CORESMA - COVID-19-Outbreak Response combining E-health, Serolomics, Modelling, Artificial Intelligence and Implementation Research

WP 3 Mathematical modelling of spread and control of COVID-19

WP 3	**Mathematical modelling of spread and control of COVID-19**
Deliverable D3.2	**Report**
Title of Deliverable:	**Working AI software for risk categorization for Corona virus patients**
Author:	**HZI**
Authors:	**Arnab Bandyopadhyay, Niklas Schwan, Sebastian Binder, Michael Meyer-Hermann**
Affiliations:	**Department of Systems Immunology and Braunschweig Integrated Centre of Systems Biology, Helmholtz Centre for Infection Research, Braunschweig, Germany;**

The goal and expected outcome of engaging AI/ML

Similar to many viruses, SARS-CoV-2 mutates over time. Such mutations may have different levels of impact on the virus' properties and activities, which can be reflected in clinical parameters of the patients like, for example, disease severity and associated expression of clinical markers. This relationship might get heavily skewed depending on the age, underlying medical conditions, and vaccination state of the patients. The goal here was to link the expression of biomarkers with disease severity. We wanted to use Machine learning tools to identify the association of biomarkers with age, vaccination, and health conditions in order to allow for a more stratified treatment of the patients.

Research questions and intended use of the AI/ML model

With the LEOSS dataset, we ask the following:

1. We want to understand the relation between the different variants of SARS-CoV-2 and its clinical outcomes on patients by comparing the clinical data of patients in different waves of the pandemic by assuming the following coordinates:

2. We want to determine the predictive power of the expression of biomarkers for disease severity by considering patients who transition from an uncomplicated disease state (uncomplicated phase – UC) to more severe disease states (complicated phase – CO; Critical phase – CR).

3. It has been widely observed in many disease contexts that the underlying medical conditions and disease type might alter the clinical parameters and patients' fate. Therefore, we would like to compare the disease severity across multiple axes shown above. For example, we want to compare biomarker expression between healthy uncomplicated vs healthy critical, or comorbid uncomplicated vs comorbid critical.

Qualitative and quantitative details on training dataset (and links to the data if possible)

Study population

A link to the data is not possible to provide for legal reasons.

The study included a total of >10,000 patients who were enrolled in the LEOSS (Lean European Open Survey on SARS-CoV-2) registry. These patients were recruited from March 18, 2020, from 122 hospitals across Europe, with the majority of them located in Germany. All patients were diagnosed with SARS-CoV-2 through positive results from PCR testing. Data collection was conducted retrospectively and anonymously, focusing on the documentation of standard of care treatment. Only patients with information available on follow-up and at the end of the treatment (recovery or death) were included in the analysis.

Upon the initial positive test result for SARS-CoV-2 and at the first point of reference with the hospitals, the patients were divided into different groups based on their clinical condition. Upon initial presentation, a patient can fall into one of three states: Uncomplicated (UC), Complicated (CO), or Critical (CR). Biomarkers were for transitioning from one state to another were evaluated.

Those in the uncomplicated phase (UC) either showed no symptoms or presented symptoms such as upper respiratory tract infection, fever, nausea, vomiting, or diarrhea. Patients in the complicated phase (CO) exhibited at least one of the following characteristics: newly requiring oxygen supplementation, need for unplanned mechanical ventilation, a clinically significant increase in prior oxygen home therapy, PaO2 (partial pressure of oxygen) at room air below 70 mmHg, SO2 (oxygen saturation) at room air below 90%, an increase in AST or ALT (liver enzymes) greater than 5 times the upper limit of normal, new cardiac arrhythmia, new pericardial effusion larger than 1 cm, new heart failure with pulmonary edema, congestive hepatopathy, or peripheral edema.

Patients in the critical phase (CR) required catecholamine support, experienced life-threatening cardiac arrhythmia, received mechanical ventilation (either invasive or non-invasive), or needed an unplanned prolongation of mechanical ventilation for more than 24 hours beyond the initially planned duration. Additionally, patients with liver failure characterized by an INR (international normalized ratio) greater than 3.5 or a quick value below 50%, a qSOFA score of 2 or higher, or acute renal failure requiring dialysis were also classified in the critical phase.

Patients are considered in recovery phase (RC) if improvement by one degree of severity according to this scheme, discharged from the hospital, no further progression of disease or re-hospitalisation.

The study received ethical approval from the responsible ethics committee at each participating study site. More information is available at https://leoss.net/

Clinical data collection

Demographic, clinical, laboratory, treatment, and outcome data were obtained from the medical records of hospitalized patients. The study examined various patient characteristics, including age, gender, BMI, smoking status, comorbidities, and the intake of pre-existing medications such as Angiotensin-converting-enzyme inhibitors (ACE inhibitors), angiotensin II type 1 (AT1) receptor antagonists, statins, ibuprofen, and immunosuppressants. The latter category includes medications used for the treatment of cancer, rheumatologic/inflammatory diseases, chronic inflammatory bowel disease, multiple sclerosis, and similar conditions.

Vital parameters such as body temperature, pulse rate, respiratory rate, oxygen saturation, and oxygen and carbon dioxide partial pressures were also examined. In addition, laboratory parameters related to acute phase responses and organ functions, as well as soluble inflammation markers, were analysed.

The parameters were recorded in categorical form. Hematologic biomarkers analysed in this study include WBC count, lymphocyte count, neutrophil count, fibrinogen, ferritin,

platelet count, haemoglobin, C-reactive protein (CRP), procalcitonin (PCT), IL-6, D-dimer, troponin, creatine kinase (CK), aspartate aminotransferase (AST) etc.

Comorbidities were divided into two categories: present or absent. The comorbidities were further classified as follows:

1. Cardiovascular disease: This category included conditions such as myocardial infarction, aortic stenosis, atrioventricular (AV) block, carotid arterial disease, chronic heart and circulation failure, peripheral vascular disease, hypertension, atrial fibrillation, and coronary artery disease.

2. Pulmonary disease: This category encompassed chronic obstructive pulmonary disease, asthma, and other chronic pulmonary diseases.

3. Hematological/oncological disease: This category included conditions such as leukemia, lymphoma, solid tumors, and stem cell transplantation.

4. Diabetes mellitus: This category was further subdivided into diabetes with and without end-organ damage.

5. Kidney disease: This category included acute kidney injury at the time of SARS-CoV-2 detection and chronic kidney disease.

6. Neurological diseases: This category comprised conditions such as hemiplegia, dementia, cerebrovascular disease, stroke, transient ischemic attack, motoneuron diseases, movement disorder, multiple sclerosis, myasthenia gravis, neuromyelitis optica spectrum disorder, other neurological autoimmune diseases, and other prior neurological diagnoses.

7. Other comorbidities: This category encompassed connective tissue disease, peptic ulcer disease, chronic liver disease, liver cirrhosis, organ transplantation, rheumatic disease, and HIV/AIDS.

A comorbidity was considered present if at least one specific comorbidity within a category was documented. If all specific comorbidities within a category were unknown or missing, the comorbidities were categorised as unknown/missing. Values recorded as unknown were defined as missing.

Detailed description and justification of the models chosen

Ensemble methods are widely used in machine learning to improve the overall performance and robustness of models by combining the predictions of multiple individual models. There are several benefits of applying an ensemble method when the data is very noisy. This method can reduce overfitting, which occurs when a model performs well on the training data but poorly on new, unseen data. This method is robust as it is based on ensemble decisions. Individual models may make errors on specific instances, but the ensemble decision, which aggregates predictions, tends to be more stable and less affected by individual outliers or noisy observations. For the same reason, it helps to reduce the

variance between the individual models and thereby, is effective at capturing complex relationships in the data.

Given the diverse physiological characteristics among over 10,000 patients in the LEOSS dataset, it is strongly advisable to employ an ensemble-based approach.

Here we used the following ensemble-based machine learning algorithms: Random Forest and XGBoost (Extreme Gradient Boosting) for classification and regression tasks. In simple words, Random Forest employs an ensemble method based on bagging (Bootstrap Aggregating) where each tree is built independently of others. In XGBoost, an ensemble method based on boosting is employed where trees are built sequentially, with each tree correcting the errors of the previous ones. Both Random Forest and XGBoost are robust and perform well on a variety of datasets and therefore well-suited for a wide range of tasks. However, XGBoost often provides better predictive performance.

Details on software implementation and links to software

All the codes developed to analyse the LEOSS dataset are now uploaded in a github public repository and can be accessed through the folowing link:
https://github.com/arnabbandyopadhyay/CORESMA_LEOSS

Machine Learning (ML) models require that all data passed as input are in numeric form. The dataset is pre-processed which assigns a number to every unique categorical value in the column. The dataset contains multiple missing values which cause an error when passed directly as an input. Thus, we removed all the missing values and started with a clean data set.

The purpose of this analysis is to accurately predict the disease state of a patient based on the biomarker expression and rank the biomarkers accordingly by their overall score. Since the ranking is crucial for the project goals, accuracy is important. Thus, for the purpose of evaluating the model we considered two evaluation metrics, accuracy and kappa, for this study.

The following terms are used in the equations: TP, True Positive; TN, True Negative; FP, False Positive; and FN, False Negative.

Accuracy

Given a dataset consisting of (*TP* + *TN*) data points, the accuracy is equal to the ratio of total correct predictions (*TP* + *TN* + *FP* + *FN*) by the classifier to the total data points. Accuracy is used to assess the performance of the classification model. Accuracy is calculated as shown below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} 0.0 < Accuracy < 1.0$$

Kappa

The kappa statistic is a measure that assesses the agreement between an observed accuracy and an expected accuracy, which represents the agreement expected by chance alone. It is commonly used to evaluate the performance of classifiers, not only individually but also in comparison to each other. The kappa statistic takes into account both the observed agreement and the agreement expected by chance, providing a more comprehensive evaluation of classifier performance. The value of Cohen's kappa ranges between -1 and 1, with higher values indicating stronger agreement.

Precision

Precision is equal to the ratio of the True Positive (*TP*) samples to the sum of True Positive (*TP*) and False Positive (*FP*) samples. Precision is also a key metric to identify the number of correctly classified patients in an imbalanced class dataset.

We have used the pre-processed dataset to train multiple ML classification models. The models included in this study include: Random Forest Classifier, XGBTree. We split the dataset into 70:30 ratio for training and testing purposes.

