



Original article

## Risk Factors of Mortality in Hospitalized Patients With COVID-19 Applying a Machine Learning Algorithm



Irene Nieto-Codesido<sup>a,\*</sup>, Uxio Calvo-Alvarez<sup>a</sup>, Carmen Diego<sup>a</sup>, Z. Hammouri<sup>b</sup>, Narmeen Mallah<sup>c</sup>, María José Ginzo-Villamayor<sup>d</sup>, Francisco Javier Salgado<sup>e,f</sup>, José Martín Carreira<sup>c</sup>, Carlos Rábade<sup>g</sup>, Gema Barbeito<sup>h</sup>, Miguel Ángel Gonzalez-Perez<sup>i</sup>, Francisco Javier Gonzalez-Barcala<sup>c,f,g,j</sup>

<sup>a</sup> Department of Respiratory Medicine, University Hospital of Ferrol, Spain

<sup>b</sup> University of Santiago de Compostela, Spain

<sup>c</sup> Department of Psychiatry, Radiology, Public Health, Nursery and Medicine, Universidad de Santiago de Compostela, Spain

<sup>d</sup> Departamento de Estadística, Análisis Matemático y Optimización, Facultad de Matemáticas, Universidad de Santiago de Compostela, Spain

<sup>e</sup> Department of Biochemistry and Molecular Biology, Faculty of Biology-Biological Research Centre (CIBUS), Universidade de Santiago de Compostela, Spain

<sup>f</sup> Translational Research In Airway Diseases Group (TRIAD)-Health Research Institute of Santiago de Compostela (FIDIS), Spain

<sup>g</sup> Department of Respiratory Medicine, University Hospital of Santiago de Compostela, Spain

<sup>h</sup> Microbiology Department, University Hospital of Santiago de Compostela, Spain

<sup>i</sup> Centro de Salud Valle-Inclán, Orense, Spain

<sup>j</sup> Spanish Biomedical Research Networking Centre-CIBERES, Madrid, Spain

### ARTICLE INFO

#### Article history:

Received 3 November 2021

Accepted 27 January 2022

Available online 6 February 2022

#### Keywords:

Prognosis

Severity

Age

Gender

Machine learning

### ABSTRACT

**Introduction:** Risk stratification of patients with COVID-19 can be fundamental to support clinical decision-making and optimize resources. The objective of our study is to identify among the routinely tested clinical and analytical parameters those that would allow us to determine patients with the highest risk of dying from COVID-19.

**Material and methods:** We carried out a retrospective cohort multicentric study by consecutively, including hospitalized patients with COVID-19 admitted in any of the 11 hospitals in the healthcare network of HM Hospitals-Spain. We collected the clinical, demographic, analytical, and radiological data from the patient's medical records.

To assess each of the biomarkers' predictive impact and measure the statistical significance of the variables involved in the analysis, we applied a random forest with a permutation method. We used the similarity measure induced by a previously classification model and adjusted the k-groups clustering algorithm based on the energy distance to stratify patients into a high and low-risk group. Finally, we adjusted two optimal classification trees to have a schematic representation of the cut-off points.

**Results:** We included 1246 patients (average age of 65.36 years, 62% males). During the study one hundred sixty-eight patients (13%) died. High values of age, D-Dimer, White Blood Cell, Na, CRP, and creatinine represent the factors that identify high-risk patients who would die.

**Conclusions:** Age seems to be the primary predictor of mortality in patients with SARS-CoV-2 infection, while the impact of acute phase reactants and blood cellularity is also highly relevant.

© 2022 Published by Elsevier España, S.L.U. on behalf of Sociedad Española de Neumología y Cirugía Torácica (SEPAR). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

\* Corresponding author.

E-mail address: [irene.nieto.codesido@gmail.com](mailto:irene.nieto.codesido@gmail.com) (I. Nieto-Codesido).

