max)		
HSL	CRP, mg/dL	2.67 $\pm$ 4.84 (0.0–46.4)	9.32 $\pm$ 8.36 (0.0–41.4)	2.4	< 0.001
	Limphocytes, %	24 $\pm$ 12 (0–90)	12 $\pm$ 10 (0–60)	0.0	< 0.001
	MCV, fL	87.79 $\pm$ 5.13 (55.8–115.2)	91.29 $\pm$ 7.14 (67.0–112.0)	0.0	< 0.001
	Platelets, $10^3/mm^3$	207.35 $\pm$ 66.20 (7.0–642.0)	174.57 $\pm$ 84.05 (13.0–553.0)	0.0	< 0.001
	RDW, %	13.20 $\pm$ 1.37 (10.9–28.6)	15.45 $\pm$ 2.33 (11.6–23.9)	0.0	< 0.001
HBP	CRP, mg/dL	2.63 $\pm$ 4.20 (0.0–41.8)	12.50 $\pm$ 10.95 (0.2–52.5)	5.3	< 0.001
	Limphocytes, %	26 $\pm$ 12 (0–90)	15 $\pm$ 10 (0–50)	0.0	< 0.001
	MCV, fL	86.43 $\pm$ 5.13 (56.6–113.5)	89.43 $\pm$ 6.49 (71.8–116.1)	0.0	< 0.001
	Platelets, $10^3/mm^3$	211.99 $\pm$ 69.15 (18.0–1196.0)	179.42 $\pm$ 91.86 (23.0–627.0)	0.1	< 0.001
	RDW, %	13.18 $\pm$ 1.22 (10.9–26.5)	15.29 $\pm$ 2.51 (11.9–24.0)	0.0	< 0.001

**Table 4**  
Performance of different oversampling scenarios.

Oversampling scenario	AUROC	Brier Score	WBA
SMOTE (k = 5 and r = 0.1)	0.91 $\pm$ 0.00	0.04 $\pm$ 0.00	0.84 $\pm$ 0.00
SMOTE (k = 5 and r = 0.5)	0.91 $\pm$ 0.00	0.07 $\pm$ 0.01	0.85 $\pm$ 0.01
SMOTE (k = 5 and r = 1.0)	0.91 $\pm$ 0.01	0.10 $\pm$ 0.01	0.85 $\pm$ 0.01
As-is (no oversampling)	0.91 $\pm$ 0.01	0.03 $\pm$ 0.00	0.84 $\pm$ 0.01

is scenario, without SMOTE oversampling. (see Table 5).

Fig. 3 shows the AUROC performance for the five scenarios of our best risk model without oversampling ( $0.91 \pm 0.01$ ). Clearly, the predictive performance varies slightly within the internal and external validation scenarios (AUROC of 0.89, 0.91, 0.91, 0.91 and 0.92 for scenarios 1–5). The reliability (calibration) curves are shown in Fig. 4, the same pattern for the brier score ( $0.03 \pm 0.00$ ) can be observed.

For the threshold calibration [32], we moved the threshold for each scenario in order to find the probability cutoff that maximized the average WBA. Table 4 shows the classification performance metrics of accuracy, balanced accuracy, sensitivity, and specificity for each one of the five validation scenarios using the calibrated threshold. In average, the models have a sensitivity of  $0.88 \pm 0.05$ , meaning that of 100 patients at higher risk, 88 would be recognized by the algorithm; and specificity of  $0.76 \pm 0.06$ , meaning that of 100 patients at lower risk, 24 would be wrongly recognized as at higher risk.

The high performance on both external testing sets may be justified by the high homogeneity of the two hospital datasets: they are both private hospitals located in the city of São Paulo, and their laboratory analyses are performed by the same company (Grupo Fleury). To evaluate if the external validations could be considered reliable, we followed the methodology in [19], and estimated the similarity between each dataset with respect to the other, as well as their AUC values, net benefit, brier score, and respective cardinality for each one of these performance metrics. We depicted these results in the external performance diagram, as shown in Fig. 5. This diagram shows that the Minimum Sample Size (MSS) has been achieved for all metrics in both datasets, meaning that the datasets sample size, namely cardinality, is sufficient to guarantee the generalization of the results. Besides, it shows that the HBP dataset

**Table 5**  
Performance metrics for each one of the scenarios for the calibrated threshold.

Scenario	Sensitivity	Specificity	Accuracy	Balanced Accuracy	WBA
1	0.84	0.80	0.80	0.82	0.83
2	0.91	0.71	0.72	0.81	0.85
3	0.89	0.75	0.75	0.82	0.84
4	0.82	0.84	0.84	0.83	0.83
5	0.95	0.68	0.69	0.81	0.86
Average $\pm$ SD	<b>0.88 <math>\pm</math> 0.05</b>	<b>0.76 <math>\pm</math> 0.06</b>	<b>0.76 <math>\pm</math> 0.05</b>	<b>0.82 <math>\pm</math> 0.01</b>	<b>0.84 <math>\pm</math> 0.01</b>

has slight similarity with respect to HSL (0.33), and for all three performance metrics, this dataset had good (net benefit of 0.67) or excellent (AUC of 0.91 and brier score of 0.04) performance. On the other hand, there is a high similarity for HSL with respect to HBP (0.72). So, although this validation achieved excellent performance in all three performance metrics (AUC of 0.92, net benefit of 0.75 and brier score of 0.03), because of the high similarity, this can only be considered as another internal validation analysis. The high similarity of HSL in respect to HBP but not of HBP in respect to HSL may be justified by the temporal origin of the datasets: while HSL was collected only until jun/2020, HBP was collected until mar/2021, when the gamma variant was highly circulating in Brazil [33]. Finally, the proposed model can be considered externally validated, when trained in HSL and tested in HBP (Scenario 4).

Fig. 6 shows a 2D representation of the two groups: red dots represent individuals who died, whereas green dots represent individuals who were discharged. To build these visualizations, we applied the t-distributed stochastic neighbor embedding (t-SNE) algorithm [34]. Unlike PCA, t-SNE is a non-linear dimensionality reduction technique that tries to preserve the local structure of data [35], thus performing usually better for machine learning visualization [36]. However, t-SNE is highly sensitive to the setting of hyperparameters, notably the perplexity [37]. Fig. 6 (a, c and e) represents the raw analytes' concentrations for each individual for HSL, HBP and HSL + HBP, respectively. No clear distinction between the two groups was found, reflecting what might be observed in attempting to draw linear correlations between the five biomarkers. On the other hand, Fig. 6 (b, d and f) represents the corresponding marginal contributions of each analyte to the models, named Shapley values [38], reflecting all the non-linear interactions between these five biomarkers involved in the decision process of our models for HSL, HBP and HSL + HBP, respectively. In this scenario, a distinction between the two groups can be seen. For coherence in the comparisons, we have used the same hyperparameters for each pair of visualizations: perplexity was defined as 1% of the size of each dataset [39]; the initialization of embedding was set to PCA; and the other hyperparameters were set as the default of sklearn package. For sensitivity analysis, we have also tested perplexities of 0.5% and 10%, and the tendencies remain the same.