10-fold cross-validation

We implemented a 10-fold cross-validation technique, which is the most common procedure and a recommended process to check the generalization ability of the model. This is usually performed to avoid over-fitting during the training process. To perform 10-fold cross-validation, data were first partitioned into 10 nearly equal-sized segments. After partitioned, 10 iterations of training and validation are performed such that, within each iteration, a different segment of the data is held-out for validation while the remaining 9 segments are used for learning and subsequently the learned models are used to make predictions about the data in the validation segment. Therefore, each time a model is trained and tested with an ‘‘unseen’’ dataset. An averaging can be used to obtain an aggregate measure from these samples.

Predictors of disease state: Uncomplicated, Complicated and Critical:

Initially, our goal was to assess which biomarkers could serve as effective classifiers among the disease states UC, CO, and CR. Notably, patient information was recorded upon positive RT-PCR test, and the timing of this test varied among individuals, occurring at different days from the onset of symptoms. This introduces inherent variation within the disease states (UC, CO, and CR) due to measurements conducted at diverse points in the symptom onset timeline.

For our analysis, we specifically focused on patients who initially presented in the UC phase, subsequently developed severe symptoms, and progressed to the CO and CR phases. We considered RF and XGB method for the analysis. In Figure 1, we present the predictions from various machine learning models. Our findings highlight that Neutrophil level 12, CRP level 7, and D-dimer level 6 emerge as reliable predictors for distinguishing between different disease states. Neutrophil/CRP/Ddimer level 12/7/6 represents neutrophil/CRP/Ddimer count in blood *>= 9,000 /µL , > 249 mg/L, 180 - 249 mg/L,* respectively.

In below we show AUC-ROC (Area Under the Receiver Operating Characteristic curve) curve that plots true positive rate (sensitivity) against the false positive rate (1-specificity). This is commonly performed to evaluate the accuracy of a machine learning model. An AUC-ROC score of 1.0 represents a perfect classifier, while a score of 0.5 suggests that the model performs no better than random guess.

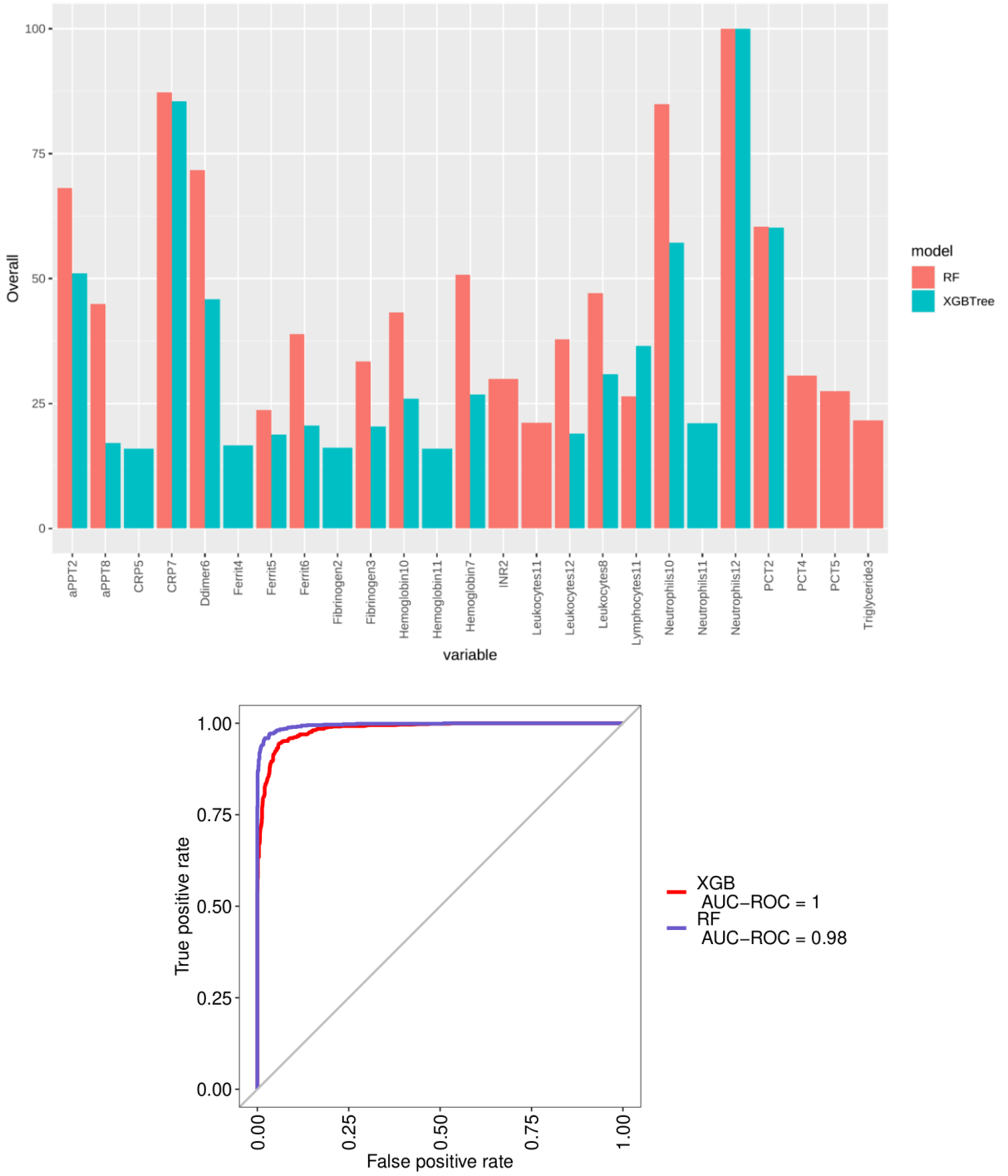


Figure 1: The machine learning prediction of the relative significance of biomarkers, serving as classifiers for distinguishing between Uncomplicated (UC), Complicated (CO), and Critical

(CR) phases. Our analysis exclusively focused on patients initially in the UC phase who subsequently progressed to more severe stages, reaching CO and CR levels. Prediction outcomes from the Random Forest and XGBoost (XGB) methods are denoted by red and cyan, respectively. Our model indicates that Neutrophil levels of 12 and 10, CRP level of 7, Ddimer level of 6, and aPPT level of 2 can effectively serve as classifiers for distinguishing among the different disease states. In the below we plot AUC-ROC, a performance metric commonly used to evaluate the accuracy of a machine learning model.

To check the ML predictions, we analysed the biomarker expressions across different disease states in Figure 8 and across different age groups in Figure 9.

Predicting critical state:

We then asked what are the biomarkers that will predict the risk of patients reaching critical state. We used XGB in this context. In Figure 2 we show the mean accuracy decrease (and associated error) that measures the reduction in average accuracy observed in a machine learning model when a specific feature is excluded, serving as an indicator of the significance of that feature in influencing overall predictive performance. We found Ferrit, CRP, Neutrophils are good predictors. At right we show the performance of the classifier in the form of confusion matrix, representing true positive, true negative, false positive, and false negative values derived from known data.

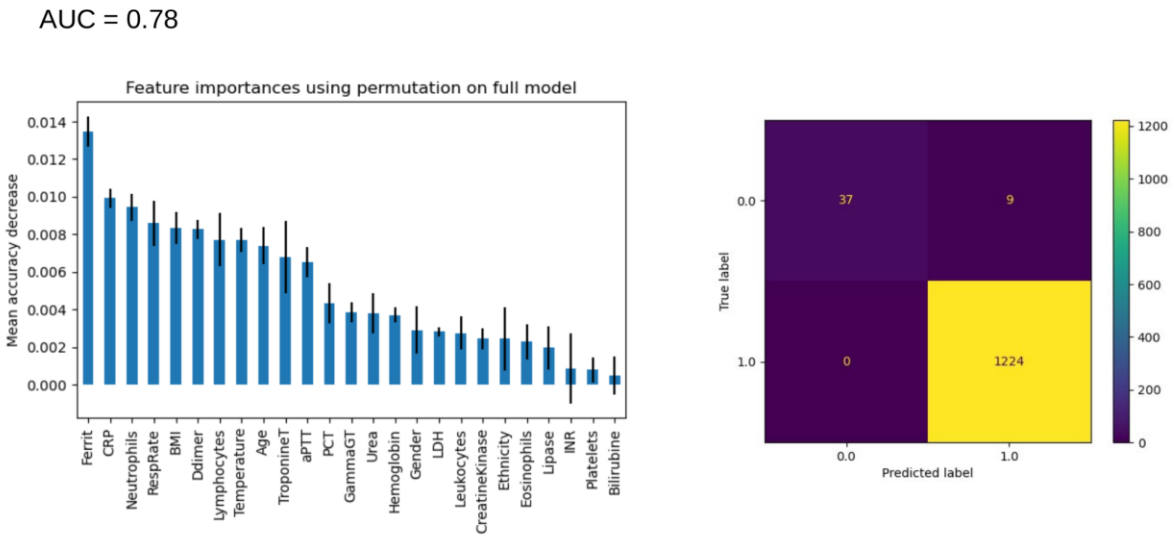


Figure 2: The mean accuracy decrease against individual features revealing that Ferrit, CRP and Neutrophil have the most significant impact. At right we plot the performance of the clssifier in form of a confusion matrix.

Healthy Uncomplicated vs Healthy Critical

We investigated what are the biomarkers that are expressed differentially in Healthy (no comorbidity) patients and contribute to the critical state. We considered only the healthy patients who were at uncomplicated phase and subsequently developed severity in

symptoms which is representitive of a CR state. Using RF and XGB model we represent relative significance of biomarkers in Figure 3. We found CRP level 7, Neutrophil level 10 and Ddimer level 6 can effectively serve as classifiers for distinguishing these two groups. In the left below panel we showed the Alpha and Kappa scores of these two models, which are commonly used to evaluate the performance of ML models. The alpha score is associated with overall accuracy, while the kappa score assesses the agreement between predicted and observed classifications, considering the possibility of chance agreement. At the right panel we show AUC-ROC curve that plots true positive rate (sensitivity) against the false positive rate (1-specificity). To check ML model predictions, we compare biomarker expression in healthy and comorbid patients in Figure 11 and across different age groups in Figure 12.

