Deploying Machine Learning to Validate Clinical Phenotypes and Factors Associated with Mortality risk in 2,022 Critically Ill Patients with COVID-19 in Spain

Alejandro Rodríguez (✉ ahr1161@yahoo.es)  
Hospital Universitari de Tarragona Joan XXIII  
https://orcid.org/0000-0001-8828-5984

Manuel Ruiz Botella  
Hospital Universitari de Tarragona Joan XXIII

Ignacio Matín-Loeches  
St James Hospital

María Jiménez Herrera  
Universitat Rovira i Virgili

Jordi Solé-Violan  
Complejo Hospitalario Doctor Negrín: Hospital Universitario de Gran Canaria Dr Negrín

Josep Gómez  
Hospital Universitari de Tarragona Joan XXIII

María Bodí  
Hospital Universitari de Tarragona Joan XXIII

Sandra Trefler  
Hospital Universitari de Tarragona Joan XXIII

Elisabeth Papiol  
Hospital Vall d’Hebron: Vall d’Hebron Hospital Universitari

Emili Díaz  
Corporació Sanitària Parc Taulí: Consorcio Corporacion Sanitaria Parc Tauli

Borja Suberviola  
Hospital Universitario Marques de Valdecilla

Montserrat Vallverdú  
Hospital Arnau de Vilanova

Eric Mayor-Vázquez  
Hospital de Tortosa Verge de la Cinta: Hospital Verge de la Cinta

Antonio Albaya Moreno  
Hospital General Universitario de Guadalajara

Alfonso Canabal Berlanga  
Hospital Universitario de la Princesa

Miguel Sánchez  
Hospital Clínico Universitario San Carlos

María del Valle Ortíz  
Centro Hospitalario Burgos: Hospital Universitario de Burgos

Juan Carlos Ballesteros  
Hospital Universitario de Salamanca: Hospital Universitario de Salamanca

Lorena Martín Iglesias  
Central University Hospital of Asturias: Hospital Universitario Central de Asturias

Judith Marín-Coral  
Hospital del Mar Institute for Medical Research: Institut Hospital del Mar d’Investigacions Mediques

Esther López Ramos  
Hospital Universitario Príncipe de Asturias: Hospital Universitario Príncipe de Asturias
Keywords: Severe SARS-CoV-2 infection, Phenotypes, Risk factors, Prognosis, Machine Learning

DOI: https://doi.org/10.21203/rs.3.rs-125422/v1

License: ©  This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background: The identification of factors associated with Intensive Care Unit (ICU) mortality and derived clinical phenotypes in COVID-19 patients could help for a more tailored approach to clinical decision-making that improves prognostic outcomes.

Methods: Prospective, multicenter, observational study of critically ill patients with confirmed COVID-19 disease and acute respiratory failure admitted from 63 Intensive Care Units (ICU) in Spain. The objective was to analyze patient's factors associated with mortality risk and use a Machine Learning (ML) to derive clinical COVID-19 phenotypes. Patient features including demographics and clinical data at ICU admission were analyzed. Generalized linear models were used to determine ICU mortality risk factors. An unsupervised clustering analysis was applied to determine presence of phenotypes. The prognostic models were validated and their performance was measured using accuracy test, sensitivity, specificity and ROC curves.

Results: The database included a total of 2,022 patients (mean age 64[IQR5-71] years, 1423(70.4%) male, median APACHE II score (13[IQR10-17]) and SOFA score (5[IQR3-7]) points. The ICU mortality rate was 32.6%. Of the 3 derived phenotypes, the C(severe) phenotype was the most common (857;42.5%) and was characterized by the interplay of older age (>65 years), high severity of illness and a higher likelihood of development shock. The A(mild) phenotype (537;26.7%) included older age (>65 years), fewer abnormal laboratory values and less development of complications and B (moderate) phenotype (623,30.8%) had similar characteristics of A phenotype but were more likely to present shock. Crude ICU mortality was 45.4%, 25.0% and 20.3% for the C, B and A phenotype respectively. The ICU mortality risk factors and model performance differed between whole population and phenotype classifications.

Conclusion: The presented ML model identified three clinical phenotypes that significantly correlated with host-response patterns and ICU mortality. Different risk factors across the whole population and clinical phenotypes were observed which may limit the application of a “one-size-fits-all” model in practice.

Funding: None

Introduction

Since the outbreak of COVID-19 disease began in December 2019 in China, soaring cases of confirmed SARS-CoV-2 are pummeling the global health system. More than 61 million people have developed SARS-CoV-2 infection, and more than 1 million have died1. The current diagnostic and monitoring procedure rely on an inaccurate and often delayed identification by the clinician who then decides the treatment upon interaction with the patient. Critical illness from COVID-19 has constrained intensive care unit (ICU) material and human resources2. As of November 30, 2020 more then 1.6 million people in Spain have been infected with SARS-CoV-2 and more than 40,000 have died3. Short-term mortality reported rate ranges from 16% to 62% of patients admitted to ICU4-6. The heterogeneity of patients that have been treated in China4, Italy8, USA5,7 or Spain9,10 may explain the wide variation of mortality rate due to their population characteristics, presence of comorbidities and healthcare systems. A recent international survey11 reported significant practice variations in the management of severe COVID-19 patients, including differences at the regional, hospital, and patient level. Therefore, it is necessary to characterize phenotypes, by extending the enrolment of patients outside of one ICU site to multiple patients being treated in different hospitals. Allowing to adequately measure mortality-related factors adjusted by the inter-hospital variation to determine clinical outcomes.

Risk factors represent the most important approach when defining treatment of hospitalized patients as these measures can inform clinical courses most likely for a patient given their a priori risk. However, risk factors can also interplay differently when they are included in different patient clusters. A single model based on general risk factors (one-size-fits-all) might be limited for clinical data interpretation and application across sites. Different combinations of risk factors may naturally cluster into previously undescribed subsets or phenotypes that may have different risks for a high mortality rate and that may therefore help to determine the response to treatments in COVID-19. We hypothesize that the presence of well-defined phenotypes in COVID-19 could help to more appropriately identify patients at risk of ICU mortality than general models for the entire population considering that this disease results in a constellation of symptoms, laboratory derangement, immune dysregulation, and clinical complications.

The primary objective was to assess which factors are independently associated with ICU mortality. The second objective was to determine the presence of distinct clinical phenotypes using unsupervised clustering methods that were applied to the datasets available on ICU admission. The added value of this large-scale multicenter prospective study lies to discover 3 phenotypes based on clinical data available at ICU admission that can help explain the variation in clinical results of COVID-19 disease in the ICU.
Material And Methods

Study design: A multicenter observational, prospective cohort study that consisted of a large-scale data source of hospital ICU admissions and patient-level clinical data. The enrolment criteria included adult patients with laboratory confirmed SARS-CoV-2 infection admitted in 63 ICUs across Spain due to acute respiratory failure between February 22, 2020 and May 11, 2020. The study was approved by the reference institutional review board at Joan XXIII University Hospital (IRB# CEIM/066/2020) and each participating site with a waiver of informed consent. All data values were anonymized prior to the phenotyping which consisted of clustering clinical variables on their association with COVID-19 mortality.

Study Sites and Patients Population: The study enrolled consecutive adult patients (>16 years) with laboratory confirmed SARS-CoV-2 infection, detected by RT-PCR positive test of nasopharyngeal, oropharyngeal swab or invasive respiratory samples according to the WHO recommendations. The follow-up of patients was scheduled until August 11, 2020, which confirmed ICU discharge or death whichever occurred first. A complete list of participating ICUs and their investigators is provided in the acknowledgements section. In this cohort, 43 patients were described in a preliminary report of a single-center case series in Tarragona, Spain.