## Factores de riesgo de mortalidad en pacientes hospitalizados con COVID-19 aplicando un algoritmo de aprendizaje automático

### R E S U M E N

#### Palabras clave:

Pronóstico  
Gravedad  
Edad  
Sexo  
Aprendizaje automático

**Introducción:** La estratificación del riesgo de los pacientes con COVID-19 puede ser fundamental para apoyar la toma de decisiones clínicas y optimizar los recursos. El objetivo de nuestro estudio es identificar, entre los parámetros clínicos y analíticos probados de forma rutinaria, aquellos que nos permitirían determinar a los pacientes con mayor riesgo de morir por COVID-19.

**Material y métodos:** Se realizó un estudio multicéntrico de cohorte retrospectiva de forma consecutiva, incluyendo pacientes hospitalizados con COVID-19 ingresados en cualquiera de los 11 hospitales de la red sanitaria de HM Hospitales-España.

Los datos clínicos, demográficos, analíticos y radiológicos se recopilaron de las historias clínicas de los pacientes.

Para evaluar el impacto predictivo de cada uno de los biomarcadores y medir la significación estadística de las variables involucradas en el análisis, se aplicó un bosque aleatorio con un método de permutación. Utilizamos la medida de similitud inducida por un modelo de clasificación previo, y ajustamos el algoritmo de agrupación de grupos  $k$  en función de la distancia de energía para estratificar a los pacientes en un grupo de alto y bajo riesgo. Finalmente, ajustamos 2 árboles de clasificación óptimos para tener una representación esquemática de los puntos de corte.

**Resultados:** Se incluyeron 1.246 pacientes (edad promedio de 65,36 años, 62% varones). Durante el estudio murieron 168 pacientes (13%). Los factores que identifican a los pacientes de alto riesgo de mortalidad son los valores elevados de edad, dímero D, glóbulos blancos, Na, PCR y creatinina.

**Conclusiones:** La edad parece ser el principal predictor de mortalidad en pacientes con infección por SARS-CoV-2, mientras que el impacto de los reactantes de fase aguda y la celularidad sanguínea también es muy relevante.

© 2022 Publicado por Elsevier España, S.L.U. en nombre de Sociedad Española de Neumología y Cirugía Torácica (SEPAR). Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introduction

In December 2019, an outbreak of pneumonia of unknown origin occurred in Wuhan, China, and then rapidly spread worldwide.<sup>1,2</sup> The new virus was named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), and its associated disease was called Coronavirus disease [COVID-19].<sup>1-3</sup> On March 11th, 2020, the World Health Organization [WHO] declared the COVID-19 outbreak a pandemic.<sup>4</sup> According to WHO data, on October 29th, 2021, 245,373,039 confirmed cases of COVID-19, causing 4,979,421 deaths.<sup>5</sup>

Patients with COVID-19 usually have mild or moderate symptoms. However, in some cases, the disease is more severe, and affected patients may develop respiratory failure, Acute Respiratory Distress Syndrome [ARDS], multi-organ dysfunction, and death.<sup>2,6-8</sup>

The coronavirus disease is expected to affect the clinical practice significantly through the imposed need to implement critical and intermediate care units, especially in specific environments such as pulmonary function laboratories or bronchoscopy, and to expand further the use of telemedicine.<sup>9-19</sup> Likewise, health systems have been overburdened which could raise ethical issues of therapeutic limitation in patients with low chances of survival.<sup>20</sup>

Risk stratification of patients with COVID-19 can be fundamental to support clinical decision-making and optimize resources.<sup>21</sup> However, some research studies' findings have poor reproducibility due to the small sample size, limitations of the study design, and simple statistical and machine learning models employed in the analysis.<sup>1,3,22-32</sup> State-of-the-art machine learning models can provide new clinical knowledge using models capable of detecting higher-order interactions between variables.

Previous knowledge of the disease suggests that older age, higher serum Lactic Acid Dehydrogenase [LDH], or lymphopenia are related to a worse evolution of COVID-19; nonetheless, there are discrepancies in the literature regarding other routinely examined parameters such as hemoglobin or platelets.<sup>1,3</sup> On the other

hand, given that COVID-19 affects all parts of the world, it would be especially relevant to identify predictors of disease evolution that are regularly used in clinical practice, are inexpensive and may be applied in any socio-economic environment.