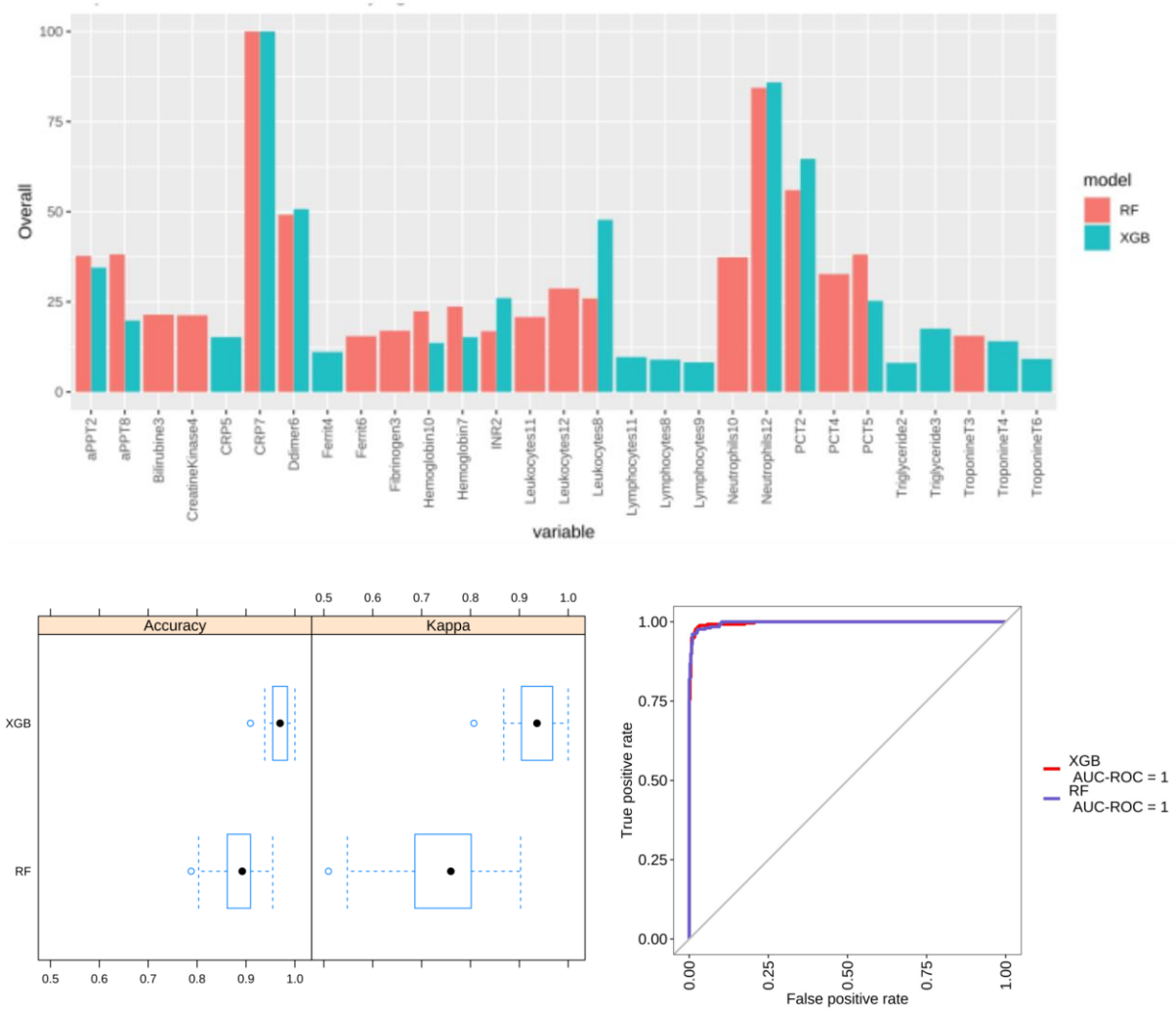


Figure 3: The machine learning prediction of the relative significance of biomarkers, serving as classifiers for distinguishing between healthy uncomplicated and healthy critical phases. Prediction outcomes from the Random Forest and XGBoost (XGB) methods are denoted by red and cyan, respectively. Our model indicates that Neutrophil levels of 12, CRP level of 7, Ddimer level of 6, and PCT level of 2 can effectively serve as classifiers for these two groups. In the left panel below we show Accuracy and Kappa score for these two models. At

right we show AUC-ROC, a performance metric commonly used to evaluate the accuracy of a machine learning model.

Comorbid Uncomplicated vs Comorbid Critical

We investigated what are the biomarkers that are expressed differentially in comorbid patients and predictor of the critical state. Similar to the previous analysis, we considered only the comorbid patients (for definition see study population section) who were at uncomplicated phase and subsequently developed severe symptoms and represents CR state. Using RF and XGB model we represent relative significance of biomarkers in Figure 4. We found CRP level 7, Neutrophil level 12 and Ddimer level 6 can effectively serve as classifiers for distinguishing these two groups. To check ML model predictions, we compare biomarker expression in healthy and comorbid patients in Figure 11 and across different age groups in Figure 12.

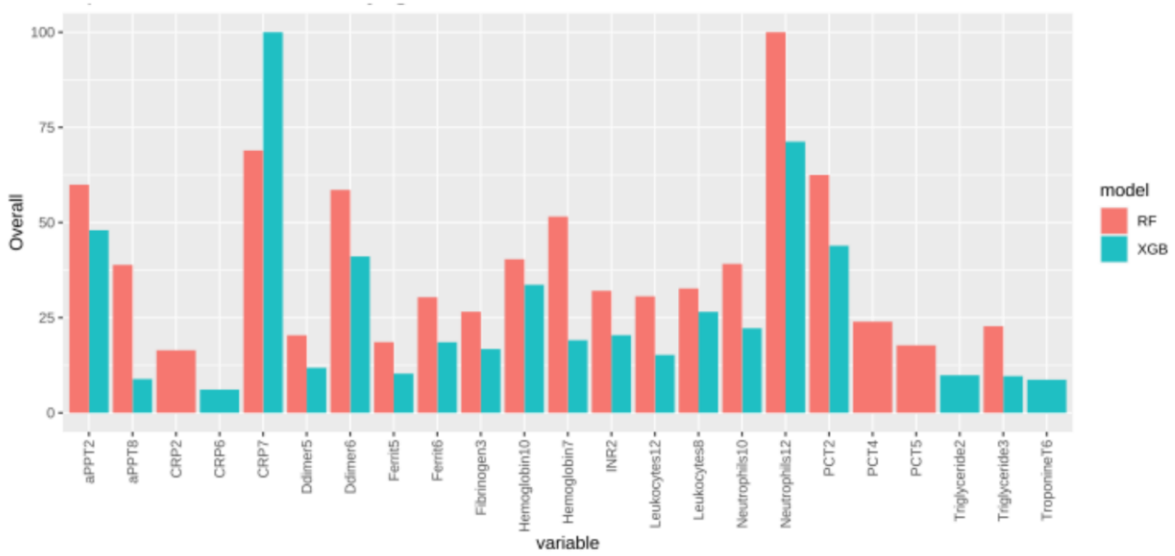


Figure 4: The machine learning prediction of the relative significance of biomarkers, serving as classifiers for distinguishing between comorbid uncomplicated and comorbid critical phases. Prediction outcomes from the Random Forest and XGBoost (XGB) methods are denoted by red and cyan, respectively. Our model indicates that Neutrophil levels of 12, CRP level of 7, Ddimer level of 6, and aPPT level of 2 can effectively serve as classifiers for distinguishing between these two groups.

Differential expression in Complicated patients infected with wt, alpha and other variants

In this context, we considered exclusively comlicated patients infected with the Wt, alpha, or other variants of concern (VoC). We are exploring the differential expression of biomarkers within the cohort of complicated patients, specifically examining how these expressions vary when infected by different virus variants. We use RF and XGB model to analyse the biomarkers and we showed Alpha and Kappa score of these two models in Figure 5. A low alpha and kappa value indicates a lower level of overall accuracy and agreement between

predicted and observed data. Therefore the reliability of the model predictions is considered to be insufficient.

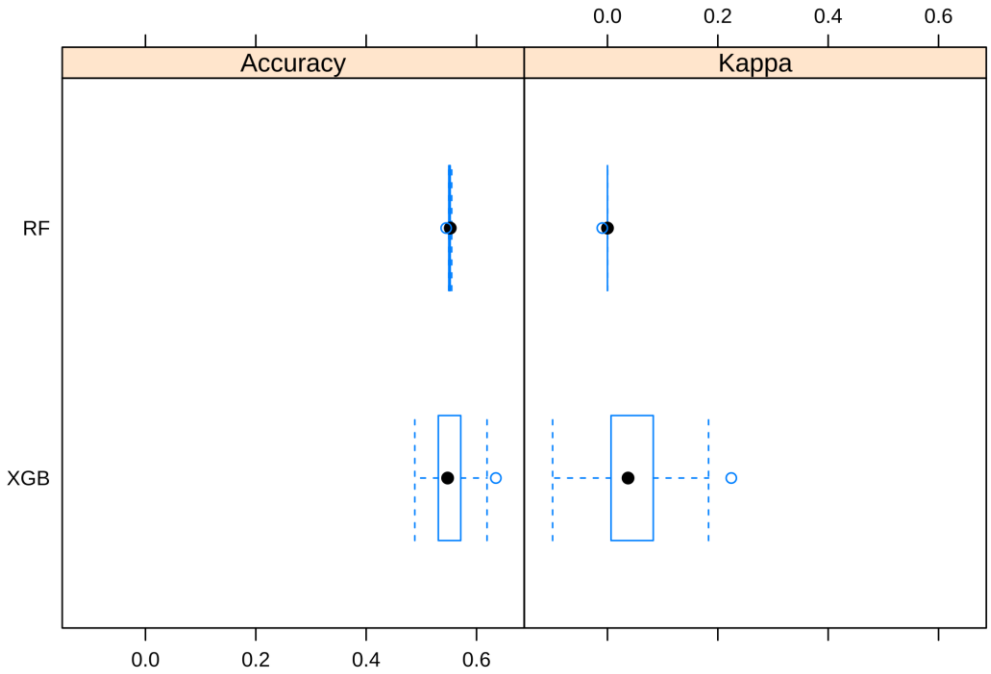


Figure 5: Accuracy and Kappa score of RF and XGB model in predicting biomarkers for Wt, alpha and other VoC infected patients in complicated cohort. A low alpha and kappa value suggests that the model's performance in terms of accuracy and agreement is not optimal.

Differential expression in Critical patients infected with wt, alpha and other variants

Similar to the above, we considered exclusively critical patients infected with the Wt, alpha, or other variants of concern (VoC). We are exploring the differential expression of biomarkers within the cohort of critical patients, specifically examining how these expressions vary when infected by different virus variants. Using RF and XGB model we represent relative significance of biomarkers in Figure 6. We found CRP level 7, PCT level 2 and Ddimer level 5 & 6 can effectively serve as classifiers for distinguishing these two groups. In the below panel we showed Alpha and Kappa score of these two models, which is commonly used to evaluate the performance of ML models. The alpha score is associated with overall accuracy, while the kappa score assesses the agreement between predicted and observed classifications, considering the possibility of chance agreement. A high alpha and kappa value suggests that the model is performing well in terms of accuracy and agreement between the predicted and observed data.