Outcomes: The primary outcome included all-causes of ICU mortality. Patients who were discharged alive from ICU were evaluated in the data as alive considering mortality was defined as any in-ICU death. All complications and outcomes were followed during ICU admission.

Data collection: Data was obtained from a voluntary registry created by Spanish Society of Intensive Care Medicine-SEMICYUC. All consecutive cases admitted to the ICU were collected. There were no patients excluded from the analysis that was enrolled to participating ICU and met criteria.

All the collected variables recorded at ICU admission are listed in the additional file (P 8-9). To determine severity of illness, the Acute Physiology and Chronic Health Evaluation (APACHE) II score and Sequential Organ Failure Assessment (SOFA) scoring were calculated for all patients within the first 24 h of ICU admission.

The ICU admission criteria considered the use of antiviral, antibiotic or co-adjuvant treatment for all patients, and also the measures that would determine the need to intubate and type of ventilator support required. These measures were not standardized between centers and were left to the discretion of the attending physician, according to SEMICYUC and National Ministry of Health and were included in the case report form and confirmed by the medical records. We also collected hospital-level data including city, county and number of hospital beds available. The study definitions used in the present analysis are shown in the additional file (P 5).

Statistical analysis

No statistical sample size calculation was performed a priori, and sample size was equal to the number of patients admitted to the participant's ICUs with confirmed COVID-19 during the study period. To describe baseline characteristics, the continuous variables were expressed as median (interquartile range [IQR]) and categorical variables as number of cases (percentage). For patient demographics and clinical characteristics, differences between groups were assessed using the chi-squared test and Fisher's exact test for categorical variables, and the Mann–Whitney U or Wilcoxon test for continuous variables. To performed the analysis, we first assessed the candidate variables, missing values, and correlation. Multiple imputation was used to account for missing data (additional file, P 5). After evaluating correlation, highly correlated variables were excluded (additional file, P 8).

An overview of the primary analysis plan is outlined in Figure 1. In a first step, a multilevel conditional logistic modelling and the intraclass correlation coefficient (ICC) was calculated (additional file, P 6) with patients nested in hospital to characterize hospital-level variation of ICU mortality and determine if a significant inter-hospital variation is present.

In a second step, a multivariate analysis (GLM: Generalized linear Regression model) was performed to investigate the association between baseline (on ICU admission) variables and ICU-mortality. The GLM model comprised factors of clinical interest and all significant covariates (p<0.05) in the univariate analysis of ICU mortality and presence of collinearity was studied by variance inflation factors (VIF). The results are presented as odds ratios (OR) and 95% confidence intervals (CI).

To determine our model, we checked adequate model performance between groups with a cross validation model (K-fold=10) and the model with better performance was chosen. For the model validation, database was randomly split into two subsets: (a) a “training set”
Lastly, to determine presence of distinct clinical phenotypes in our population of COVID-19 patients, an unsupervised clustering analysis was applied to the database at ICU admission. In order to carry out this analysis, a discretization of the numerical variables into categorical ones was done. The information provided by each variable regarding ICU mortality was defined using the Information Value (IV). A IV greater than 0.03 was considered clinically important and this variable was included in the multivariate logistic regression analysis. Model performance was examined using accuracy test, Sensitivity, Specificity and AUC modeling. Subsequently, the unsupervised cluster analysis was performed using the important variables. The Podani distance was used to calculate the distance between patients and the “partition around medoids” (PAM) algorithm to perform the clustering. The optimal number of clusters were determined after studying the silhouette and the PAM objective for different numbers of clusters (additional file, P 15). Each of these clusters represents a specific patient’s phenotype. To visualize the clusters in a lower dimensional space, we used a Principal Component Analysis (PCA). We obtain important variables (IV) for each phenotype, and the OR of these variables were obtained after applying a GLM analysis. Multinomial regression models were fit to further compare patient comorbidities across phenotype classification. Model performance in each phenotype was examined using accuracy test, Sensitivity, Specificity and AUC. Data analysis was performed using R software (cran.r-project.org).

Results

Patients Characteristics at ICU admission

From February 29, 2020 to June 11, 2020 a total of 2,022 critically ill patients from 63 ICUs were enrolled in the present analysis. Forty percent of ICUs belonged to hospitals with more than 500 beds, 40% to hospitals between 200 and 500 beds and the remaining 20 percent to hospitals with fewer than 200 beds.

The median (IQR) age was 64 (55-71) years, and 1,423 (70.3%) were men. The median of time between the onset of symptoms and diagnosis was 7 (4-9) days. A total of 1,467 (72.6%) patients had at least one coexisting comorbidity. Arterial hypertension (936[46.7%]), obesity (655[32.4%]) and diabetes mellitus (420[20.7%]) were the most frequently co-morbid conditions reported. The severity of illness was high according to the APACHE II (14; IQR 11-18) and SOFA (5.7; IQR 4-7.3) scores. The PaO2/FiO2 ratio on the day of ICU admission was 132(IQR 96-163) and 1,131 (55·9%) patients meet criteria of severe and 198(9·8%) of mild acute distress respiratory syndrome (ARDS) profile at ICU admission, 1,174 (58.0%) patients required invasive mechanical ventilation (MV), 906 (44.8%) developed shock and 580 (28.7%) meet criteria of acute kidney injury (AKI). The overall ICU length of stay (LOS) was 14 (8-26) days and similar for survivors (14[8-27]) days and non-survivors (14[8-24]) days (p=0·10).

The most frequent prescribed co-adjuvant treatments for COVID-19-related infection were hydroxychloroquine (1866[92.3%]) and lopinavir/ritonavir (1662[82.2%]). Empiric antibiotic treatment was administered in 1,818 (89·9%) of the patients and intravenous corticosteroids in 1,174 (58·6%). Further clinical characteristics of patients and laboratory finding are shown in Table 1.

ICU mortality

Overall, 660 patients (32.6%) died. The crude ICU mortality increased significantly with the increase in predefined age cut-off and was greater than 80% in patients over 80 years old. (additional file, P 10). Age, male sex, severity of illness (APACHE II and SOFA), presence of arterial hypertension, diabetes, coronary arterial disease, chronic obstructive pulmonary disease (COPD), chronic kidney disease (CKD), immunosuppression and hematologic disease markers were significantly higher in ‘non-survivor’ patients. Non-survivor patients compared with those that survived had higher levels of D-Lactate dehydrogenase (LDH), white blood cells, serum creatinine, C-reactive protein (CRP), Procalcitonin (PCT), serum lactate, serum D-dimer and serum ferritin. Non-survivor patients developed more frequent complications such as shock, kidney and myocardial dysfunction at ICU admission. High Flow Nasal Cannula (HFNC) was more frequent in survivors, while invasive mechanical ventilation (MV) was more common in non-survivors. Mortality for those who received MV during their ICU stay (n= 1554; 76.8%) was 37.3% (n=580) higher than observed in patients who did not require MV 17.0% (80/468, p<0.001). Complete characteristics of patients according outcome are shown in Table 2.

Construction of the ICU Mortality multivariate model
Of the 42 variables measured at ICU admission, 25 variables that were statistically significant in the univariate analysis (Table 2) were included in the model. The initial dataset of patients was randomly split in two subgroups “Training group” with 1,618 patients (80%) and “Test group” with 404 patients (20%). The characteristics of patients included into each subgroup are shown in Table 1. No significant differences were observed between the subgroups. Inclusion of these 25 variables in a GLM model for the training group, resulted in 10 variables that were independently associated with ICU mortality (Figure 2). No presence of collinearity between explanatory variables was observed (additional file, P 11) and the Hosmer-Lemeshow Goodness-of-fit test (X-squared = 5.53, df = 8, p-value = 0.69) established no discrepancy between the observed values and those that would have been expected in the model. The validation of model in the test group demonstrated adequate performance with an accuracy of 0.78, a precision test of 0.73, sensitivity of 0.88, specificity of 0.45 and an AUC ROC of 0.82 (95%CI 0.78-0.86) (additional file, P 11).