Our study's objective is to identify among the routinely tested clinical and analytical parameters that would allow us to determine patients with the highest risk of dying from COVID-19.

## Materials and methods

We carried out a retrospective cohort multicentric study, carried-out according to usual clinical practice including consecutively hospitalized patients with COVID-19. Nasopharyngeal swab or sputum samples were collected from each patient admitted between March 1st and April 25th, 2020, in any of the 11 hospitals in the healthcare network of HM Hospitals, a leading hospital group in Spain, in the regions of Madrid, Catalonia, Castilla-León and Galicia. The diagnosis of COVID-19 was then confirmed by Real-Time Reverse-Transcription Polymerase Chain Reaction (RT-PCR).

We collected the clinical data from the patients' medical records and considered only the earliest test results that were performed during hospitalization. In addition, for each patient, we registered demographic, analytical, and radiological data. Laboratory tests were performed at hospital admission and consisted of a complete blood count, blood chemical analysis, coagulation testing, renal function assessment, serum sodium measures, C-reactive protein, lactate dehydrogenase, and D-dimer. We determined the presence of a radiologic abnormality by relying on the description of the chest radiography results that was registered in the clinical records.

The Ethics Committee of Galicia approved the study [2020/225].

## Statistical analysis

To assess each of the biomarkers' predictive impact and measure the statistical significance of the variables involved in the

**Table 1**  
Basic statistics of the continuous biomarkers in live and alive patients. The mean and standard deviation are shown.

	Total (n = 1246)		Live (n = 1078)		Alives (n = 168)	
	Mean	SD	Mean	SD	Mean	SD
Age, years	65.36	± 14.959	62.889	± 14.152	81.214	± 9.182
Creatinine, mg/dL	0.985	± 0.563	0.909	± 0.322	1.471	± 1.192
D-Dimer, ng/mL	2088.562	± 7450.976	1572.801	± 5371.999	5398.03	± 14,664.197
Eosinophils-Abs, ng/mL	0.029	± 0.058	0.032	± 0.061	0.015	± 0.032
Haemoglobin, g/dL	13.939	± 1.684	14.043	± 1.619	13.277	± 1.927
LDH, U/L	615.387	± 360.271	578.662	± 253.47	851.042	± 699.092
White blood cell, Cel/μL	7.221	± 4.331	6.874	± 3.382	9.448	± 7.767
Na, Mmol/L	136.7	± 4.286	136.532	± 3.695	137.78	± 6.897
Neutrophils-Abs, Cel/μL	5.425	± 3.315	5.106	± 2.993	7.469	± 4.394
C Reactive protein, mg/L	105.275	± 99.359	94.476	± 90.351	174.572	± 123.971
Platelets, Cel/μL	220.152	± 89.75	222.391	± 90.002	205.78	± 87.009

Abs, absolute number; LDH, lactate dehydrogenase.

analysis in a multivariate model, we applied a Random Forest algorithm with a permutation's method. Moreover, to assess the predictive capacity of Random Forest in a fair situation where there is no imbalance between live and dead patients, we applied the SMOTE technique, and using a ROC curve, we estimated the AUC over the original sample. Next, we used the similarity measure induced by the Random Forest<sup>33,34</sup> and adjusted the k-groups clustering<sup>35,36</sup> based on the energy distance to stratify patients into high and low-risk groups. Finally, we adjusted two optimal classification trees to have a schematic representation of the cut-off points that differentiate between: (i) high and low-risk patients; (ii) high and low-risk patients who survive, and high-risk patients who die from COVID-19. Moreover, we introduce a nonparametric kernel density estimator of each continuous biomarker from each below group to assess each variable's impact in the cluster discrimination.

For all the calculated p-values, we applied a correction for multiple comparisons according to the criterion of False Discovery Rate [FDR] under dependency assumptions. All the statistical analyses were carried out using R software. In specific, we used the energy package, Random Forest, and functions implemented by default in the software.