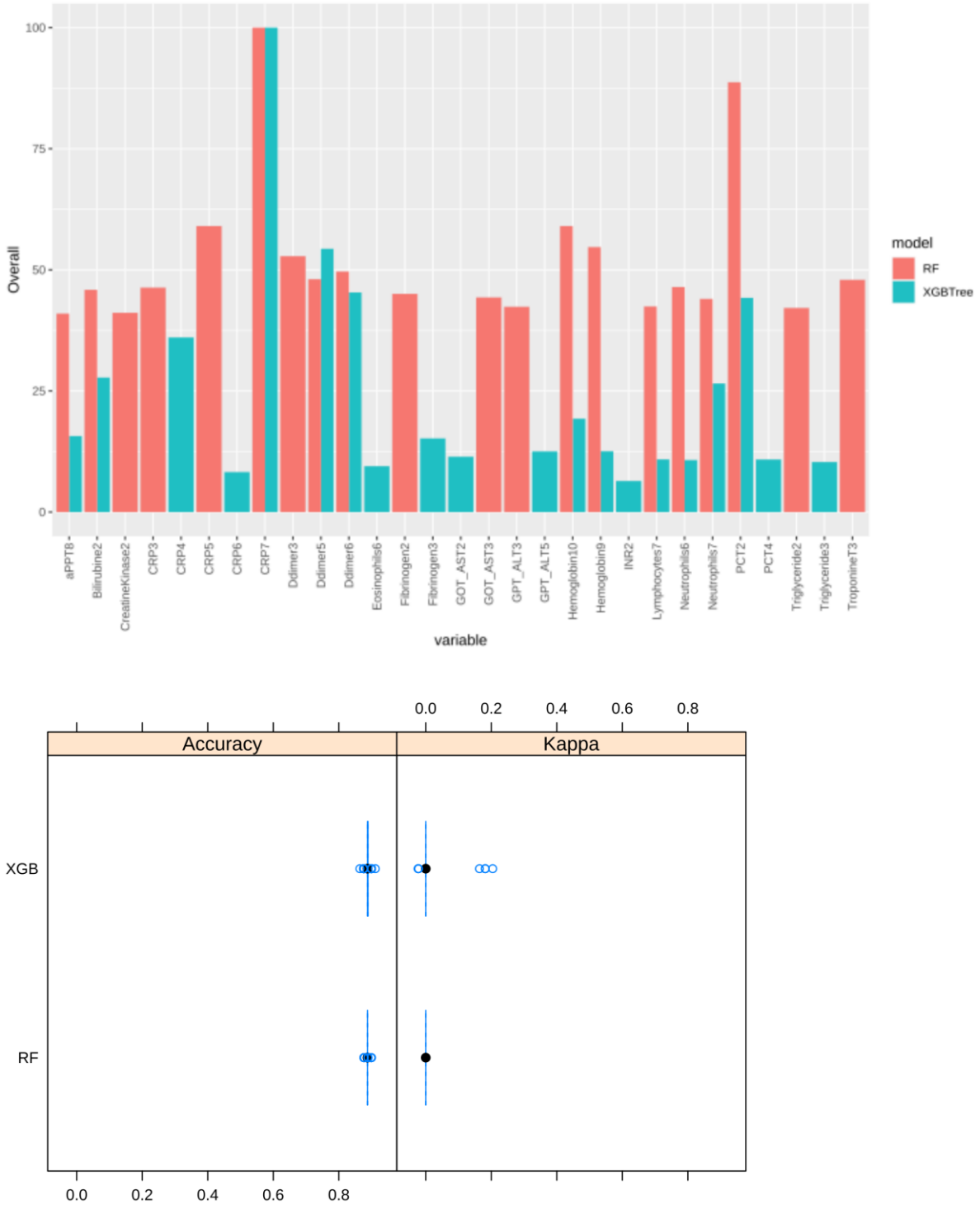


Figure 6:The RF and XGB model prediction of the relative significance of biomarkers, serving as classifiers for distinguishing between Wt, alpha and other VoC infection among critical patients. Prediction outcomes from the RF and XGBoost (XGB) methods are denoted by red and cyan, respectively. Our model indicates that CRP level of 7, Ddimer level of 5&6, and PCT level of 2 can effectively serve as classifiers for these two groups. In the below panel we show Accuracy and Kappa score for these two models.

Discussion on potential bias due to the training data set, details on explainability and ethical considerations raised by the use of AI/ML technology

Our machine learning models are trained on a representative and fair dataset that consists of more than 12000 patients information collected across 122 German hospitals. In this data set all genders, all age groups and all patient conditions (Uncomplicated, Complicated, Critical and Recovered) are well represented. More information is available at https://leoss.net/. Here, in Figure 7 we present the distribution of age and gender within the dataset.

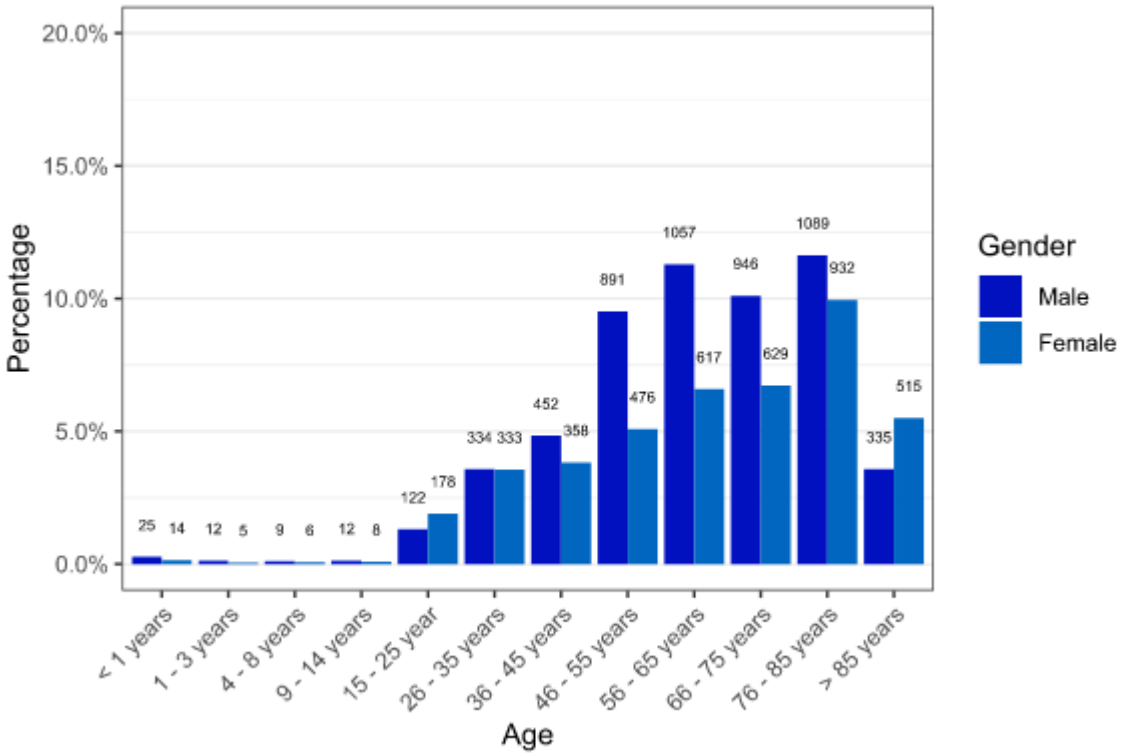


Figure 7: Histograms illustrating the distribution of age and gender in the LEOSS dataset.

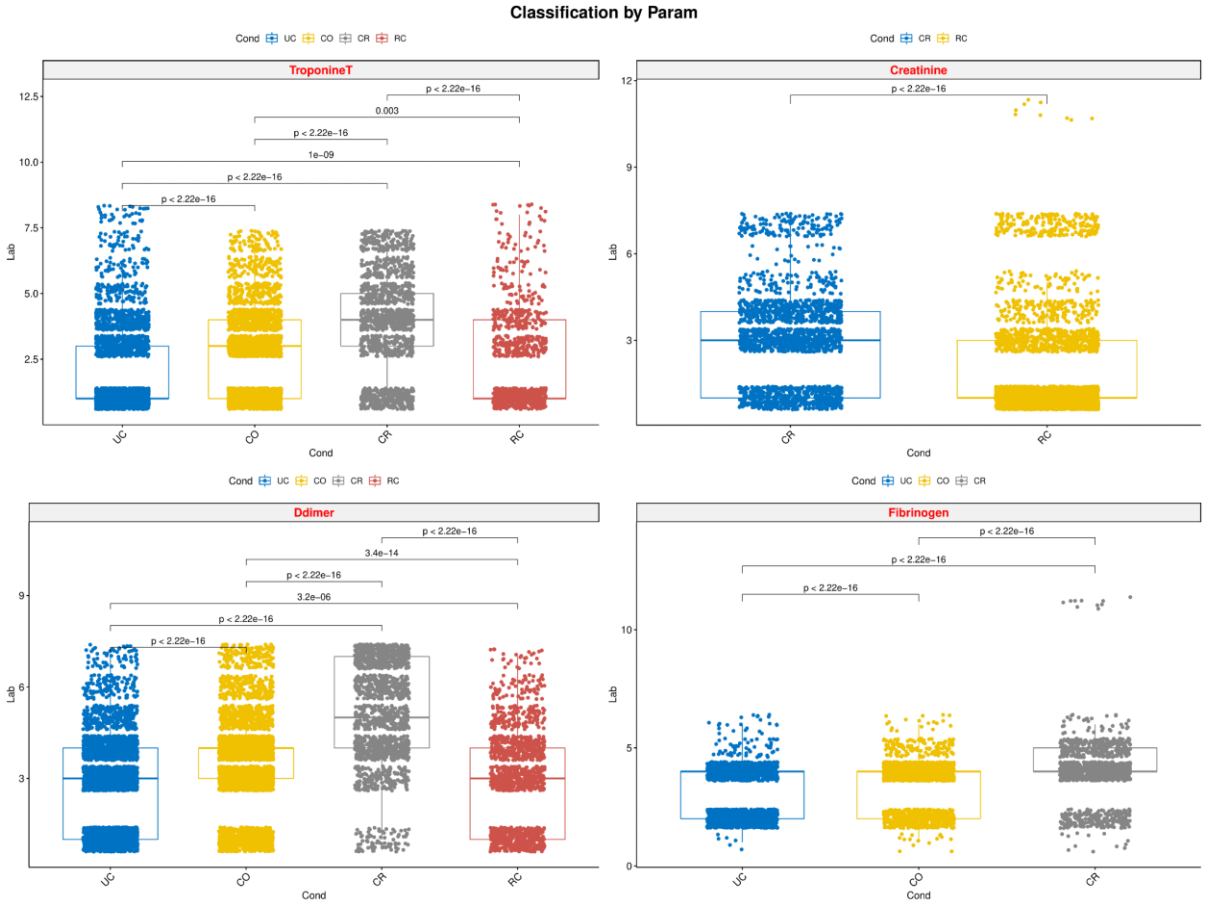
Importantly the machine learning methods (Random Forest and XGBoost) that we used to analyse the data, involve regularization techniques that penalize the model for making biased predictions. This encourages the model to learn patterns that are fair across various groups, stop overfitting, and contribute to a more balanced and unbiased model. However, the data set used stems from German hospitals, thus, introducing a bias towards European, primarily Caucasian subpopulations and potentially does not allow for a generalisation to the World population.