In order to assess the importance of the features and thus extract intuitive insights from the prediction, the SHAP algorithm [38] was applied to the model. Briefly, SHAP calculates the importance of each feature by estimating the effect of its absence on the model's decision. The importance of each feature for every individual was graphically represented, and these results are shown in the SHAP Summary Plot (Fig. 7), where features are depicted in order of importance. Red dots are associated with individuals for which the corresponding biomarker (feature) shows a relatively higher value. On the other hand, blue dots are associated with individuals for which the corresponding biomarker shows a relatively lower value. Further, there is a vertical line separating patients - the dots located on the left side are those for which the model

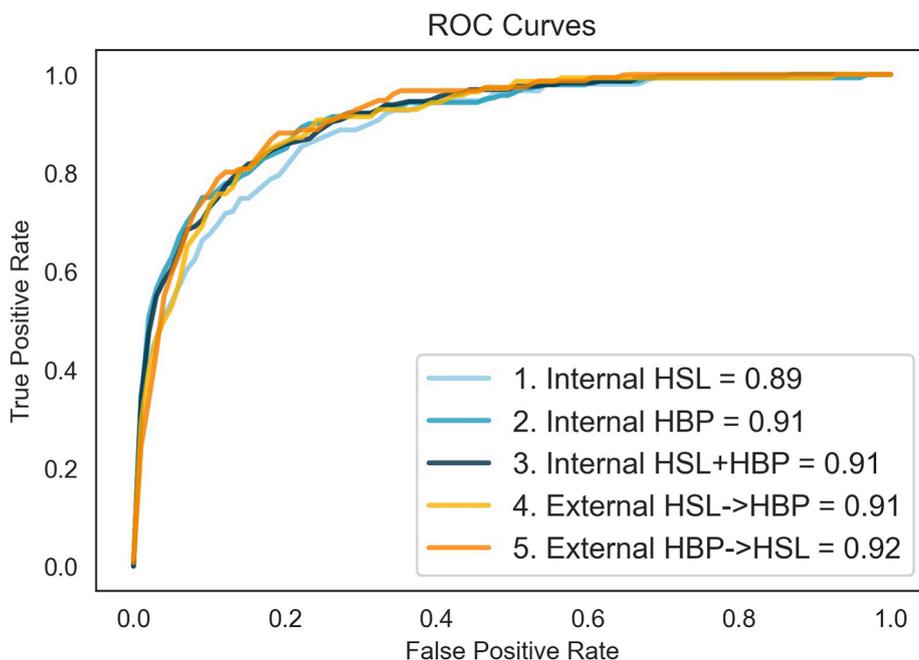


Fig. 3. AUROC curves for the five validation scenarios.

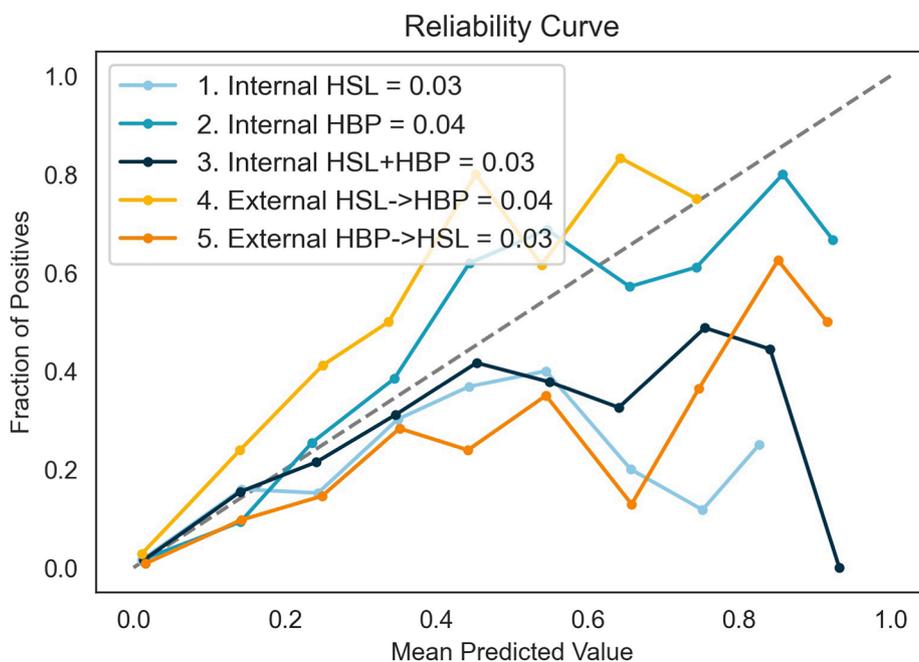


Fig. 4. Reliability curves for the five validation scenarios.

provided a negative decision (discharge) and, on the right, those related to death. Fig. 7a shows the SHAP Summary Plot for the model trained on the HSL database, 7b for the model trained on HBP database, and 7c for the model trained on HSL and HBP databases together.

Fig. 7 presents the five biomarkers that compose the models shown in order of importance for the HSL, HBP and HSL + HBP datasets. Although the order of importance of the features differs in the three presented SHAP summary plots, we can see that the SHAP values largest, smallest and mean observations for each feature in each one of the three models are extremely close. This indicates that there is no significant difference among the three models' decision mechanisms. Furthermore, in all three SHAP summary plots, we can observe the same tendencies for each

feature, reinforcing the consistency between the three models. We can see more red dots on the right and blue dots on the left side for RDW, CRP and MCV, meaning that individuals most likely to die due to COVID-19 usually have higher values of these variables than individuals who survive. In the cases of lymphocytes and platelets the inverse pattern is observed, that is, most of the blue dots are concentrated on the right, and the red dots on the left side, meaning that individuals with a higher probability of dying due to COVID-19 usually tend to have a lower count of these two hematological parameters than surviving individuals. However, it is noteworthy that, although the variables are evaluated individually, their corresponding importance is estimated taking into account the non-obvious interactions among all features