**Results**

With an average age of 65.36 years, one thousand two hundred forty-six patients were included (average age 65.36 years, 62% of them males). One hundred sixty-eight patients (13%) died during the hospital admission, most of whom were older adults (Table 1).

Through Random Forest, we established each variable's impact using a model that included all adjusted markers concurrently (Table 2 and Fig. 1). Age has the highest predictive capacity in estimating mortality, followed by LDH, D-Dimer, C-Reactive Protein (CRP), and Sodium (Na), which jointly have almost one-third of the predictive capacity. Gender, diagnosis of pneumonia, platelets, and hemoglobin are not statistically significant variables for mortality predictors.

The AUC on the original sample and applying previously the SMOTE unbalance method is 0.97.

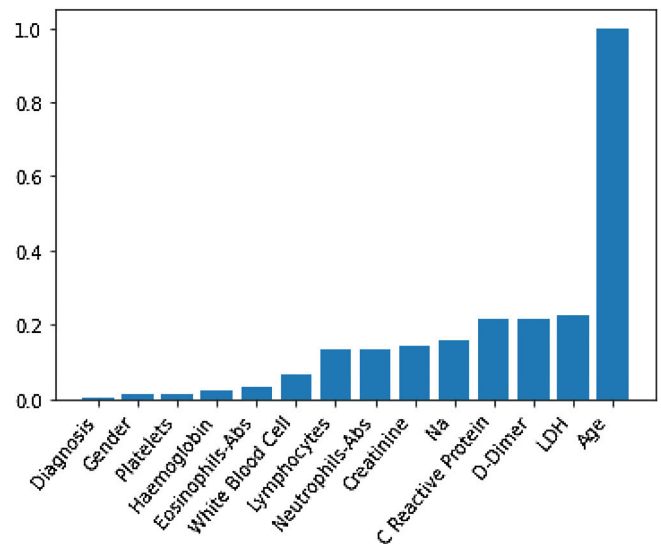
A semi-supervised cluster analysis was carried out using methods based on energy distance from the previously adjusted random forest. We obtained two groups: patients at risk of death (98% of the deceased patients were in this group) were at low risk of death and would survive, patients who were at high risk of death but who would survive, and patients who were at high risk of death and would die from COVID-19.

Patients with high D-Dimer values, age, LDH, CRP, and creatinine belong to the risk group. In contrast, high values of age, D-Dimer, White Blood Cell, Na, CRP, and creatinine represent the factors that

**Table 2**  
Predictive capacity of each variable in the Random Forest together with the p-value of testing the null hypothesis that there is no effect of the biomarker in the multivariate model. The impact capacity is normalized respect the score obtained by the age, the variable more influential.