We have not systematically investigated explainability of the methods used here yet. However, random forests allow for systematic evaluation of decision trees in general and, thus, allow insight into the algorithm's decision process.

Differential expression of biomarkers in patients across disease state

In order to understand the clinical significance of a biomarker and its correlation with the progression of the disease, our initial step involved analysing the expression of the biomarker across different disease states. As all data are presented as categorical variables, we performed a non-parametric Wilcoxon test for all possible groups and a significance value is attached (Figure 8). For all the statistical analyses, $p < 0.05$ was considered to be significant. Analyses were performed in R version 4.2.3.

Our findings revealed that Troponin, D-dimer, Fibrinogen, Lymphocytes, Haemoglobin, and CRP (C-reactive protein) exhibited significantly higher expression levels in patients with complicated (CO) and critical (CR) disease states compared to patients in the uncomplicated (UC) disease state.




Cond UC CO CR
Leukocytes
p < 2.22e−16
p < 2.22e−16
p < 2.22e−16
Lab
Cond
UC
CO
CR
Cond UC CO CR RC
Lymphocytes
p < 2.22e−16
p < 2.22e−16
6.6e−15
3e−12
p < 2.22e−16
p < 2.22e−16
Lab
Cond
UC
CO
CR
RC
Cond UC CO CR RC
Neutrophils
p < 2.22e−16
p < 2.22e−16
2.4e−15
0.99
p < 2.22e−16
p < 2.22e−16
Lab
Cond
UC
CO
CR
RC
Cond UC RC
Platelets
0.0006
Lab
Cond
UC
RC
Cond UC CO CR RC
Hemoglobin
p < 2.22e−16
0.0014
p < 2.22e−16
1.4e−07
p < 2.22e−16
p < 2.22e−16
Lab
Cond
UC
CO
CR
RC
Cond UC CO CR RC
Eosinophils
p < 2.22e−16
p < 2.22e−16
0.45
p < 2.22e−16
p < 2.22e−16
p < 2.22e−16
Lab
Cond
UC
CO
CR
RC
Cond UC CO CR RC
CRP
p < 2.22e−16
8.8e−08
p < 2.22e−16
p < 2.22e−16
3.9e−08
5.7e−10
Lab
Cond
UC
CO
CR
RC
Cond CO RC
IL6
p < 2.22e−16
Lab
Cond
CO
RC

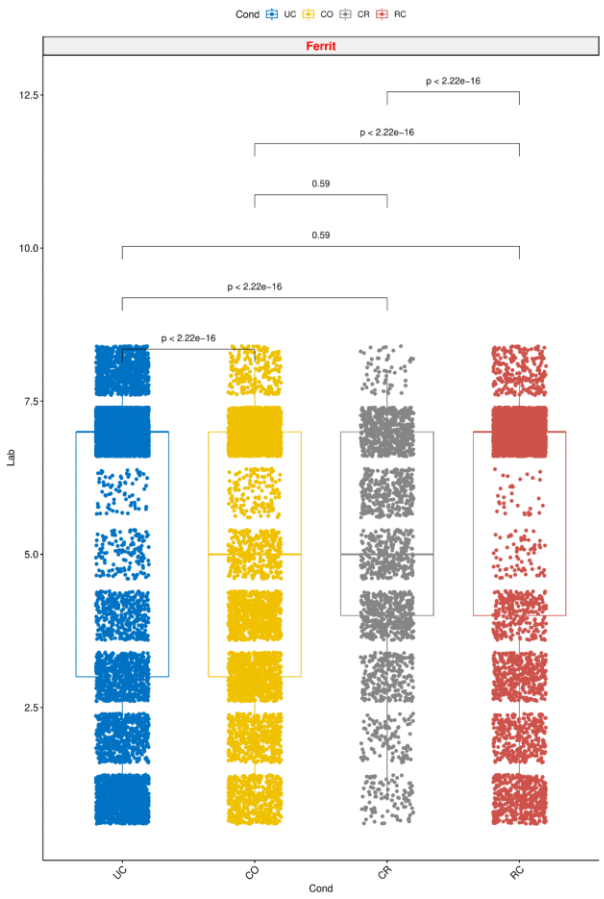
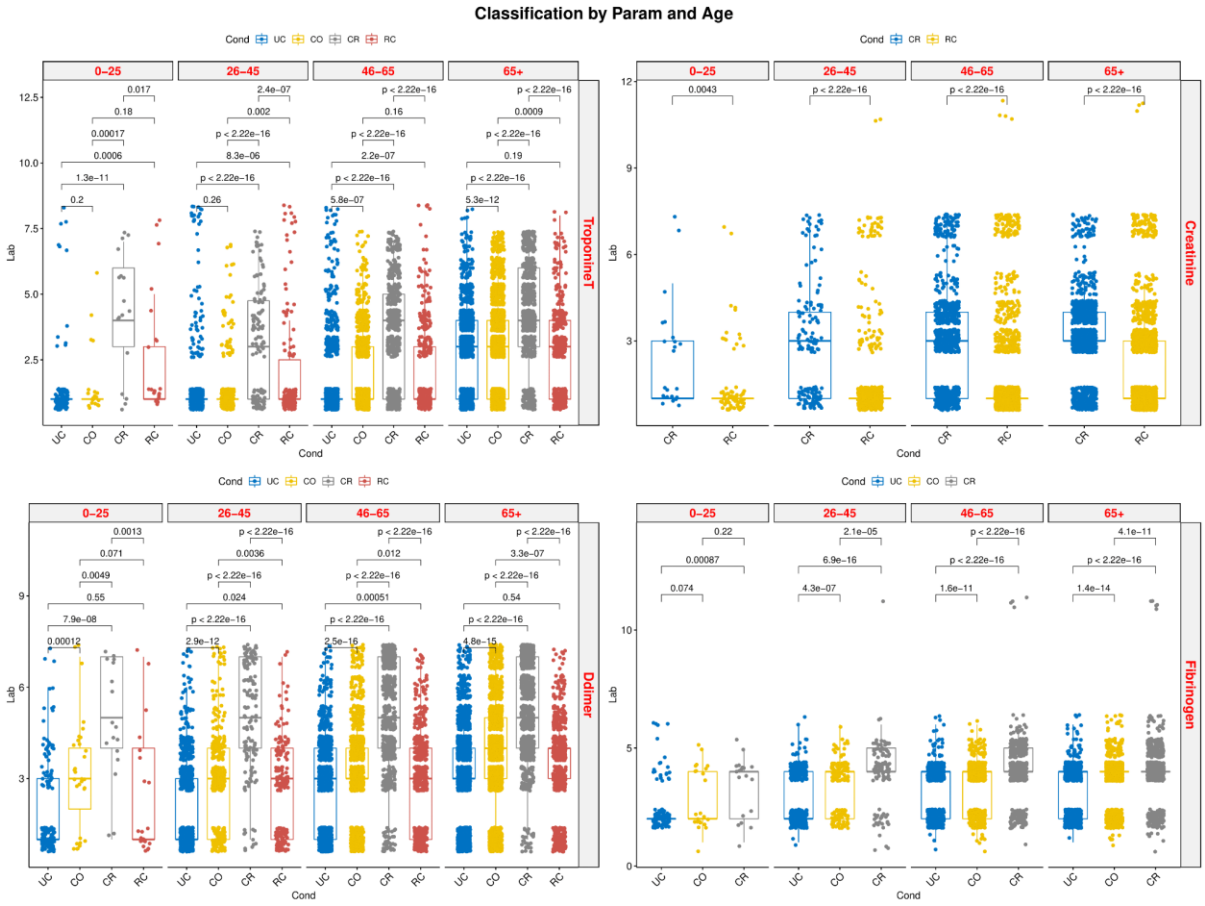
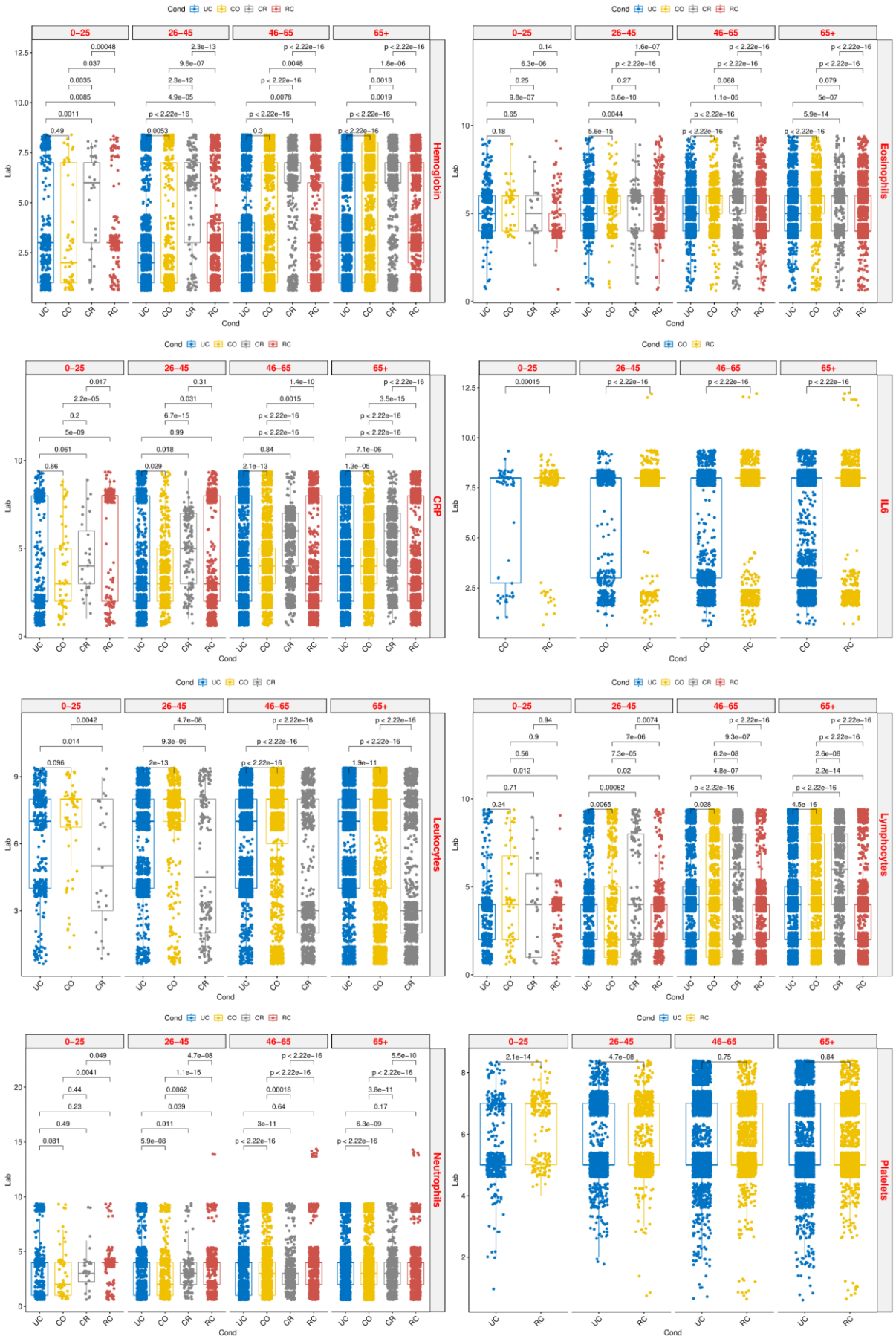