Unsupervised analysis (cluster) to determine different phenotypes in critically ill patients

Once the variables were categorized, 5 patients (0.24%) were excluded for outlier’s data, and the analysis was performed with 2,017 patients. Of the 50 variables considered, only 25 were considered as predictors according to the IV (additional file, P 12) and were included in the model. Remarkably, no treatment option was a predictive factor for ICU-mortality. The categorized variables independently associated with ICU-mortality are shown in additional file (P 14). The performance of the model was adequate with an accuracy of 0.77, sensitivity of 0.88, specificity of 0.54 and AUC of 0.82. According to the Podani’s distance and the Silhouettes and PAM plots (additional file, P 15) the optimal number of clusters in our dataset was 3. Cluster A included 537 patients (26.7%), cluster B included 623 (30.8%) and cluster C included 857 patients (42.5%). The clusters in a lower dimensional space are shown in the additional file (P 16). The size and characteristics of the phenotypes in the 3-class model appear in the additional file (P 13). Patients with the cluster A phenotype (mild COVID-19 disease) had <65 years, lower severity of illness, fewer abnormal laboratory values and less development of complications, with a crude ICU mortality of 20.3%; those with the cluster B phenotype (moderate COVID-19 disease) had similar characteristics as seen in the A phenotype but were more likely to present shock at ICU admission with a crude ICU-mortality of 25.5%. Patients with the cluster C phenotype (severe COVID-19 disease) had >65 years, a high level of severity of illness, more likely to have elevated measures of inflammation (e.g. D dimer, LDH and ferritin), high frequency of shock, AKI and myocardial dysfunction, with a crude ICU mortality of 45.4%. The clinical characterization of each observed phenotype can be seen in Figure 3. By including these important variables (additional file, P 17) in a regression model for each cluster, we observed that the discrimination of each model was higher than general model except for C phenotype (Table 3). The Variables independently associated with mortality were different between global and cluster models (Figure 4 and additional file, P 18-19).

Discussion

The main finding of our study is that among patients with COVID-19, 3 clinical phenotypes were derived using habitual clinical and laboratory variables at ICU admission. The ability of identifying phenotypes using a small set of variables is a crucial step towards clinical application and has important implications for possible differential treatment guided by phenotypes and validated prognostic scoring systems18,19.

Our C phenotype was associated with more than double the ICU mortality than each of the remaining two phenotypes. This C phenotype was characterized by the interplay of older age (>65 year), a high severity (APACHE II > 15 and SOFA >5), greater burden of risk factors (hematologic disease and coronary disease) and a higher likelihood of developing further complications (shock and AKI).

Previous studies have implemented clustering techniques to analyze various data sources relating to demographic, geographic, environment, and socioeconomic determinants of health and disease. There are studies that have evaluated treatment decisions and characterized clinical phenotypes associated with complications, ICU admission and mortality risk in critically ill COVID-19 patients. According to the Situation Report & Public Health Guidance published by Johns Hopkins University on March 19th, 2020, people over 60 and those with chronic health conditions are at the highest risk for COVID-19 complications20. To our knowledge, this is the first study with a high number of critically ill patients to analyze the presence of phenotypes in patients with SARS-CoV-2 infection. This multicenter cohort study of 2,022 critically ill patients found that 660 patients (32.6%) died at ICU discharge. Our ICU mortality rates was significantly lower as reported in Yang X et al.4 in Wuhan, China (61.5%), by Myers L et al7 in California, USA (50.0%), by Arentz M et al.6 in Washington, USA (67%) and by Richardson S et al.5 in New York, USA (78%), but slightly higher to reported by Grasselli G et al8. In Lombardy region, Italy (26%). These observed differences in ICU mortality could respond to different healthcare models and important practice variations in the management of severe COVID-19 patients11, but it can also depend on the frequency of presentation of the different phenotypes.
In our study, a great variability in model performance and risk factors were observed during cross-validation to choose the best model to use. This might be related to the presence of a very heterogeneous population of patients, which is revealed during random partitioning performed (80%/20%) for validation of each model. In this context, three clinical phenotypes of COVID-19 patients were derived using routinely available clinical data at ICU admission by an unsupervised cluster analysis. The phenotypes were multidimensional, differed in their demographics, laboratory abnormalities, patterns of organ dysfunction, and associated with ICU mortality. In addition, the phenotypes were not homologous with traditional patient groupings such as the presence of clinical complications\textsuperscript{21,22}, or type of ARDS\textsuperscript{23}. Our COVID-19 phenotypes can be identified at the time of the ICU admission, and thus could be useful in facilitating early tailored therapy and improve prognosis.

Only routinely available clinical and laboratory data were used in the clustering models, and the phenotypes were derived from a large observational multicenter cohort to ensure generalizability. Importantly, we have observed that the variables associated with the ICU mortality varied between the global model and the models developed for each phenotype. The discrimination power (AUC) of A and B phenotypes models improved in comparison to the global model. However, for the C phenotype (severe COVID-19 disease), the performance of the model was not superior respect of the global model. The C phenotype was most strongly correlated with abnormal values of biomarkers as well as clinical features of cardiovascular dysfunction, AKI and subsequently a higher ICU mortality. Recently, several authors have proposed different clinical phenotypes of COVID-19 patients\textsuperscript{21-23}. Rello J et al.\textsuperscript{21} speculated that COVID-19 has five phenotypic presentations based on physiological and clinical features from published studies. Garcia-Vidal et al.\textsuperscript{22} describe the main clinical complications of hospitalized patients with COVID-19 through classification into three pattern groups (inflammatory, co-infection and thrombotic). However, as the authors acknowledge, the cut-off points of the different biomarkers for defining phenotypes have been arbitrary and not scientifically supported. Finally,Gattinoni L et al.\textsuperscript{23} proposing two phenotypes for COVID-19 patients, 1) “Type L” characterized by high compliance and low lung recruitability and 2) “Type H” with low compliance and high lung recruitability as a two “extremes” of a spectrum of respiratory failure in COVID-19 pneumonia. Despite the importance related to clinical experience in each of these approaches, none of these studies have been developed through a machine learning process to determine phenotypes nor have they been tested for validation.

Although hypoxemia is a proposed severity marker for phenotype differentiation\textsuperscript{21,23}, in our study the PaO2/FiO2 ratio to ICU admission was an independent risk factor for ICU mortality in the multivariate analysis but with a very small protector effect (OR=0.99) and only was closely associated with ICU mortality in C phenotype. Other variables such as advanced age, SOFA score, development of AKI or need for mechanical ventilation were more strongly related to ICU mortality than PaO2/FiO2 at ICU admission.

Our results should be interpreted in the context of the study limitations. First, although phenotypes were found to be generalizable in our population, risk factors and characteristics of clinical phenotypes were derived initially from data at ICU admission of multicenter observational study in Spain. However, these risk factors are similar to those that have been reported by other investigators\textsuperscript{4-8}. The cross-validation carried out and the high discrimination observed for each of the models built for phenotypes, suggests their applicability to other populations, but it should be examined considering the high variability observed in patients with COVID-19 and in the support measures applied. Second, because missing data were common for some variables included in the clustering models, multiple imputation was used in the primary analysis. However, variables with high missing values were excluded and the missing threshold used was reported elsewhere\textsuperscript{18}. Third, only routinely available clinical data at ICU admission were used to identify risk factors and clinical phenotypes, and the inclusion of other data related to clinical evolution of patients in the ICU could change risk factors or phenotype assignments. However, our objective was to study early risk factors and phenotypes at ICU admission that may allow for early treatment implementation and as a result improve patient outcome.