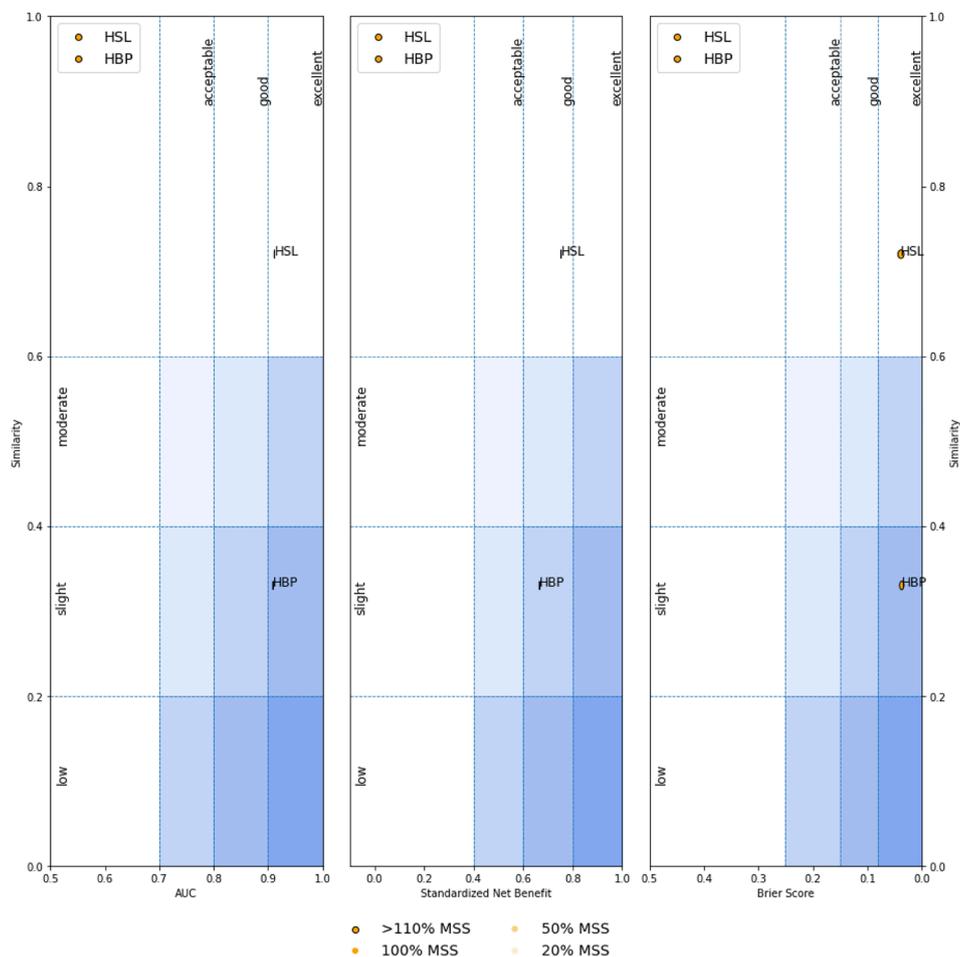


Fig. 5. External performance diagram displays the results of the validations of scenario 4 (train HSL → test HBP) and scenario 5 (train HBP → test HSL). Information about the MSS is rendered in terms of hue brightness. The width of the ellipses is equal to the width of the 95% confidence interval with reference to the given performance metrics.

within the model.

### 3. Discussion

Currently, there are several studies around the world using machine learning technique based on different characteristics easily obtained from patients with COVID-19 [13,18,40] and different outcomes such as need for intensive care, mechanical ventilation and death, among others. Among these characteristics, laboratory test data have also been used in machine learning prediction models, rather than subjective data that could vary between geographic regions, ethnic characteristics, observers and institutions.

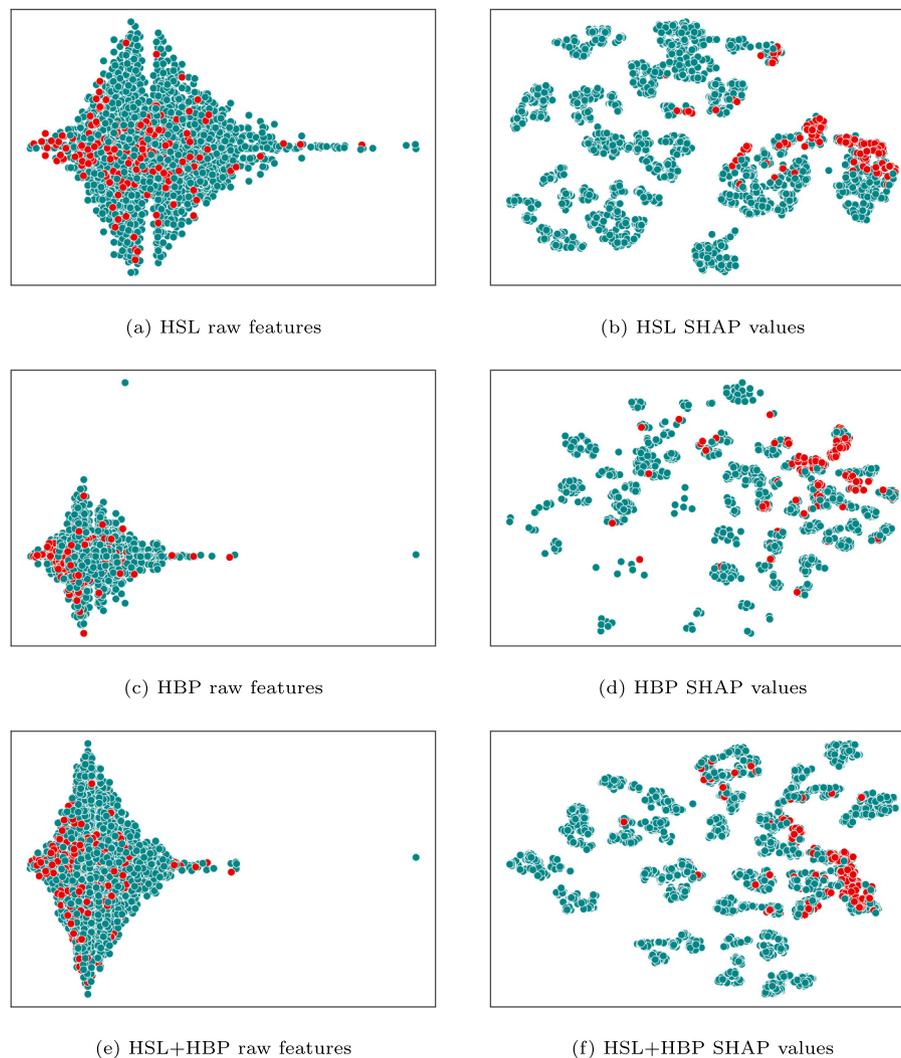
Using machine learning, the present study evaluated routine laboratory variables of patients with COVID-19 treated and/or admitted in two large hospitals in Brazil. The objective was to investigate whether such variables would be able to predict the risk of COVID-19 death in order to assist clinicians to early discriminate those truly at higher risk of death. The proposed machine learning model is composed of five biomarkers, namely RDW, MCV, CRP, lymphocytes and platelets, was reliably externally validated using a diverse external database against the training set, as well as showing sufficient cardinality and good performance in the complementary dimensions of discrimination, calibration, and utility.