Figure 8. Expression of biomarkers across disease states, categorized as Uncomplicated (UC - blue), Complicated (CO - yellow), Critical (CR - grey), and Recovered (RC - red). The biomarker names are indicated at the top, and the Y-axis reflects the biomarker levels across the disease states. Given that biomarker levels are categorical variables, we conducted a non-parametric Wilcoxon test for all possible group combinations, with corresponding significance values (p-values) provided on top of the group comparisons. The rectangular "box" represents the interquartile range (IQR) of the data, with a line inside indicating the median.

Given that the elderly population is known to be particularly vulnerable to COVID-19, we specifically examined whether the expression of biomarkers was consistent across different age groups or if there was a notable difference in biomarker expression within the elderly population. We split the patients into four age groups: 0-25, 26-45, 46-65, 65+. In Figure 9 we showed biomarker expression in different age groups. We found that Tropanine, D-dimer, Fibrinogen, Lymphocytes, Haemoglobin, and CRP (C-reactive protein) exhibited higher expression levels across different age groups.


Classification by Param and Age
Cond UC CO CR RC
0–25
26–45
46–65
65+
TroponinT
Creatinine
Ddimer
Fibrinogen
Lab
Cond


Cond UC CO CR RC
0–25
26–45
46–65
65+
Hemoglobin
Eosinophils
CRP
IL6
Leukocytes
Lymphocytes
Neutrophils
Platelets
Lab
Cond
UC
CO
CR
RC

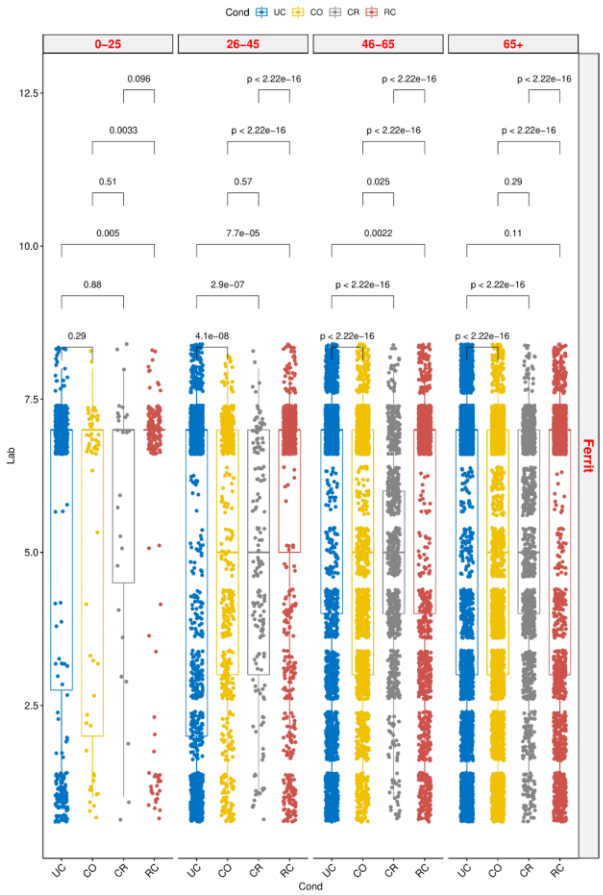
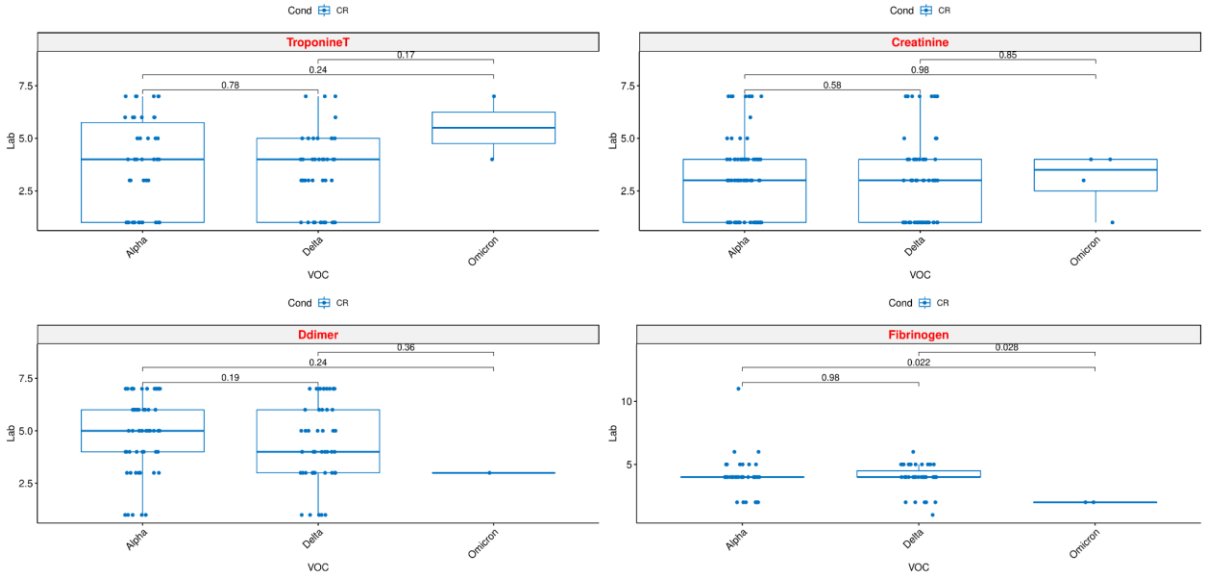


Figure 9. Expression of biomarkers in different age-groups and across disease states, categorized as Uncomplicated (UC - blue), Complicated (CO - yellow), Critical (CR - grey), and Recovered (RC – red). Age-groups considered are represented at the top of the panel and the biomarker of interest is indicated at the right. The Y-axis reflects the biomarker levels. Given that biomarker levels are categorical variables, we conducted a non-parametric Wilcoxon test for all possible group combinations within each age group, with corresponding significance values (p-values) provided on top of the group comparisons. The rectangular "box" represents the interquartile range (IQR) of the data, with a line inside indicating the median.

Differential expression of biomarkers across virus variant type

We next want to explore the expression of biomarkers across different virus variant types (alpha, delta, omicron) in critical patients. In Figure 10 we plot different biomarker expression across different VoC infection in CR patients. We do not find significant differences in expression of Tropanine, Creatinine, D-dimer, Lymphocytes, and CRP levels across variant types. However, Fibrinogen expression is lower in omicron compared to alpha and delta. Haemoglobin expression is lower in delta compared to alpha.