**Conclusion**

To our knowledge this is the largest study that describe different phenotypes of patients with confirmed COVID-19 that were admitted to ICU to date. We not only characterized three novel clinical phenotypes, but extended findings outside of a single site ICU by characterizing the association of comorbidities with clinical phenotype and the association of clinical phenotypes with clinical outcomes. Different risk factors for the global population and clinical phenotypes were observed, possibly due to the heterogeneity of patients, which may limit the application of a single predictive model for all patients with COVID-19. Further research is needed to determine the application of these phenotypes in clinical practice and for clinical trial design.

**Abbreviations**
ADRS
Acute distress respiratory syndrome

AKI
Acute kidney injury

APACHE II
Acute Physiology And Chronic Health Evaluation II

AUC
Area Under the ROC Curve

CKD
Chronic kidney disease

CI
Confidence interval

COPD
Chronic obstructive pulmonary disease

CRP
C-Reactive Protein

GLM
Generalized linear regression model

HFNC
High flow nasal cannula oxygen therapy

ICC
Interclass correlation coefficient

ICU
Intensive Care Unit

IQR
Interquartile range

IV
Information value

LDH
D-lactate dehydrogenase

LOS
Length of stay

MV
Mechanical ventilation

PAM
partition around medoids

PCA
Principal component analysis

PCT
Procalcitonin

ROC
Receiver Operator Characteristic curve

RT-PCR
real time polymerase chain reaction

SEMICYUC
Sociedad Española de Medicina Intensiva, Critica y Unidades Coronarias

SOFA
Sequential Organ Failure Assessment

VIF
Variance inflation factors

WHO
World Health Organization
Declarations

- Ethical Approval and Consent to participate
The study was approved by the reference institutional review board at Joan XXIII University Hospital (IRB# CEIM/066/2020) and each participating site with a waiver of informed consent

- Consent for publication
Not applicable

- Availability of supporting data
The anonymized database collected for the study by the SEMICYUC, and the data dictionary that defines each field in the set, will be made available to reviewers if they consider it necessary prior confidentiality agreement,

- Competing interests
All authors declare that they have no conflicts of interest

- Funding
This study was supported by the Spanish Intensive Care Society (SEMICYUC) and Ricardo Barri Casanovas Foundation. The study sponsors have no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

- Authors’ contributions
AR, MRB, IML, MJH, JSV, JG, MB, ST, EP, ED, BS, MV, EMV, AAM, JASG, MDVO, JCB, LMI, JMC, ELR, VHV, LV, SSC, FJGM, SHG, CCSP, JCPL, and RRG had substantial contributions to conception and design of the work.
AR, ST, MJH, MRB and JG had substantial contribution for data acquisition
AR, MRB, IML, JG, MB had substantial contribution for data analysis and interpretation of data for the work
AR, IML, MB, JSV and ED drafting of the manuscript
RF, IML, MB, ED, BS, JCB, SSC, FJGDM critically reviewed the draft manuscript
The corresponding author (AR) had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors approved the final version of the manuscript. The views expressed in this article are those of the authors and not necessarily those of the SEMICYUC.

- Acknowledgements
To Alexis Garduno for English manuscript edition and to the COVID-19 SEMICYUC Investigators:

Andalucía:

UCI Hospital Universitario Virgen de Valme (Sevilla): Ana Loza; UCI Hospital Quirón (Huelva): Diego Matallana Zapata; UCI Hospital Universitario Puerto Real (Cádiz): Isabel Díaz Torres, Sonia Ibañez Cuadros, María Recuerda Nuñez, Maria Luz Carmona Pérez, Jorge Gómez Ramos, Alba Villares Casas; UCI Hospital Universitario Virgen de la Macarena (Sevilla): María Luisa Cantón, José Javier González Contreras, Helena Pérez Chomón, Nerissa Alvarez Chicote, Alberto Sousa González; UCI Hospital Universitario Reina Sofía (Córdoba): María De Alba Aparicio, Juan Carlos Pozo Laderas; UCI Hospital Universitario de Jerez (Jerez de la Frontera): Angel Estella, Sara moreno Cano. UCI Hospital Infanta Elena (Huelva): Diego Matallana Zapata

Aragón:

UCI Hospital Nuestra Señora de Gracia (Zaragoza): Ruth Jorge García; UCI Hospital Clínico Universitario Lozano Blesa (Zaragoza): Laura Sánchez Montori, Sandra Herrero García, Paula Abanses Moreno, Carlos Mayordomo García. UCI Hospital General San

**Asturias:**

UCI Hospital Universitario San Agustín (Avilés): Ana Luz Balán Mariño, María José Gutiérrez Fernández, Marta Martín Cuadrado, Belén García Arias; UCI Hospital Universitario Central de Asturias (Oviedo): Lorena Forcelledo Espina, Lucía Vifña Soria, Lorena Martín Iglesias, Lucía López Amor, Elisabet Fernández Rey, Emilio García Prieto. UCI Hospital Cabueñes (Gijón): Débora Fernández Ruíz, Carla Martínez González

**Baleares:**

UCI hospital Universitario Son Llatzer (Palma de Mallorca): Lorenzo Socias, Marcio Borges-Sá, María Aranda Pérez, Antonia Socias. UCI Hospital Quirón Salud Pálmmaplanas (Palma de Mallorca): José Mª Bonell Goytisolo, Inmaculada Alcalde Mayayo, Carlos Corradini, Isabel Ceniceros, Edwin Rodríguez; UCI Hospital Universitario Son Espases (Palma de Mallorca): Jose Ignacio Ayestarrán Rota, Mariana Andrea Novo Novo, Joaquim Colomina Climent, Albert Figueras Castilla, Tomàs Leal Rullan, Maria Magdalena Garcia Sastre; UCI Hospital Comarcal d’Inca (Inca): Rossana Pèrez Senoff; UCI Hospital Mateu Orla (Mao): Ramón Fernández-Cid Bouza

**Canarias:**

UCI Complejo Hospitalario Universitario Insular - Materno Infantil (Las Palmas de G.C): Juan Carlos Martín González, Carmen Pérez Ortiz, José Luciano Cabrera Santana, Juan José Cáceres Agra, Domingo González Romero, Ana Casamitjana Ortega; UCI Hospital General de la Palma (Tenerife): Luis Alberto Ramos Gómez, Carolina Montelongo Ojeda; UCI Hospital Universitario Dr. Negrín (Las Palmas de G.C): Jordi Solé-Violán.