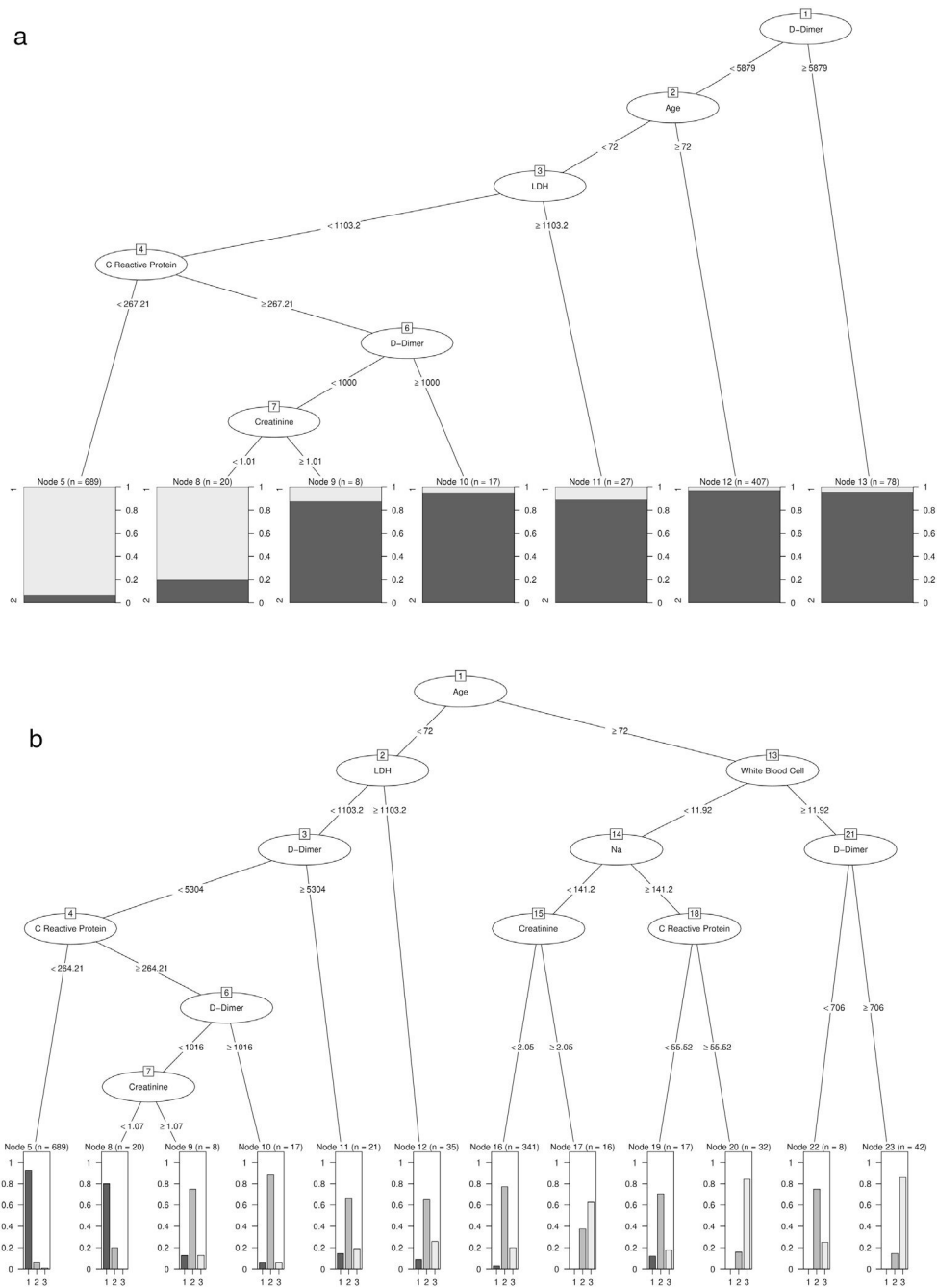
	Impact variable (normalized by age)	p-Value
Age, years	1	<0.020
Creatinine, mg/dL	0.14	<0.020
D-Dimer, ng/mL	0.22	<0.020
Eosinophils-Abs, ng/mL	0.033	0.042
Haemoglobin, g/dL	0.021	0.099
LDH, U/L	0.22	<0.020
White blood cell, Cel/μL	0.07	<0.020
Na, Mmol/L	0.16	<0.020
Neutrophils-Abs, Cel/μL	0.14	<0.020
C reactive protein, mg/L	0.21	<0.020
Platelets, Cel/μL	0.012	<0.87
Gender	0.012	0.082
Diagnosis	0.002	0.501

Abs, absolute number; LDH, lactate dehydrogenase.



**Fig. 1.** Predictive capacity of each variable, according to Random Forest. The values are normalized by the age score, the variable with the highest predictive power.

identify high-risk patients who would die. Further details on cut-off points are available in Fig. 2a and b. Moreover, the univariate distribution of each continuous variable in each group under consideration is shown in Fig. 3.



**Fig. 2.** (a) Result of estimating the optimal classification tree that discriminates between the risk and non-risk group. (b) Results of estimation of the optimal classification tree that discriminates between the risk patients who died and lived, and the patients included in the non-risk group.