Classification by Param



Classification by Param

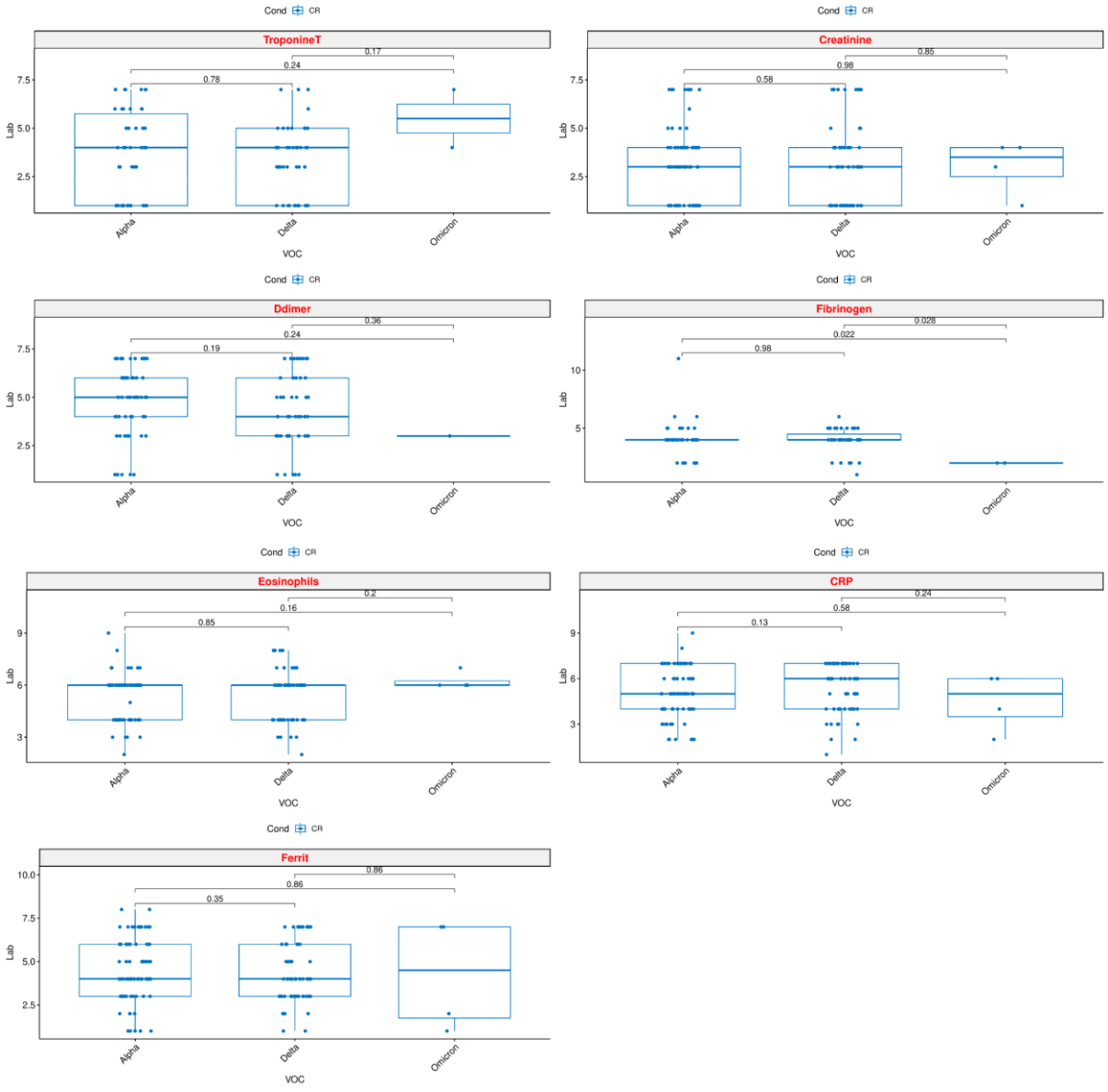
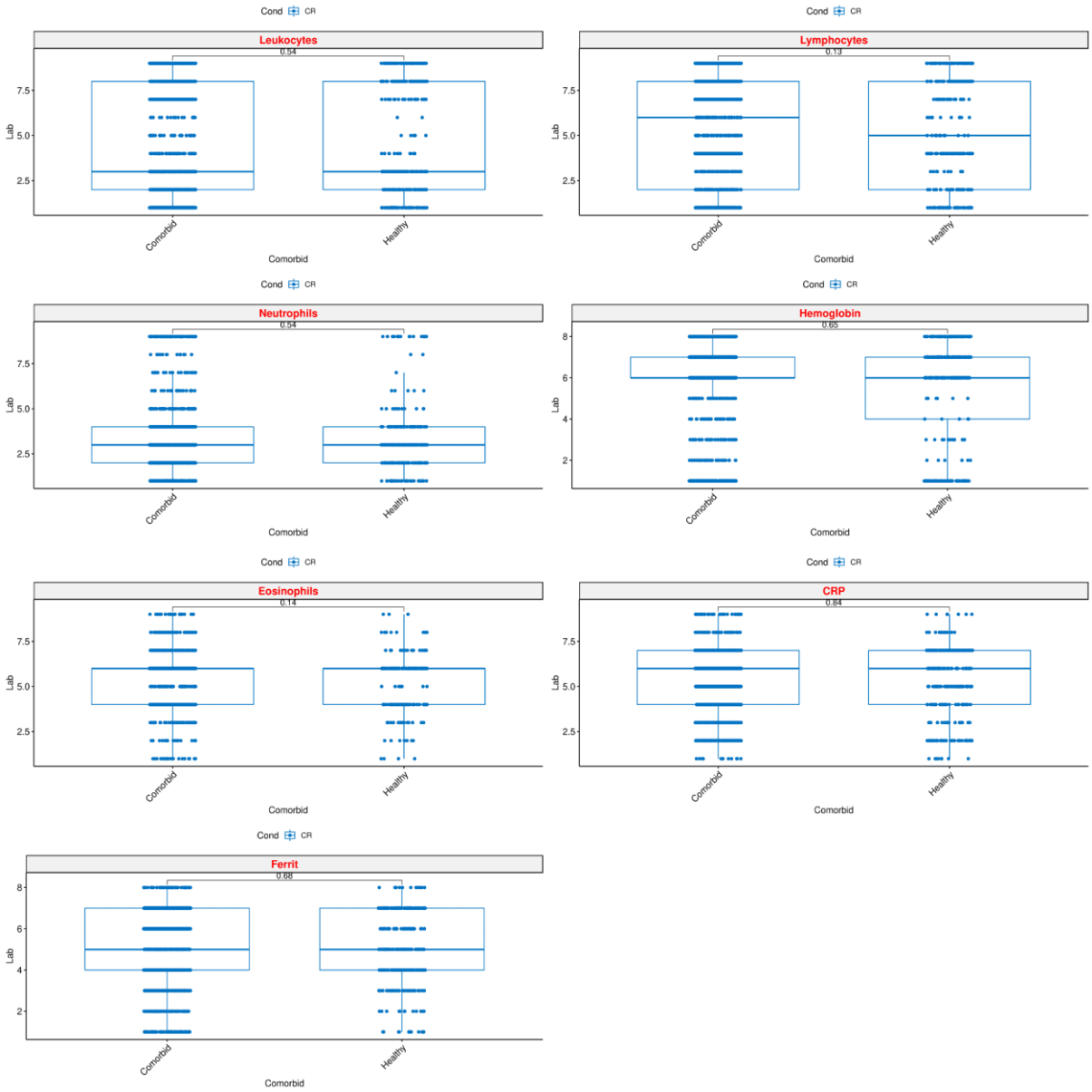


Figure 10. Expression of biomarkers across different VoC infection among the critical patients. Y-axis reflects the biomarker levels. A non-parametric Wilcoxon test for all possible group combinations was performed, with corresponding significance values (p-values) provided on top of the group comparisons. The rectangular "box" represents the interquartile range (IQR) of the data, with a line inside indicating the median.

Differential expression of biomarkers in comorbid patients

We next want to explore the expression of biomarkers in comorbid patients. For this we split the data into two groups, healthy and comorbid. Patients are considered comorbid if there is an entry in Cardiovascular disease, Pulmonary disease, Hematological/oncological disease. Diabetes mellitus, Kidney disease,Neurological diseases Other comorbidities. In Figure 11 we show biomarker expressions in the healty and comorbid patients and in Figure 12 we show biomarker expression of healthy and comorbid patients across diffent age groups. Only Troponin and Creatinine are found differentially expressed into these groups.



Classification by Param

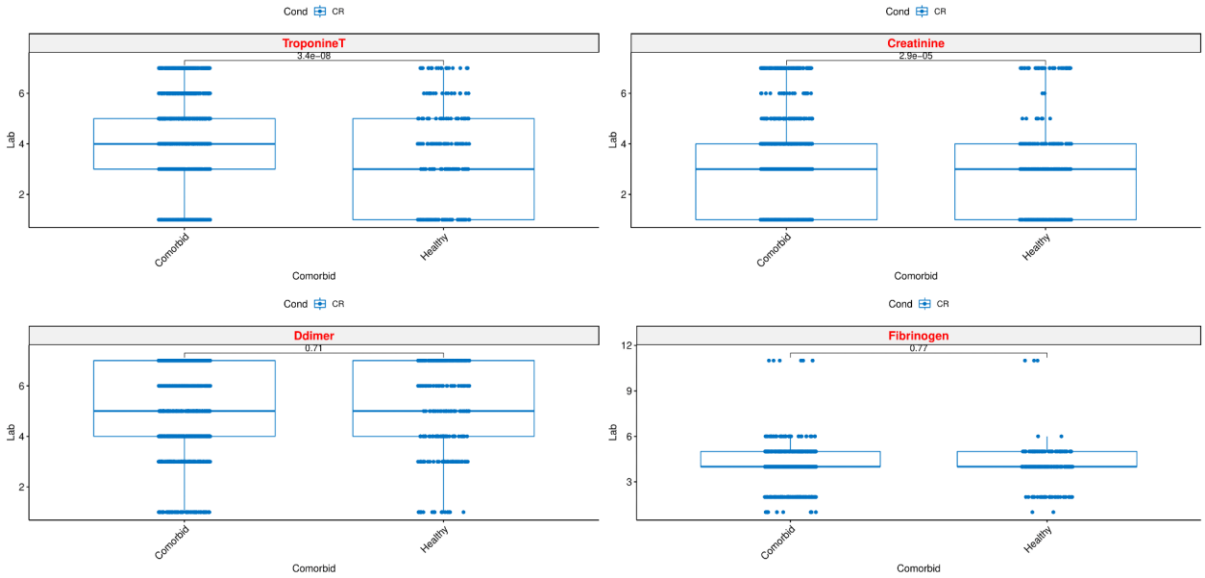


Figure 11. Expression of biomarkers in Healthy and Comorbid patients. Y-axis reflects the biomarker levels. A non-parametric Wilcoxon test was performed and corresponding significance values (p-values) are attached. The rectangular "box" represents the interquartile range (IQR) of the data, with a line inside indicating the median.