**Cataluña:**

UCI Hospital Universitari de Tarragona Joan XXIII (Tarragona): Alejandro Rodríguez, María Bodí, Gerard Moreno, Sandra Trefler, Laura Claverias, Raquel Carbonell, Erika Esteve, Montserrat Olona, Xavier Teixidó. UCI Hospital Universitari Arnau de Vilanova (Lleida): Monserrat Vallverdú Vidal, Beogaña Balsaera Garrido. UCI Hospital Universitari Vall d’Hebron (Barcelona): Elisabeth Papiol Gallofré, Raquel Albertos Martell, Rosa Alcaráz Peñarrocha, Xavier Nuvials Casals, Ricard Ferrer Roca; UCI Hospital Verge de la Cinta (Tortosa): Eric Adrián Mayor Vázquez, Ferrán Roche Campo, Pablo Concha Martínez, Diego Franch Llasat; UCI Hospital del Mar (Barcelona): Joan Ramón Masclanz, Judith Marin-Coral, Purificación Pérez, Rosana Muñoz, Clara Vila; UCI Hospital Mutua de Terrasa (Terrasa): Francisco Javier González de Molina, Elisabeth Navas Moya, Josep Trenado; UCI Hospital Sant Joan (Reus): Imma Vallverdú, Eric Castañé; UCI Hospital Parc Taulí (Sabadell): Emili Díaz Santos, Gemma Goma, Edgar Moglia;

**Cantabria**

UCI Hospital Universitario Marqués de Valdecillas (Santander): Borja Suberviola

**Castilla La Mancha:**

UCI Hospital Universitario de Guadalajara (Guadalajara): Antonio Albaya Moreno, Carlos Marian Crespo. UCI Hospital Nuestra Señora del Prado (Toledo): Carmen Carolina Sena Pérez, Francisca Arbol Linde

**Castilla y León:**

UCI Hospital Virgen de la Concha (Zamora): Diana Monge Donaire, Vega Losada Martínez, Nuria Rodrigo Castroviejo, Gerardo Ferrigno, Reyes Beltrán, Carolina Sanmartino, Concepción Tarancón Maján, Alfredo Marcos Gutiérrez; UCI Complejo Asistencial de Segovia (Segovia): Virginia Hidalgo Valverde, Caridad Martín López; UCI Hospital universitario de Burgos (Burgos): Oihane Badallo, María del Valle Ortiz, Rebeca Var Arlanzón, David Iglesias Posadilla; UCI Hospital Clínica de Salamanca (Salamanca): María Teresa Recio, Juan Carlos Ballesteros; UCI Complejo Asistencial Universitario de Palencia (Palencia):

**Ceuta**

UCI Hospital Universitario de Ceuta: Enrique Laza Laza
Extremadura
UCI Hospital San Pedro de Alcántara (Cáceres): Elena Gallego Curto, Mª Carmen Sánchez García; UCI Hospital de Mérida (Mérida): Miguel Díaz-Tavora, Rosa Mancha

Galicia:
UCI Hospital Montecelo (Ponteverda): Ana Ortega Montes, Isabel Gallego Barbachano, Eva Sanmartín Mantiñán. UCI CHUAC A Coruña (A Coruña): María Lourdes Cordero, Raquel María Rodríguez García, Jorge Gámez Zapata, María Gestal Vázquez. UCI Centro Hospital Universitario de Ferrol (Ferrol): María José Castro Orjales, María Isabel Álvarez Diéguez. UCI Hospitalario Clínico Universitario de Santiago (Santiago de Compostela): Carmen Rivero Velasco, Beatriz Lence Massa; REA CHUAC A Coruña (A Coruña): María Gestal Vázquez. UCI Hospital Lucus Augusti (Lugo): Ignacio Martínez Varela

Huelva:
UCI Hospital Infanta Elena (Huelva): Diego Matallana Zapata,

Madrid:
UCI IFEMA (Madrid): Alberto Hernández Tejedor; UCI Hospital Príncipe de Asturias (Madrid): Esther Mª López Ramos, Laura Alcázar Sánchez Elvira, Rocío Molina Montero, Mª Consuelo Pintado Delgado, María Trascasa Muñoz de la Peña, Yaiza Betania Ortiz de Zárate Ansoetegui, Alejandra Acha Aranda, Juan Higuera Lucas; UCI Hospital de la Princesa (Madrid): Juan Antonio Sanchez Giralt, Marta Chicot Llano, Nuria Arevallillo Fernández, Marta Sánchez Galindo, Ricardo Andino Ruiz, Alfonso Canabal Berlanga; UCI Hospital Clínico San Carlos (Madrid): Miguel Sánchez, Mercedes Nieto; UCI Hospital HLA la Moncloa (Madrid): Eduardo Arias Sarmiento, Adoración Bueno Blázquez, Rosa María de la Casa, Fátima Martín, Samuel González López,

Murcia
UCI Hospital Morales Meseguer (Murcia): Elena Martínez Quintana, Bernardo Gil Rueda, Áurea Higon Cañigral, Laura López Gómez, Pablo Safwat Bayoumi Delis, Augusto Montenegro Muore, Ángel Andrés Agamez Luengas; UCI Hospital Clínico Universitario Virgen de la Arrixaca (Murcia): Enrique Atande Soler, Ana Beatriz Pérez Pérez, José Higinio de Gea García, Rubén Jara Rubio, Silvia Sánchez Cámara, Alba Moreno Flores, José Moya Sánchez, Daniel Francisco Pérez Martínez, Mª Desamparados del Rey Carrión; UCI Hospital Reina Sofía (Murcia): María José Rico Lledó, Juana María Serrano Navarro, Juan Francisco Martín Ruiz, Julián Triviño Hidalgo, África López Ferrer, Isabel Cremades Navalón; UCI Hospital Santa Lucia (Cartagena): Josefa Murcia Payá, JM Allegre Gallego; UCI Hospital Rafael Méndez (Lorca): María del Carmen Lorente; UCI Hospital Universitario Mar Menor (San Javier): Marta Gonsalvez

Navarra
UCI Hospital Reina Sofía (Tudela): Ruth González Natera, Raquel Garrido López de Murillo, Tania Ojuel Gros, Raquel Flecha Viguera, Isabel López González; UCI Hospital García Orcoyen (Estella-Lizarra): Adriana García Herrera

País Vasco
UCI Hospital Universitario de Donostia (Donostia): Loreto Vidaur Tello, Maialen Aseguiolaza, Itziar Eguibar;

Sevilla:
UCI Hospital Universitario Virgen de la Macarena (Sevilla): María Luisa Cantón Bulnes, Jose Javier González Contreras, Helena Pérez Chomón, Nerissa Álvarez Chicote, Alberto Sousa González

Valencia:
Ferreres Franco, Ainhoa Serrano Lazaro, Mar Juan Díaz, María Luisa Blasco Cortés; UCI Hospital Virgen de los Lirios de Alcoy (Alicante): Laura Fayos, Julia Giménez, Gaspar Soriano, Ricardo Navarro. UCI Hospital Arnau de Vilanova (Valencia): Sonia Mas, Elena Bisbal, Laura Albert, Johncard Romero, Juan Fernández Cabreara; UCI Hospital Comarcal de Vinarós (Vinarós): Andrea Ortíz