As observed in the model explanations (Fig. 5), the concentrations of these five biomarkers may increase or decrease the risk of death due to COVID-19 infection. In an attempt to speculate and summarize the relation between lab variables and possible mechanisms of severe SARS-

CoV-2 infection, we discussed each one of these biomarkers below. We firstly discuss RDW, MCV, and CRP, biomarkers whose higher concentrations may increase the risk of death. Then we discuss lymphocytes and platelets, biomarkers that have the opposite trend.

Bellan et al. (2021) [41] predicted in-hospital mortality by COVID-19 through the analysis of RDW, using a linear cutoff value for defining the prognosis. Similarly, our results showed a greater risk for patients with relatively higher RDW, which indicates a variation in the volume of erythrocytes conventionally known as anisocytosis. [41,42]. Particularly, in the context of inflammation, the cytokine storm that occurs in the most severe cases of COVID-19 [43] may be closely related to the variation in the volume of erythrocytes. As reviewed by Ganz & Nemeth (2015) [44], infectious or inflammatory stimuli account for the characteristic hypoferrremia of inflammation that develops a few hours after systemic infection. Thus, there is a lower iron supply to the bone marrow for hemoglobin production, which ultimately contributes to the reduction of globular volume resulting in microcytosis. In contrast, other factors contribute to the increase in erythrocyte volumes such as comorbidities, folate deficiency, and the use of drugs such as folate inhibitors and others. Such factors, overall, prevail over microcytosis given the increased values of the MCV indicating macrocytosis. Therefore, the coexistence of erythrocytes that became microcytic and macrocytic for different underlying causes and at different times of infection, in our view, partially explains the increase in both RDW and MCV values.

Still referring to RDW, a priori, the increase in RDW already signals the beginning of a disorder in erythropoiesis and/or abnormal red blood



**Fig. 6.** (6a, c and e) t-SNE visualization of the individuals, clustered based on their raw features values (i.e., analytes concentrations in the blood). Red dots represent individuals who died, whereas blue dots represent individuals who were discharged. (6b, d and f) t-SNE visualization of the individuals, clustered based on Shapley values. This visualization represents the ability of our model to separate these two groups of individuals.

cell survival [45], while all other blood count parameters may still be normal. Thus, the RDW measurement is a parameter of considerable clinical importance given its early change. Moreover, the RDW value is being considered a strong and independent risk factor for death in the general population [45]. Intravascular hemolysis may also be an under-recognized complication of COVID-19 [46]. Early reticulocyte release would increase not only MCV but also RDW. The reduction in erythrocyte half-life may be a consequence of the presence of the virus itself as well as of the lung damage resulting from the SARS-CoV-2 infection maximized by the inflammatory process. In fact, RDW has been used as an indicator of ineffective red cell production or hemolysis which has recently been identified as a predictor of poor prognosis in different cardiovascular and noncardiovascular diseases [47]. Hemostatic disorders during severe COVID [48] may also indirectly imply a reduction in erythrocyte survival due to the possible occurrence of thrombotic microangiopathy, defined by a set of manifestations including thrombocytopenia, microangiopathic hemolytic anemia, and multiple organ failure [49].

Our model also found increased CRP levels as an important biochemical variable in the prognosis of COVID-19. CRP levels are increased in a number of inflammatory conditions [50]. Particularly in those with severe COVID, CRP is elevated in 75–93% of them [51]. As reported in a systematic review and meta-analysis [52], the most

common laboratory abnormalities in COVID patients were elevated CRP (68.6%), lymphopenia (57.4%), and elevated lactate dehydrogenase (LDH) (51.6%). Another meta-analysis [53] reported a significant association between lymphopenia, thrombocytopenia and elevated levels of CRP and COVID-19 severity.

Low relative values of lymphocytes and platelets were identified as indicative of poor prognosis in patients infected with SARS-CoV-2 in our models. Lymphopenia has been very frequently observed in patients with COVID-19 infection [54,55] probably indicating a diminished immune response to the virus [56]. It has been described that SARS-CoV-2 infection may primarily affect T lymphocytes [57]. Since T cells are important for dampening overactive innate immune responses during infection [58], their loss during SARS-CoV-2 infection may result in a more severe inflammatory response. According to Guan et al., 2020, the vast majority of patients (82.1%) have experienced SARS-CoV-2-induced peripheral blood lymphopenia suggesting possible pulmonary infiltration of lymphocytes and/or cell damage through apoptosis or pyroptosis [59]. In addition, the expression of ACE2 in lymphocytes turns them into potential targets of SARS-CoV-2, which results in cell death of both CD4+ and CD8+ T cells [57]. In this scenario, the analysis of the lymphocyte count is therefore a reliable indicator of the COVID-19 severity, and really useful in the monitoring and therapeutic decisions. Moreover, after clinical improvement lymphocyte count is corrected