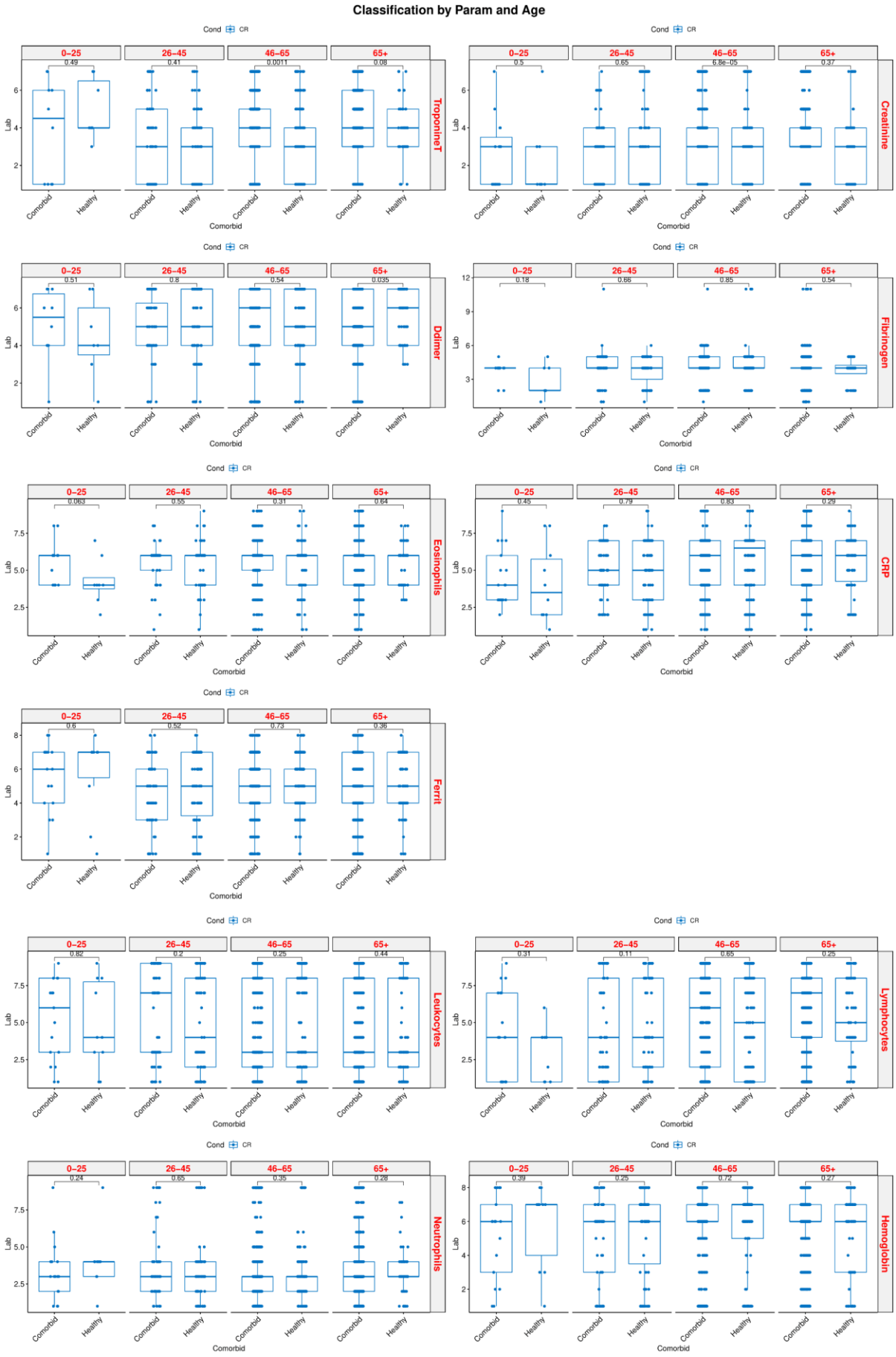

Classification by Param and Age
Cond CR
0–25
26–45
46–65
65+
0.49
0.41
0.0011
0.08
TroponineT
0.5
0.65
6.8e−05
0.37
Creatinine
0.51
0.8
0.54
0.035
Ddimer
0.18
0.66
0.85
0.54
Fibrinogen
0.063
0.55
0.31
0.64
Eosinophils
0.45
0.79
0.83
0.29
CRP
0.6
0.52
0.73
0.36
Ferrit
0.82
0.2
0.25
0.44
Leukocytes
0.31
0.11
0.65
0.25
Lymphocytes
0.24
0.65
0.35
0.28
Neutrophils
0.39
0.25
0.72
0.27
Hemoglobin
Lab
Comorbid
Healthy

Figure 12. Expression of biomarkers in Healthy and Comorbid patients across different age groups. Y-axis reflects the biomarker levels. A non-parametric Wilcoxon test were performed for group comparison and corresponding significance values (p-values) provided. The rectangular "box" represents the interquartile range (IQR) of the data, with a line inside indicating the median.

LEOSS group collects data via an electronic case report form (eCRF) provided by using the ClinicalSurveys.net online platform of the University Hospital of Cologne (UHC). ClinicalSurveys.net is hosted by QuestBack, Oslo, Norway on servers in Cologne, Germany as part of a software-as-a-service agreement. All data transmissions are encrypted via TLS 1.2 with an AES 256 GCM bit key and ECDHE RSA key exchange; certificate provided by COMODO RSA Domain Validation Server. Data is documented anonymously; no directly identifying data are stored on QuestBack servers. Regular data-backup, hierarchic management of rights and authentication protocols ensure the protection of data from unauthorized access and loss. Neither we have access nor we store any patient-related information.

Proper critical discussion of the findings and their expected impact

We found that Neutrophil, CRP, Ddimer, aPPT and PCT are good predictors for different disease states. Recognizing these biomarkers as predictive factors holds the promise of identifying patients at an elevated risk of progressing to severe forms of COVID-19 at an early stage. This, consequently, enables timely and suitable interventions, such as ongoing health condition monitoring, prompt medical attention, or specific therapeutic interventions etc, potentially leading to enhanced patient outcomes. The clinical relevance of these variables for the assessment of disease severity is outlined below.

Association with Neutrophil count:

Exploring the connection between COVID-19 severity and neutrophil levels is a key area of interest for unraveling the immune response and clinical implications associated with the disease. Numerous studies have noted a correlation between heightened neutrophil counts and the severity of COVID-19 [Ref 1, 5, 7]. In severe cases, an exacerbated immune response, commonly known as a cytokine storm, may occur, resulting in high expression of Neutrophils. Neutrophils play a crucial role in this response, releasing inflammatory mediators that contribute to damage in the lungs that can result in respiratory complications. The balance between an effective immune response and the potential hyperactivation of neutrophils is a critical aspect of COVID-19 pathophysiology.

Regular monitoring of neutrophil levels is integrated into standard clinical assessments for COVID-19 patients and our analysis underlines the importance of neutrophil counts for the clinical categorisation of disease severity.

Association with CRP level:

The correlation between COVID-19 severity and C-reactive protein (CRP) levels is also a subject of interest in comprehending the clinical manifestations and immune responses associated with the disease. Several studies have suggested a link between higher CRP levels and the severity of COVID-19 cases [Ref 2, 5, 7]. In response to pro-inflammatory cytokines (particularly interleukin-6 (IL-6)), the liver synthesizes and releases acute-phase proteins, including C-reactive protein (CRP). CRP serves as an essential component of the innate immune system. CRP binds to damaged cells and microorganisms, marking them for removal by other components of the immune system. In the context of severe COVID-19, where inflammation can lead to tissue damage, elevated CRP levels may also indicate the extent of tissue injury and the need for an intensified immune response.

Monitoring CRP levels is a routine part of clinical assessments for COVID-19 patients and following changes in CRP levels over time can provide insights into the progression of the disease and the effectiveness of treatments.

Association with DDimer level:

Upon infection, the virus can cause endothelial dysfunction affecting the inner lining of blood vessels. This contributes to a pro-inflammatory and pro-coagulant state. In response to this endothelial dysfunction, the coagulation cascade, a series of events that lead to blood clot formation, is triggered. This process involves the conversion of fibrinogen to fibrin, forming a mesh-like structure that helps in clot formation. As clots are formed, the body simultaneously initiates fibrinolysis, a process that breaks down fibrin clots. D-dimer is a fragment produced during fibrinolysis, and its levels rise when there is increased clot formation and breakdown. In severe cases of COVID-19, there is evidence of excessive blood clotting, both in the microvasculature and larger vessels. This hypercoagulable state can lead to the formation of blood clots in various organs, contributing to organ damage and complications.

Monitoring D-dimer levels is an important aspect of clinical care for COVID-19 patients. Persistently high D-dimer levels may suggest an increased risk of complications, including thrombosis and organ damage [Ref 3, 5, 7]. Clinicians use D-dimer levels, along with other clinical markers, to guide treatment decisions and assess the severity of the disease. Our models underscore the importance of this criterium.

Association with aPTT:

The activated partial thromboplastin time (aPTT) is a diagnostic test used to measure the time it takes for blood to coagulate. COVID-19 is associated with disruptions in normal blood clotting, and severe cases often manifest a heightened tendency for clot formation in both large and small blood vessels. In the context of COVID-19, an extended aPTT may signal irregularities in the intrinsic pathway of coagulation, which is closely linked to the increased clotting tendencies observed in severe cases. Severe instances of COVID-19 may give rise to disseminated intravascular coagulation (DIC), a condition marked by widespread activation of coagulation processes, resulting in the formation of microclots throughout the body. Prolonged aPTT serves as one of the indicators suggesting the occurrence of DIC.

Monitoring aPTT levels is part of the clinical assessment of COVID-19 patients, especially those at risk of developing coagulation abnormalities [Ref 4]. Abnormal aPTT levels, along with other coagulation markers, can guide clinicians in managing anticoagulation therapies and assessing the risk of thrombotic events.

Association with PCT:

PCT is a precursor peptide of calcitonin, a hormone involved in calcium homeostasis. Under normal conditions, PCT levels in the blood are very low. PCT levels are known to rise significantly in response to bacterial infections. Viruses, including SARS-CoV-2 (the virus that causes COVID-19), may not stimulate PCT production to the same extent as bacterial infections. However, elevated PCT levels are often considered a sign of systemic inflammation. In COVID-19, as the severity of the disease is often associated with an exaggerated immune response and inflammation, a heightened PCT levels could also be observed. This is consistent with other studies and clinical findings [Ref 5,6,7].

Most of the markers described above are already in use in current practice. Hence, our analysis is in agreement with the approach to diagnosis and prognosis currently employed by clinicians. No novel or unexpected predictive marker of severe disease could be identified in our analysis. However, the feature ranking provides important insights into prioritization of different diagnostic criteria. Further, it emphasizes the importance of further research into disease events that impact on these quantities.

References:

1. Cavalcante-Silva, Luiz Henrique Agra, et al. "Neutrophils and COVID-19: The road so far." *International immunopharmacology* 90 (2021): 107233.
2. Wang, L. "C-reactive protein levels in the early stage of COVID-19." *Medecine et maladies infectieuses* 50.4 (2020): 332-334.
3. Rostami, Mehrdad, and Hassan Mansouritorghabeh. "D-dimer level in COVID-19 infection: a systematic review." *Expert review of hematology* 13.11 (2020): 1265-1275.
4. Jin, Xin, et al. "The values of coagulation function in COVID-19 patients." *PLoS One* 15.10 (2020): e0241329.
5. Malik, Preeti, et al. "Biomarkers and outcomes of COVID-19 hospitalisations: systematic review and meta-analysis." *BMJ evidence-based medicine* 26.3 (2021): 107-108.
6. Hu, Rui, et al. "Procalcitonin levels in COVID-19 patients." *International journal of antimicrobial agents* 56.2 (2020): 106051.
7. Ponti, Giovanni, et al. "Biomarkers associated with COVID-19 disease progression." *Critical reviews in clinical laboratory sciences* 57.6 (2020): 389-399.