Principado de Andorra

ICU Hospital Nostra Señyora de Meritxell (Les Esclades): Antonio MargaritRibas

References


http://dx.doi.org/10.1016/j.medin.2020.04.002


Tables
<table>
<thead>
<tr>
<th>Variablea,b</th>
<th>Overall</th>
<th>Train population</th>
<th>Test population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients (%)</td>
<td>2022 (100.0)</td>
<td>1618 (80.0)</td>
<td>404 (20.0)</td>
</tr>
<tr>
<td>Age, median (IQR), y</td>
<td>64(55–71)</td>
<td>64(55–71)</td>
<td>65(56–72)</td>
</tr>
<tr>
<td>Age &gt; 60 years, No. (%)</td>
<td>1254(62.0)</td>
<td>1006(62.2)</td>
<td>248(61.4)</td>
</tr>
<tr>
<td>Male, No. (%)</td>
<td>1423(70.4)</td>
<td>1129(69.7)</td>
<td>294(72.7)</td>
</tr>
<tr>
<td>APACHE II, median (IQR)</td>
<td>13(10–17)</td>
<td>14(11–18)</td>
<td>15(11–18)</td>
</tr>
<tr>
<td>SOFA, median (IQR)</td>
<td>5.0(3–7)</td>
<td>5.6(4.0–7.7)</td>
<td>5.2(4.0–7.0)</td>
</tr>
<tr>
<td>D-Lactate dehydrogenase, median (IQR), U/L</td>
<td>537(417–707)</td>
<td>535(410–690)</td>
<td>528(417–680)</td>
</tr>
<tr>
<td>White blood cell, median (IQR), x10^9</td>
<td>8.8(6.2–12.2)</td>
<td>8.8(6.2–12.3)</td>
<td>9.0(6.2–11.7)</td>
</tr>
<tr>
<td>Serum Creatinine, median (IQR), mg/dL</td>
<td>0.8(0.7–1.1)</td>
<td>0.8(0.7–1.1)</td>
<td>0.8(0.7–1.1)</td>
</tr>
<tr>
<td>C-Reactive Protein, median (IQR), mg/mL</td>
<td>15.5(9.1–24.3)</td>
<td>15.6(9.3–23.8)</td>
<td>15.7(8.6–24.5)</td>
</tr>
<tr>
<td>Procalcitonin, median (IQR), ng/mL</td>
<td>0.3(0.1–2.0)</td>
<td>0.3(0.1–0.8)</td>
<td>0.3(0.1–0.8)</td>
</tr>
<tr>
<td>Serum lactate, median (IQR), mmol/L</td>
<td>1.5(1.1–2.0)</td>
<td>1.5(1.1–2.0)</td>
<td>1.6(1.2–2.0)</td>
</tr>
<tr>
<td>Ferritin ng/mL, median (IQR),</td>
<td>1617(1290–2240)</td>
<td>1481(1200–1780)</td>
<td>1476(1252–1740)</td>
</tr>
<tr>
<td>Corticosteroids, No. (%)</td>
<td>1174(58.1)</td>
<td>942(58.2)</td>
<td>232(57.4)</td>
</tr>
<tr>
<td>Antibiotics, No. (%)</td>
<td>1818(89.9)</td>
<td>1452(89.7)</td>
<td>366(90.6)</td>
</tr>
<tr>
<td>Lopinavir/ritonavir, No. (%)</td>
<td>1662(82.2)</td>
<td>1321(81.6)</td>
<td>341(84.4)</td>
</tr>
<tr>
<td>Hydroxychloroquine, No. (%)</td>
<td>1866(92.3)</td>
<td>1492(92.2)</td>
<td>374(92.5)</td>
</tr>
<tr>
<td>Tocilizumab, No. (%)</td>
<td>577(28.5)</td>
<td>477(29.5)</td>
<td>100(24.7)</td>
</tr>
<tr>
<td>Interferon β, No. (%)</td>
<td>715(35.4)</td>
<td>568(35.1)</td>
<td>174(43.0)</td>
</tr>
<tr>
<td>Arterial hypertension, No. (%)</td>
<td>936(46.3)</td>
<td>752(46.5)</td>
<td>184(45.5)</td>
</tr>
<tr>
<td>Obesity, No. (%)</td>
<td>655(32.4)</td>
<td>527(35.2)</td>
<td>128(31.7)</td>
</tr>
<tr>
<td>Diabetes, No. (%)</td>
<td>420(20.8)</td>
<td>339(20.9)</td>
<td>81(20.1)</td>
</tr>
<tr>
<td>Coronary arterial disease, No. (%)</td>
<td>124(6.1)</td>
<td>96(5.9)</td>
<td>28(6.9)</td>
</tr>
<tr>
<td>COPD, No. (%)</td>
<td>148(7.3)</td>
<td>111(6.8)</td>
<td>37(9.1)</td>
</tr>
<tr>
<td>Chronic renal disease, No. (%)</td>
<td>85(4.2)</td>
<td>71(4.4)</td>
<td>14(3.5)</td>
</tr>
<tr>
<td>Hematologic disease, No. (%)</td>
<td>73(3.6)</td>
<td>62(3.8)</td>
<td>11(2.7)</td>
</tr>
<tr>
<td>Asthma, No. (%)</td>
<td>121(5.9)</td>
<td>98(6.0)</td>
<td>23(5.7)</td>
</tr>
<tr>
<td>HIV, No. (%)</td>
<td>5(2.0)</td>
<td>4(0.2)</td>
<td>1(0.2)</td>
</tr>
<tr>
<td>Pregnancy, No. (%)</td>
<td>4(0.2)</td>
<td>3(0.18)</td>
<td>1(0.2)</td>
</tr>
<tr>
<td>Variable(^{a,l})</td>
<td>Overall</td>
<td>Train population</td>
<td>Test population</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------</td>
<td>------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Autoimmune disease, No. (%)</td>
<td>74(3.7)</td>
<td>56(3.5)</td>
<td>18(4.4)</td>
</tr>
<tr>
<td>Chronic heart disease, No. (%)(^{g})</td>
<td>57(2.8)</td>
<td>46(2.8)</td>
<td>11(2.7)</td>
</tr>
<tr>
<td>Neuromuscular disease, No. (%)</td>
<td>16(0.8)</td>
<td>12(0.7)</td>
<td>4(1.0)</td>
</tr>
<tr>
<td>Other immunosuppression, No. (%)(^{h})</td>
<td>53(2.6)</td>
<td>46(2.8)</td>
<td>7(1.7)</td>
</tr>
</tbody>
</table>

Oxygenation and ventilator support

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall</th>
<th>Train population</th>
<th>Test population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxygen mask, No. (%)</td>
<td>333(6.4)</td>
<td>267(16.5)</td>
<td>66(16.3)</td>
</tr>
<tr>
<td>High Flow nasal cannula, No. (%)</td>
<td>375(18.5)</td>
<td>298(18.4)</td>
<td>77(19.1)</td>
</tr>
<tr>
<td>Non-invasive ventilation, No. (%)</td>
<td>140(6.9)</td>
<td>107(6.6)</td>
<td>33(8.1)</td>
</tr>
<tr>
<td>Invasive mechanical ventilation, No. (%)</td>
<td>1174(58.0)</td>
<td>941(58.1)</td>
<td>233(57.6)</td>
</tr>
</tbody>
</table>

Complications

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall</th>
<th>Train population</th>
<th>Test population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shock, No. (%)(^{i})</td>
<td>906(44.8)</td>
<td>729(45.0)</td>
<td>177(43.8)</td>
</tr>
<tr>
<td>Acute kidney dysfunction, No. (%)(^{j})</td>
<td>580(28.7)</td>
<td>477(29.5)</td>
<td>103(25.5)</td>
</tr>
<tr>
<td>Community-acquired co-infection(^{k}), No. (%)</td>
<td>190(9.4)</td>
<td>149(9.2)</td>
<td>41(10.1)</td>
</tr>
<tr>
<td>&gt;2 Quadrant infiltrates in chest x-ray, No. (%)</td>
<td>1333(65.9)</td>
<td>1062(65.6)</td>
<td>271(67.0)</td>
</tr>
<tr>
<td>ICU crude mortality, No. (%)</td>
<td>660(32.6)</td>
<td>528(32.6)</td>
<td>132(32.6)</td>
</tr>
</tbody>
</table>

Abbreviations: IQR, interquartile range; APACHE II, Acute Physiology and Chronic Health Evaluation II; SOFA, Sequential Organ Failure Assessment; BMI, body mass index; COPD, Chronic obstructive pulmonary disease; HIV, human immunodeficiency viruses; PaO2/FiO2, Partial pressure arterial oxygen/ fraction of inspired oxygen

\(^{a}\) Corresponds to minimum or maximum value, as appropriate, within 12 hours of ICU admission. The variables in this Table were no transformed for your comparison.