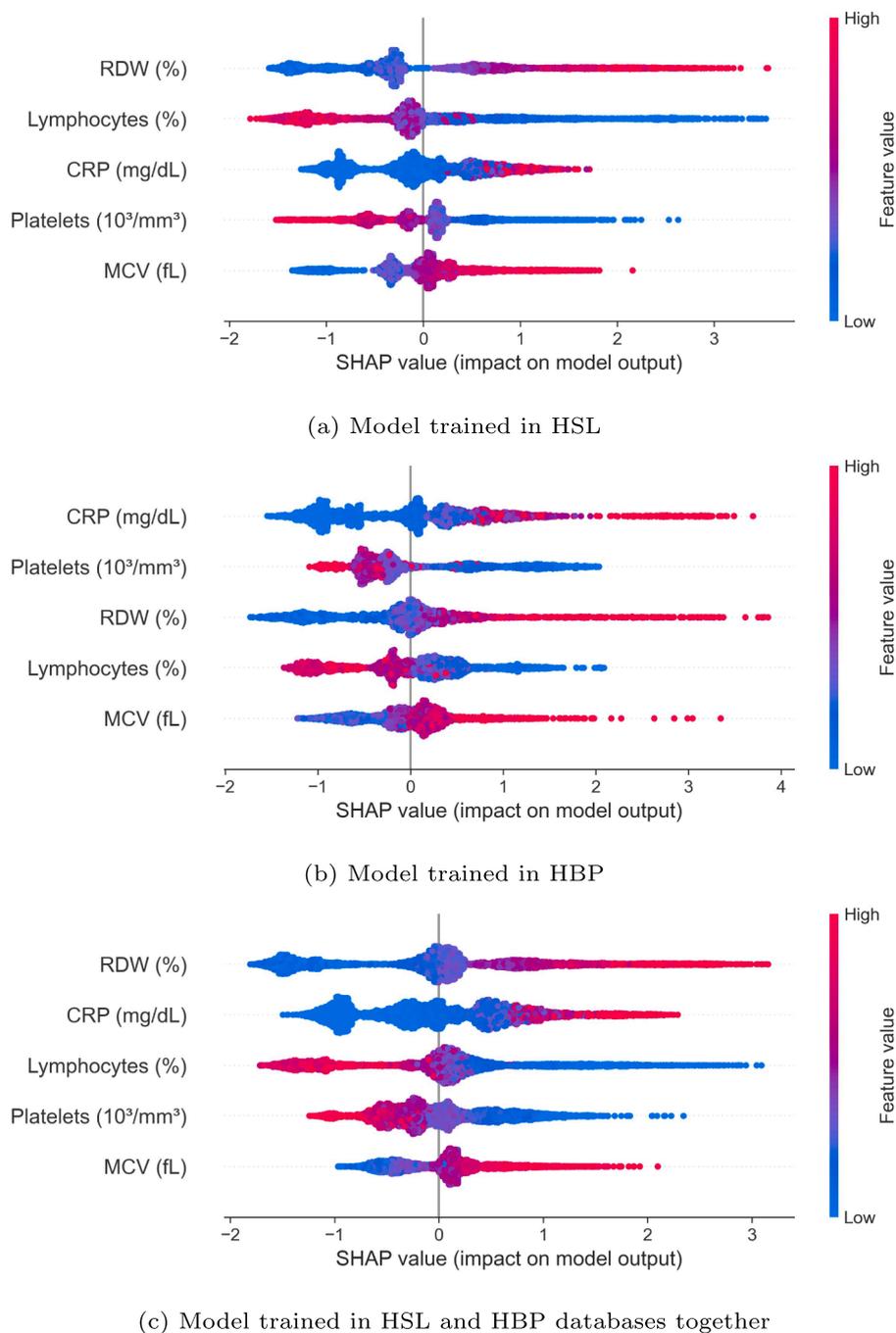


Fig. 7. SHAP summary plots showing the effect of each feature in predicting the risk of death by COVID-19.

[60]. In summary, lymphopenia has been associated with severe COVID-19 which may reflect systemic inflammation and response to pneumonia [61].

As for thrombocytopenia, several authors [62,63] have reported that the drop in the number of platelets may indicate an unfavorable prognostic factor. Three hypotheses related to decreased platelet number and its structure are proposed in severe COVID-19 [64]. Firstly, SARS-CoV-2 infection may reduce platelet production by bone marrow; secondly, there may be increased destruction of platelets by the immune system; and thirdly, platelet consumption due to aggregation in the lungs. As described by Tang et al., 2020, 71.4% of non-survivors had overt disseminated intravascular coagulation (DIC) during their hospitalization compared to only 0.6% of survivors. As is widely known, platelet consumption may occur during DIC implying a reduction in its number

[65]. It is also important to emphasize that the use of drugs to treat COVID-19 may be interfering with megakaryopoiesis, such as azithromycin [66] and hydroxychloroquine [67], causing drug-induced immune thrombocytopenia. Furthermore, heparin-induced thrombocytopenia (HIT), a rare complication of heparin to thrombosis treatment in COVID-19 patients, associated with increased in vivo thrombin generation [68], cannot be ruled out.

To investigate the risk of sex bias, we have evaluated the models considering the sex dimension. We have found that the models perform a little better for females, with an average AUROC of  $0.92 \pm 0.02$ , vs  $0.90 \pm 0.02$  for males. The reason for this difference may come from sex differences in COVID-19 case fatality, given that not only biological factors but also behavioural risk place men at a greater risk for death as a consequence of COVID-19 [69]. However, this idea should be further

investigated.

#### 4. Limitations

Although our study was performed based on thousands of hospitalized patients and was robustly validated in five scenarios, including a reliable external one in terms of data similarity and data cardinality, some limitations should be noted. First, our study did not exclude patients affected by liver, cardiovascular, renal, or malignant diseases, or other previous comorbidities and/or conditions that could bias our study. Second, the used databases were collected before vaccination, and thus the models may not be applicable to vaccinated individuals. Third, the external validation was performed using a dataset from the same country, unlike the study by Cabitza [12], which used datasets from several countries. Finally, it is noteworthy that the results used in the present study were obtained through RT-PCR, the gold standard method for the diagnosis of SARS-CoV-2 [70]. As this method involves the amplification of a small segment of the genetic material of the virus, it has a high specificity. However, the hypothesis of cross-reaction with other viruses that cause acute respiratory syndrome cannot be ruled out.

#### 5. Conclusions

Our findings contribute to consolidating previous studies focused on the prognosis of COVID-19, and highlight the importance of a simple CBC and other routine biomarkers for risk stratification and prediction of in-hospital mortality. In this scenario, we attempted to speculate and summarize possible mechanisms of severe SARS-CoV-2 infection and corresponding altered lab variables in order to assist clinicians to early discriminate those truly at higher risk of death, with a view to reducing mortality. However, it is essential that our algorithm can be tested in other populations in different geographic regions, exploring other contexts, since the natural history of SARS-CoV-2 infection and its consequences on the hematopoietic system and other organs is still quite recent. Finally, from the point of view of clinical contribution, our findings help to optimize the interpretation of data provided by a simple blood count and other routine biomarkers, allowing early risk stratification and prediction of hospital mortality. In light of this knowledge, the adoption of appropriate measures by the medical staff could reduce the mortality of patients with severe COVID-19.

#### Summary Table

What was already known on the topic?.

- Prediction of COVID-19 disease severity is hard to assess.
- Older age, male gender, smoking, genetic factors, and other conditions such as hypertension and obesity are risk factors for severe illness or death by COVID-19.
- Routine biomarkers could help on risk stratification for COVID-19 stratification.

What this study added to our knowledge?.

- Higher values of RDW and/or MCV may increase the risk of death by COVID-19, suggesting abnormal erythropoiesis and/or shorter red blood cell survival due to respiratory failure and systemic inflammation.
- Higher values of CRP may increase the risk of death by COVID-19, suggesting systemic inflammation.
- Lower number of lymphocytes may increase the risk of death by COVID-19, reflecting a more severe inflammatory response likely due to a diminished immune response to the virus;
- Lower number of platelets may increase the risk of death by COVID-19, suggesting a possible occurrence of thrombotic microangiopathy and/or consumption of this cell.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.ijmedinf.2022.104835>.