\(^{b}\) APACHE II score to the severity of illness, the score is obtained by adding the following components 1) 12 clinical and laboratory variables each with a score range of 0 to 4 points (APS). The APS is determined from the worst physiologic values during the initial 24 h after ICU admission, 2) age with a range 0 to 6 points and 3) Chronic health points if the patients has history of severe organ system insuciency or is immunocompromised assign 5 points if the patients is no operative or emergency postoperative and 2 points for elective postoperative patients with a total score range of 0 to 71.

\(^{c}\) SOFA score corresponds to the severity of organ dysfunction, reflecting six organ systems each with a score range of 0 to 4 points (cardiovascular, hepatic, hematologic, respiratory, neurological, renal), with a total score range of 0 to 24,

\(^{d}\) Defined as a body mass index (calculated as weight in kilograms divided by height in meters squared) of 30 or greater.

\(^{e}\) Baseline eGFR < 60 on at least two consecutive values at least 12 weeks apart prior or hemodialysis

\(^{f}\) included acute leukemia, myelodysplastic syndrome and Lymphomas.

\(^{g}\) According to the New York Heart Association (NYHA) Functional Classification III and IV

\(^{h}\) included Chronic corticosteroid treatment (> 20 mg prednisolone /day or equivalent dose), chemotherapy or therapy with biological agents

\(^{i}\) Defined as patients in whom adequate fluid resuscitation therapy are unable to restore hemodynamic stability and need any dose of vasopressor drugs.

\(^{j}\) Define as an abrupt and sustained (more than 24 hours) decrease in kidney function and categorized according to RIFLE criteria

\(^{k}\) Was considered in patients with confirmation of SARS-CoV-2 infection showing recurrence of fever, increase in cough and production of purulent sputum plus positive bacterial/fungal respiratory or blood cultures at ICU admission

\(^{l}\) Kruskal-Wallis, ANOVA, or chi-square p-value as appropriate comparing train and test population, p > .01 for all comparison
<table>
<thead>
<tr>
<th>Variable</th>
<th>Survival</th>
<th>No survival</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General characteristics and severity of illness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients(%)</td>
<td>1362(67.4)</td>
<td>660(32.6)</td>
<td></td>
</tr>
<tr>
<td>Age, median (IQR), y</td>
<td>61(53–69)</td>
<td>69(63–74)</td>
<td>.001</td>
</tr>
<tr>
<td>Male, No.(%)</td>
<td>939(68.9)</td>
<td>484(79.8)</td>
<td>.001</td>
</tr>
<tr>
<td>APACHE II(^b), median (IQR)</td>
<td>13(10–17)</td>
<td>17(12–21)</td>
<td>.001</td>
</tr>
<tr>
<td>SOFA(^c), median (IQR)</td>
<td>5(3–7)</td>
<td>7(5–8)</td>
<td>.001</td>
</tr>
<tr>
<td><strong>Laboratory findings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-Lactate dehydrogenase, median (IQR), U/L</td>
<td>507(397–667)</td>
<td>600(475–808)</td>
<td>.001</td>
</tr>
<tr>
<td>White blood cell, median (IQR), x10(^9)</td>
<td>8.4(6.1)</td>
<td>9.7(6.6–13.6)</td>
<td>.001</td>
</tr>
<tr>
<td>Serum Creatinine, median (IQR), mg/dL</td>
<td>0.8(0.6–1.0)</td>
<td>0.9(0.7–1.3)</td>
<td>.001</td>
</tr>
<tr>
<td>C-Reactive Protein, median (IQR) mg/mL</td>
<td>15(8.7–23.5)</td>
<td>17(10.2–26.4)</td>
<td>.001</td>
</tr>
<tr>
<td>Procalcitonin, median (IQR),ng/mL</td>
<td>0.3(0.1–0.7)</td>
<td>0.5(0.2–1.4)</td>
<td>.001</td>
</tr>
<tr>
<td>Serum lactate, median (IQR) mmol/L</td>
<td>1.4(1.1–1.9)</td>
<td>1.7(1.3–2.4)</td>
<td>.001</td>
</tr>
<tr>
<td>D dimer, median (IQR), ng/mL</td>
<td>1280(634–2900)</td>
<td>2230(1000–5190)</td>
<td>.001</td>
</tr>
<tr>
<td>Ferritin median (IQR), ng/mL</td>
<td>1530(1200–2000)</td>
<td>1900(145–2500)</td>
<td>.001</td>
</tr>
<tr>
<td><strong>Treatments</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticosteroids, No.(%)</td>
<td>777(57.0)</td>
<td>397(65.5)</td>
<td>.20</td>
</tr>
<tr>
<td>Antibiotics, No.(%)</td>
<td>1223(89.8)</td>
<td>595(98.1)</td>
<td>.86</td>
</tr>
<tr>
<td>Lopinavir/ritonavir, No.(%)</td>
<td>1111(81.6)</td>
<td>551(90.9)</td>
<td>.32</td>
</tr>
<tr>
<td>Hydroxychloroquine, No.(%)</td>
<td>1268(93.1)</td>
<td>598(98.6)</td>
<td>.06</td>
</tr>
<tr>
<td>Tocilizumab, No.(%)</td>
<td>407(29.9)</td>
<td>170(28.0)</td>
<td>.06</td>
</tr>
<tr>
<td>Interferon (\beta), No.(%)</td>
<td>456(33.5)</td>
<td>259(42.7)</td>
<td>.01</td>
</tr>
<tr>
<td><strong>Coexisting condition and Comorbidities</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arterial hypertension, No.(%)</td>
<td>561(41.2)</td>
<td>375(61.9)</td>
<td>.001</td>
</tr>
<tr>
<td>Obesity (d), No.(%)</td>
<td>431(31.6)</td>
<td>224(37.0)</td>
<td>.32</td>
</tr>
<tr>
<td>Diabetes, No.(%)</td>
<td>245(17.9)</td>
<td>175(28.8)</td>
<td>.001</td>
</tr>
<tr>
<td>Coronary arterial disease, No.(%)</td>
<td>58(4.2)</td>
<td>66(10.9)</td>
<td>.001</td>
</tr>
<tr>
<td>COPD, No.(%)</td>
<td>74(5.4)</td>
<td>74(12.2)</td>
<td>.001</td>
</tr>
<tr>
<td>Chronic renal disease (e), No.(%)</td>
<td>42(3.0)</td>
<td>43(7.0)</td>
<td>.001</td>
</tr>
<tr>
<td>Hematologic disease (f), No.(%)</td>
<td>36(2.6)</td>
<td>37(6.1)</td>
<td>.001</td>
</tr>
<tr>
<td>Asthma, No.(%)</td>
<td>83(6.0)</td>
<td>38(6.3)</td>
<td>.84</td>
</tr>
<tr>
<td>HIV, No.(%)</td>
<td>4(0.3)</td>
<td>1(0.1)</td>
<td>.89</td>
</tr>
<tr>
<td>Pregnancy, No.(%)</td>
<td>4(0.3)</td>
<td>0(0.0)</td>
<td>.38</td>
</tr>
<tr>
<td>Autoimmune disease, No.(%)</td>
<td>48(3.5)</td>
<td>26(4.3)</td>
<td>.73</td>
</tr>
<tr>
<td>Variable</td>
<td>Survival</td>
<td>No survival</td>
<td>p-value</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>Chronic heart disease</td>
<td>34(2.5)</td>
<td>23(3.8)</td>
<td>.26</td>
</tr>
<tr>
<td>Neuromuscular disease, No. (%)</td>
<td>7(0.5)</td>
<td>9(1.5)</td>
<td>.07</td>
</tr>
<tr>
<td>Other immunosuppression[^h], No. (%)</td>
<td>25(1.8)</td>
<td>28(4.6)</td>
<td>.002</td>
</tr>
</tbody>
</table>

**Oxygenation and ventilator support**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Survival</th>
<th>No survival</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PaO2/FiO2, median(IQR), mmHg</td>
<td>135(101–170)</td>
<td>121(85–151)**</td>
<td>.001</td>
</tr>
<tr>
<td>Oxygen mask, No. (%)</td>
<td>235(17.2)</td>
<td>93(15.3)</td>
<td>.08</td>
</tr>
<tr>
<td>High Flow nasal cannula, No. (%)</td>
<td>295(21.6)</td>
<td>80(13.2)**</td>
<td>.001</td>
</tr>
<tr>
<td>Non-invasive ventilation, No. (%)</td>
<td>98(7.2)</td>
<td>42(6.9)</td>
<td>.55</td>
</tr>
<tr>
<td>Invasive mechanical ventilation, No. (%)</td>
<td>716(52.2)</td>
<td>458(75.5)**</td>
<td>.001</td>
</tr>
</tbody>
</table>

**Complications**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Survival</th>
<th>No survival</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shock [^i], No. (%)</td>
<td>539(39.6)</td>
<td>367(60.5)</td>
<td>.001</td>
</tr>
<tr>
<td>Acute kidney dysfunction [^j], No. (%)</td>
<td>258(18.9)</td>
<td>322(53.1)</td>
<td>.001</td>
</tr>
<tr>
<td>Community-acquired co-infection [^k], No. (%)</td>
<td>117(8.6)</td>
<td>73(12.0)</td>
<td>.08</td>
</tr>
<tr>
<td>&gt; 2 Quadrant infiltrates in chest x-ray, No. (%)</td>
<td>865(63.5)</td>
<td>468(77.2)</td>
<td>.001</td>
</tr>
<tr>
<td>Cardiac dysfunction, No. (%)</td>
<td>73(5.3)</td>
<td>96(15.8)</td>
<td>.001</td>
</tr>
</tbody>
</table>

**Abbreviations:** IQR, interquartile range; APACHE II, Acute Physiology and Chronic Health Evaluation II; SOFA, Sequential Organ Failure Assessment; BMI, body mass index; COPD, Chronic obstructive pulmonary disease; HIV, human immunodeficiency viruses; PaO2/FiO2, Partial pressure arterial oxygen/ fraction of inspired oxygen

\[^a\] Corresponds to minimum or maximum value, as appropriate, within 12 hours of ICU admission. The variables in this Table were not transformed for your comparison.

\[^b\] APACHE II score to the severity of illness, the score is obtained by adding the following components 1) 12 clinical and laboratory variables each with a score range of 0 to 4 points (APS). The APS is determined from the worst physiologic values during the initial 24 h after ICU admission, 2) age with a range 0 to 6 points and 3) Chronic health points if the patients has history of severe organ system insuciency or is immunocompromised assign 5 points if the patients is no operative or emergency postoperative and 2 points for elective postoperative patients with a total score range of 0 to 71.

\[^c\] SOFA score corresponds to the severity of organ dysfunction, reflecting six organ systems each with a score range of 0 to 4 points (cardiovascular, hepatic, hematologic, respiratory, neurological, renal), with a total score range of 0 to 24,

\[^d\] Defined as a body mass index (calculated as weight in kilograms divided by height in meters squared) of 30 or greater.

\[^e\] Baseline eGFR < 60 on at least two consecutive values at least 12 weeks apart prior or hemodialysis

\[^f\] included acute leukemia, myelodysplastic syndrome and Lymphomas.

\[^g\] According to the New York Heart Association (NYHA) Functional Classification III and IV

\[^h\] included Chronic corticosteroid treatment (> 20 mg prednisolone /day or equivalent dose), chemotherapy or therapy with biological agents

\[^i\] Defined as patients in whom adequate fluid resuscitation therapy are unable to restore hemodynamic stability and need any dose of vasopressor drugs.

\[^j\] Define as an abrupt and sustained (more than 24 hours) decrease in kidney function and categorized according to RIFLE criteria

\[^k\] Was considered in patients with confirmation of SARS-CoV-2 infection showing recurrence of fever, increase in cough and production of purulent sputum plus positive bacterial/fungal respiratory or blood cultures at ICU admission

\[^l\] Kruskal-Wallis, ANOVA, or chi-square p-value as appropriate comparing survivors vs. non-survivors.
Table 3
Performance of global and clustering models (GLM model: generalized linear models with variables according ICU mortality table; Global model: GLM model with important variables according to information value analysis for ICU mortality; A, B and C Phenotypes: GLM models with important variables according to information value analysis for each cluster)

<table>
<thead>
<tr>
<th>Variable</th>
<th>GLM model</th>
<th>Global Model</th>
<th>A Phenotype</th>
<th>B Phenotype</th>
<th>C Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. patients</td>
<td>2022</td>
<td>2017</td>
<td>537</td>
<td>623</td>
<td>857</td>
</tr>
<tr>
<td>Variables included</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.78</td>
<td>0.77</td>
<td>0.86</td>
<td>0.80</td>
<td>0.72</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.88</td>
<td>0.88</td>
<td>0.94</td>
<td>0.92</td>
<td>0.71</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.45</td>
<td>0.54</td>
<td>0.55</td>
<td>0.47</td>
<td>0.73</td>
</tr>
<tr>
<td>AUC ROC</td>
<td>0.83</td>
<td>0.83</td>
<td>0.90</td>
<td>0.85</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Figures

Figure 1
Overview of the primary analysis plan. ICU= intensive care units, PAM= partition around medoids clustering analysis, GLM= Generalized Linear model
Figure 2

Variables independently associated with ICU mortality in multivariable analysis (GLM: Generalized Linear Model). Data are show as OR (Odds Ratio) and 95% Confidence interval. (SOFA: Sequential Organ Failure Assessment; PCT: Procalcitonin; PaO2/FiO2: Partial pressure arterial oxygen/ fraction of inspired oxygen; Dysf: Dysfunction; LDH: D-Lactate dehydrogenase; MV: Mechanical ventilation; AKI: Acute Kidney injury; >2 infiltrates: > 2 infiltrates in chest-x ray)
Figure 3

Phenotype clinical characterization (APACHE II= Acute Physiology and Chronic Health Evaluation II; SOFA= Sequential Organ Failure Assessment; LDH= D-Lactate dehydrogenase, U/L; AKI= Acute Kidney injury;)

Figure 4

(A) Chord Diagrams Showing Abnormal Clinical Variables by Phenotype. A: mild COVID-19 disease; B: moderate COVID-19 disease and C: severe COVID-19 disease. (B) Chord Diagrams Showing Abnormal Clinical Variables by Phenotype differentiating survivors (green) from non-survivors (red). (APACHE II= Acute Physiology and Chronic Health Evaluation II; SOFA= Sequential Organ Failure Assessment; PCT = Procalcitonin; >3 chest x-ray= more than 3 quadrants infiltrates in the chest x-ray; Miocard Dys = Myocardial dysfunction; Hydroxichloroq.= Hydroxychloroquine; GAP antiviral= Time in days from onset of symptoms to first dose of antiviral; DD= D dimer; AKI= Acute Kidney injury ; LDH= D-Lactate dehydrogenase, U/L; COPD= Chronic Pulmonary Obstructive Disease; Pa/Fi = Partial pressure arterial oxygen/ fraction of inspired oxygen; Hemat. Dis= Hematologic disease; GAP_UCI= Time in days from Hospital to ICU admission; Coronary dis.= Coronary disease)
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile.pdf