#### References

- [1] W. Guo, M. Li, Y. Dong, H. Zhou, Z. Zhang, C. Tian, R. Qin, H. Wang, Y. Shen, K. Du, L. Zhao, H. Fan, S. Luo, D. Hu, Diabetes is a risk factor for the progression and prognosis of COVID-19, 2020.
- [2] K.I. Zheng, F. Gao, X.-B. Wang, Q.-F. Sun, K.-H. Pan, T.-Y. Wang, H.-L. Ma, Y.-P. Chen, W.-Y. Liu, J. George, M.-H. Zheng, Letter to the editor: Obesity as a risk factor for greater severity of COVID-19 in patients with metabolic associated fatty liver disease, *Metabolism* 108 (2020) 154244.
- [3] L. Hu, S. Chen, Y. Fu, Z. Gao, H. Long, H.-W. Ren, Y. Zuo, J. Wang, H. Li, Q.-B. Xu, W.-X. Yu, J. Liu, C. Shao, J.-J. Hao, C.-Z. Wang, Y. Ma, Z. Wang, R. Yanagihara, Y. Deng, Risk factors associated with clinical outcomes in 323 coronavirus disease 2019 (COVID-19) hospitalized patients in wuhan, china, *Clin. Infect. Dis.* 71 (16) (2020) 2089–2098.
- [4] COVID-19 Host Genetics Initiative, Mapping the human genetic architecture of COVID-19, *Nature* 600 (2021) 472–477, <https://doi.org/10.1038/s41586-021-03767-x>.
- [5] D. Ding, X. Chen, L. Zhang, M. Zhou, Y. Xu, J. Zhao, Y. Zhou, Y. Wang, Clinical course of severe and critically ill patients with coronavirus disease 2019 (COVID-19): A comparative study, *J Infect* 81 (2020) e82–e84.
- [6] M.C. Castro, S. Kim, L. Barberia, A.F. Ribeiro, S. Gurzenda, K.B. Ribeiro, E. Abbott, J. Blossom, B. Rache, B.H. Singer, Spatiotemporal pattern of COVID-19 spread in Brazil, *Science* 372 (6544) (2021) 821–826.
- [7] H. Ritchie, E. Mathieu, L. Rod s-Guirao, C. Appel, C. Giattino, E. Ortiz-Ospina, J. Hasell, B. Macdonald, D. Beltekian, M. Roser, Coronavirus pandemic (COVID-19), *Our World in Data* (Mar. 2020).
- [8] G. Zuin, D. Araujo, V. Ribeiro, M.G. Seiler, W.H. Prieto, M.C. Pintao, C. Dos Santos Lazari, C.F.H. Granato, A. Veloso, Prediction of SARS-CoV-2-positivity from million-scale complete blood counts using machine learning, *Commun. Med.* 2 (2022) 72.
- [9] N.N. Qomariyah, A. Andi Purwita, S.D. Atas Asri, D. Kazakov, A tree-based mortality prediction model of COVID-19 from routine blood samples, in: 2021 International Conference on ICT for Smart Society (ICISS), IEEE, 2021.
- [10] R. Dabbagh, A. Jamal, M.-H. Tamsah, J.H.B. Masud, M. Titi, Y. Amer, M. Alkubeyyer, T. Alhazmi, F. Baothman, L. Hneiny, Machine learning models for predicting diagnosis or prognosis of COVID-19: A systematic review, *Comput. Methods Programs Biomed.* 205 (2021) 105993.
- [11] L. Yan, H.-T. Zhang, J. Goncalves, Y. Xiao, M. Wang, Y. Guo, C. Sun, X. Tang, L. Jing, M. Zhang, X. Huang, Y. Xiao, H. Cao, Y. Chen, T. Ren, F. Wang, Y. Xiao, S. Huang, X. Tan, N. Huang, B. Jiao, C. Cheng, Y. Zhang, A. Luo, L. Mombaerts, J. Jin, Z. Cao, S. Li, H. Xu, Y. Yuan, An interpretable mortality prediction model for COVID-19 patients, *Nat Mach. Intell* 2 (2020) 283–288.
- [12] A. Carobene, F. Milella, L. Famigliani, F. Cabitza, How is test laboratory data used and characterised by machine learning models? a systematic review of diagnostic and prognostic models developed for COVID-19 patients using only laboratory data, *Clin. Chem. Lab. Med.* (May 2022).
- [13] A.L. Booth, E. Abels, P. McCaffrey, Development of a prognostic model for mortality in COVID-19 infection using machine learning, *Mod. Pathol.* 34 (3) (2020) 522–531.
- [14] J. Luo, L. Zhou, Y. Feng, B. Li, S. Guo, The selection of indicators from initial blood routine test results to improve the accuracy of early prediction of COVID-19 severity, *PLoS One* 16 (6) (2021) e0253329.
- [15] N. Darapaneni, M. Gupta, A.R. Paduri, R. Agrawal, S. Padasali, A. Kumari, P. Purushothaman, A novel machine learning based screening method for high-risk covid-19 patients based on simple blood exams, in: 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), IEEE, 2021.
- [16] S. Aktar, M.M. Ahamad, M. Rashed-Al-Mahfuz, A. Azad, S. Uddin, A. Kamal, S. A. Alyami, P.-I. Lin, S.M.S. Islam, J.M. Quinn, V. Eapen, M.A. Moni, Machine learning approach to predicting COVID-19 disease severity based on clinical blood test data: Statistical analysis and model development, *JMIR Med Inform* 9 (4) (2021) e25884.
- [17] F.T. Fernandes, T.A. de Oliveira, C.E. Teixeira, A.F. de Moraes Batista, G.D. Costa, A.D.P. Filho, A multipurpose machine learning approach to predict COVID-19 negative prognosis in s ao paulo, Brazil, *Sci. Rep.* 11 (1) (2021) 1–7.
- [18] L. Famigliani, A. Campagner, A. Carobene, F. Cabitza, A robust and parsimonious machine learning method to predict ICU admission of COVID-19 patients, *Med. Biol. Eng. Comput.* (2022) 1–13.
- [19] F. Cabitza, A. Campagner, F. Soares, L. Garc a de Guadiana-Romualdo, F. Challa, A. Sulejmani, M. Seghezzi, A. Carobene, The importance of being external. methodological insights for the external validation of machine learning models in medicine, *Computer Methods and Programs in Biomedicine* 208 (2021) 106288.

- doi:<https://doi.org/10.1016/j.cmpb.2021.106288>. URL <https://www.sciencedirect.com/science/article/pii/S016926072100362X>.
- [20] FAPESP, FAPESP COVID-19 Data Sharing/BR, accessed: 2021-5-5 (2020). <https://repositoriodatasharingfapesp.uspdigital.usp.br>.
- [21] R. Pan, T. Yang, J. Cao, K. Lu, Z. Zhang, Missing data imputation by K nearest neighbours based on grey relational structure and mutual information, *Appl. Intell.* 43 (3) (2015) 614–632.
- [22] R. van den Goorbergh, M. van Smeden, D. Timmerman, B. Van Calster, The harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression (2022).
- [23] Y. Elor, H. Averbuch-Elor, To SMOTE, or not to SMOTE?, *arXiv* (2022), doi: 10.48550/arXiv.2201.08528.
- [24] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, Lightgbm: A highly efficient gradient boosting decision tree, *Advances in neural information processing systems* 30 (2017) 3146–3154.
- [25] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, Association for Computing Machinery, New York, NY, USA, 2016*, p. 785–794, doi:10.1145/2939672.2939785.
- [26] B. Schölkopf, J. Platt, T. Hofmann, *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*, MIT Press, 2007.
- [27] T. Fawcett, An introduction to ROC analysis, *Pattern Recognit. Lett.* 27 (2006) 861–874, doi:10.1016/j.patrec.2005.10.010.
- [28] F. Cabitza, A. Campagner, The need to separate the wheat from the chaff in medical informatics: Introducing a comprehensive checklist for the (self)-assessment of medical AI studies, *Int J Med Inform* 153 (2021) 104510, <https://doi.org/10.1016/j.ijmedinf.2021.104510>.
- [29] A.A. Bradley, A. Allen Bradley, S.S. Schwartz, T. Hashino, Sampling uncertainty and confidence intervals for the brier score and brier skill score, *Weather Forecast.* 23 (2008) 992–1006.
- [30] F. Cabitza, A. Campagner, L.M. Sconfienza, As if sand were stone. New concepts and metrics to probe the ground on which to build trustable AI, *BMC Med. Inform. Decis. Mak.* 20 (1) (2020) 219.
- [31] Y. Sasaki, Yutaka, The truth of the F-measure, *Teach Tutor, Mater* (2007).
- [32] R. Sahoo, S. Zhao, A. Chen, S. Ermon, Reliable decisions with threshold calibration, *Adv. Neural Inf. Process. Syst.* 34 (2021) 1831–1844.
- [33] C.M. Michelon, Main SARS-CoV-2 variants notified in Brazil, *RBAC* 53 (2) (2021).
- [34] L. van der Maaten, G. Hinton, Visualizing data using t-SNE, *J. Mach. Learn. Res.* 9 (86) (2008) 2579–2605.
- [35] Y. Cao, L. Wang, Automatic selection of t-SNE perplexity, *ArXiv* (2017).
- [36] Difference between PCA VS t-SNE, <https://www.geeksforgeeks.org/difference-between-pca-vs-t-sne/>, accessed: 2022-7-12 (May 2020).
- [37] M. Wattenberg, F. Viégas, I. Johnson, How to use t-sne effectively, *Distill* (2016), <https://doi.org/10.23915/distill.00002>. URL <http://distill.pub/2016/misread-tsne>.
- [38] R.F. Berry, J.L. Hellerstein, A unified approach to interpreting measurement data in performance management applications, *Proceedings of 1993 IEEE 1st Int. Workshop Syst. Man.* (1993) 81–89, doi: 10.1109/IWSM.1993.315286.
- [39] D. Kobak, P. Berens, The art of using t-SNE for single-cell transcriptomics, *Nat. Commun.* 10 (1) (2019) 5416.
- [40] H.-J. Kim, D. Han, J.-H. Kim, D. Kim, B. Ha, W. Seog, Y.-K. Lee, D. Lim, S.O. Hong, M.-J. Park, J. Heo, An easy-to-use machine learning model to predict the prognosis of patients with covid-19: Retrospective cohort study, *J Med Internet Res* 22 (11) (2020) e24225, <https://doi.org/10.2196/24225>. URL <http://www.jmir.org/2020/11/e24225/>.
- [41] M. Bellan, D. Azzolina, E. Hayden, G. Gaidano, M. Pirisi, A. Acquaviva, G. Aimaretti, P. Aluffi Valletti, R. Angilletta, R. Arioli, G.C. Avanzi, G. Avino, P.E. Balbo, G. Baldon, F. Baorda, E. Barbero, A. Baricich, M. Barini, F. Barone-Adesi, S. Battistini, M. Beltrame, M. Bertoli, S. Bertolin, M. Bertolotti, M. Betti, F. Bobbio, P. Boffano, L. Boggione, S. Borrè, M. Bruccoli, E. Calzaducça, E. Cammarata, V. Cantaluppi, R. Cantello, A. Capponi, A. Carriero, G.F. Casciaro, L.M. Castello, F. Ceruti, G. Chichino, E. Chirico, C. Cisari, M.G. Cittone, C. Colombo, C. Comi, E. Croce, T. Daffara, P. Danna, F. Della Corte, S. De Vecchi, U. Dianzani, D. Di Benedetto, E. Esposito, F. Faggiano, Z. Falaschi, D. Ferrante, A. Ferrero, I. Gagliardi, A. Galbiati, S. Gallo, P.L. Garavelli, C.A. Gardino, M. Garzaro, M.L. Gastaldello, F. Gavelli, A. Gennari, G.M. Giacomini, I. Giacone, V. Giai Via, F. Giolitti, L.C. Gironi, C. Gramaglia, L. Grisafi, I. Inserra, M. Invernizzi, M. Krengli, E. Labella, I.C. Landi, R. Landi, I. Leone, V. Lio, L. Lorenzini, A. Maconi, M. Malerba, G.F. Manfredi, M. Martelli, L. Marzari, P. Marzullo, M. Mennuni, C. Montabone, U. Morosini, M. Mussa, I. Nericci, A. Nuzzo, C. Olivieri, S.A. Padelli, M. Panella, A. Parisini, A. Pasché, F. Patrucco, G. Patti, A. Pau, A.R. Pedrinelli, I. Percivale, L. Ragazzoni, R. Re, C. Riganonti, E. Rizzi, A. Rognoni, A. Roveta, L. Salamina, M. Santagostino, M. Saraceno, P. Savoia, M. Sciarra, A. Schimmenti, L. Scotti, E. Spinoni, C. Smirne, V. Tarantino, P.A. Tillio, S. Tonello, R. Vaschetto, V. Vassia, D. Zagaria, E. Zavattaro, P. Zeppegno, F. Zottarelli, P.P. Sainaghi, Simple parameters from complete blood count predict In-Hospital mortality in COVID-19, *Dis. Markers* 2021 (May 2021).
- [42] M. Bellan, A. Giubertoni, C. Piccinino, A. Dimaghi, F. Grimoldi, M. Sguazzotti, M. E. Burlone, C. Smirne, D. Sola, P. Marino, M. Pirisi, P.P. Sainaghi, Red cell distribution width and platelet count as biomarkers of pulmonary arterial hypertension in patients with connective tissue disorders, *Dis. Markers* 2019 (2019).
- [43] B. Hu, S. Huang, L. Yin, The cytokine storm and COVID-19, *J. Med. Virol.* 93 (1) (2021) 250–256.
- [44] T. Ganz, E. Nemeth, Iron homeostasis in host defence and inflammation, *Nat. Rev. Immunol.* 15 (8) (2015) 500–510.
- [45] G.L. Salvagno, F. Sanchis-Gomar, A. Picanza, G. Lippi, Red blood cell distribution width: A simple parameter with multiple clinical applications, *Crit. Rev. Clin. Lab. Sci.* 52 (2) (2015) 86–105.
- [46] G. Lancman, B.K. Marcellino, S. Thibaud, K. Troy, Coombs-negative hemolytic anemia and elevated plasma hemoglobin levels in COVID-19, *Ann. Hematol.* 100 (3) (2021) 833–835.
- [47] L.L. Abraham IV, J.D.A. Ramos, E.L. Cunanán, M.D.A. Tiongson, F.E.R. Punzalan, Red cell distribution width and mortality in patients with acute coronary syndrome: A Meta-Analysis on prognosis, *Cardiol. Res. Pract.* 9 (3) (2018) 144–152.
- [48] T. Iba, J.H. Levy, M. Levi, J.M. Connors, J. Thachil, Coagulopathy of coronavirus disease 2019, *Crit. Care Med.* 48 (9) (2020) 1358–1364.
- [49] R. Lopes da Silva, Viral-associated thrombotic microangiopathies, *Hematol. Oncol. Stem Cell Ther.* 4 (2) (2011) 51–59.
- [50] M.B. Pepys, G.M. Hirschfield, C-reactive protein: a critical update, *J. Clin. Invest.* 111 (12) (2003) 1805–1812.
- [51] G. Lippi, M. Plebani, Laboratory abnormalities in patients with COVID-2019 infection, *Clin. Chem. Lab. Med.* 58 (7) (2020) 1131–1134.
- [52] L. Fu, B. Wang, T. Yuan, X. Chen, Y. Ao, T. Fitzpatrick, P. Li, Y. Zhou, Y.-F. Lin, Q. Duan, G. Luo, S. Fan, Y. Lu, A. Feng, Y. Zhan, B. Liang, W. Cai, L. Zhang, X. Du, L. Li, Y. Shu, H. Zou, Clinical characteristics of coronavirus disease 2019 (COVID-19) in China: A systematic review and meta-analysis, *J. Infect.* 80 (6) (2020) 656–665.
- [53] P. Malik, U. Patel, D. Mehta, N. Patel, R. Kelkar, M. Akrmah, J.L. Gabrilove, H. Sacks, Biomarkers and outcomes of COVID-19 hospitalisations: systematic review and meta-analysis, *BMJ Evid Based Med* 26 (3) (2021) 107–108.
- [54] M. Ozbalak, S.K. Besisik, Y.B. Tor, A. Medetalibeyoglu, M. Kose, N. Senkal, E. Aksoy, A. Cagatay, M. Erelel, A. Gul, F. Esen, S.S. Yavuz, U.I. Alkac, T. Tukek, Initial complete blood count score and predicting disease progression in COVID-19 patients, *Am. J. Blood Res.* (2021) 77–83.
- [55] J.L. Frater, G. Zini, G. d'Onofrio, H.J. Rogers, COVID-19 and the clinical hematology laboratory, *Int. J. Lab. Hematol.* 42 (Suppl 1) (2020) 11–18.
- [56] G. Lippi, M. Plebani, The critical role of laboratory medicine during coronavirus disease 2019 (COVID-19) and other viral outbreaks, *Clin. Chem. Lab. Med.* 58 (7) (2020) 1063–1069.
- [57] G. Chen, D. Wu, W. Guo, Y. Cao, D. Huang, H. Wang, T. Wang, X. Zhang, H. Chen, H. Yu, X. Zhang, M. Zhang, S. Wu, J. Song, T. Chen, M. Han, S. Li, X. Luo, J. Zhao, Q. Ning, Clinical and immunological features of severe and moderate coronavirus disease 2019, *J. Clin. Invest.* 130 (5) (2020) 2620–2629.
- [58] K.D. Kim, J. Zhao, S. Auh, X. Yang, P. Du, H. Tang, Y.-X. Fu, Adaptive immune cells temper initial innate responses, *Nat. Med.* 13 (10) (2007) 1248–1252.
- [59] G. Huang, A.J. Kovalic, C.J. Graber, Prognostic value of leukocytosis and lymphopenia for coronavirus disease severity - volume 26, number 8—august 2020 - emerging infectious diseases journal - CDC.
- [60] F. Wang, J. Nie, H. Wang, Q. Zhao, Y. Xiong, L. Deng, S. Song, Z. Ma, P. Mo, Y. Zhang, Characteristics of peripheral lymphocyte subset alteration in COVID-19 pneumonia, *J. Infect. Dis.* 221 (11) (2020) 1762–1769.
- [61] Q. Zhao, M. Meng, R. Kumar, Y. Wu, J. Huang, Y. Deng, Z. Weng, L. Yang, Lymphopenia is associated with severe coronavirus disease 2019 (COVID-19) infections: A systemic review and meta-analysis, *Int. J. Infect. Dis.* 96 (2020) 131–135.
- [62] S.-Q. Jiang, Q.-F. Huang, W.-M. Xie, C. Lv, X.-Q. Quan, The association between severe COVID-19 and low platelet count: evidence from 31 observational studies involving 7613 participants, *Br. J. Haematol.* 190 (1) (2020) e29–e33.
- [63] G. Lippi, M. Plebani, B.M. Henry, Thrombocytopenia is associated with severe coronavirus disease 2019 (COVID-19) infections: A meta-analysis, *Clin. Chim. Acta* 506 (2020) (2019) 145–148.
- [64] P. Xu, Q. Zhou, J. Xu, Mechanism of thrombocytopenia in COVID-19 patients, *Ann. Hematol.* 99 (6) (2020) 1205–1208.
- [65] N. Tang, D. Li, X. Wang, Z. Sun, Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia, *J. Thromb. Haemost.* 18 (4) (2020) 844–847.
- [66] M.U. Butt, A. Jabri, S.C. Elayi, Azithromycin-Induced thrombocytopenia: A rare etiology of Drug-Induced immune thrombocytopenia, *Case Rep. Med.* 2019 (2019) 6109831.
- [67] D. Demir, F. Öcal, M. Abanoz, H. Dermenci, A case of thrombocytopenia associated with the use of hydroxychloroquine following open heart surgery, *Int. J. Surg. Case Rep.* 5 (12) (2014) 1282–1284.
- [68] M. Sartori, B. Cosmi, Heparin-induced thrombocytopenia and COVID-19, *Hematol. Rep.* 13 (1) (2021) 8857.
- [69] C. Wenham, J. Smith, R. Morgan, Gender and COVID-19 Working Group, COVID-19: the gendered impacts of the outbreak, *Lancet* 395 (10227) (2020) 846–848.
- [70] P. Habibzadeh, M. Mofatteh, M. Silawi, S. Ghavami, M.A. Faghghi, Molecular diagnostic assays for COVID-19: an overview, *Crit. Rev. Clin. Lab. Sci.* 58 (6) (2021) 385–398.