**Discussion**

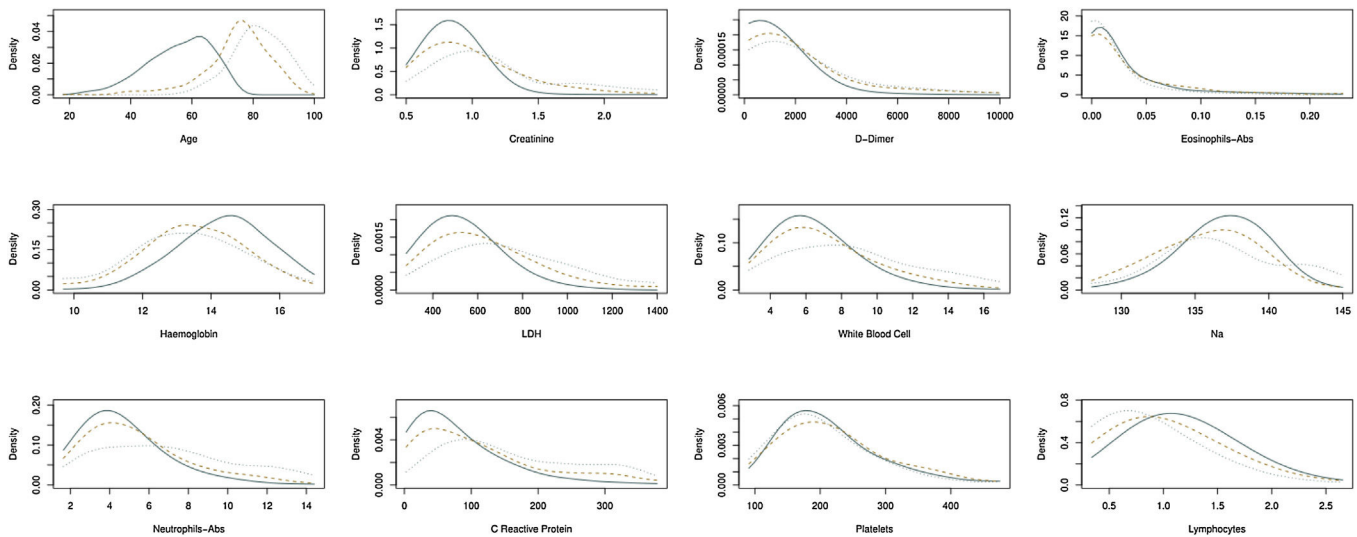
Early identification of COVID-19 patients at high risk of death may prove useful in optimizing patients' management and healthcare administration. The latter is especially relevant in large epidemics, where a high number of patients request health care simultaneously.

In our study, in addition to age and sex, we included widely available basic analytical parameters, as well as data from simple chest radiology.

Among this study participants, 13% died, and most of them were elderly patients (mean age: 81 years). This study's mortality rate is in the average of other published studies.<sup>1,3,22-32</sup> This mortality rate might be considered acceptable given that the

average age of our patients is higher than that reported in previous publications.<sup>1,3,22-32</sup>

Older age has been consistently associated with increased mortality from COVID-19 in studies conducted in different countries.<sup>1,23-26,32</sup> However, upon the inclusion of seriously ill patients, no significant relationship was observed between age and mortality.<sup>27,28</sup> Various factors seem to be associated with a higher mortality rate among elderly patients. With age advancement, the immune system deteriorates. Aging disrupts innate immunity functions such as antigen presentation, phagocytosis, or the expression of pattern recognition receptors. It also reduces the production and the activity of T and B cells. These alterations also contribute to increased neutrophil recruitment and inflammatory activity.<sup>36</sup>



**Fig. 3.** Estimation of the density function in continuous biomarkers for patients that belong in three groups under consideration: (i) patients who belong to the risk group and live; (ii) patients who belong to the risk group and die, and (iii) patients who belong to the non-risk group. Continuous line, patients belong to the non-risk group. Dotted line, patients who belong to the risk group and die. Another case, patients who belong to the risk group and do not die.

Furthermore, with aging, lung function decreases, and the pulmonary epithelial barrier deteriorates, increasing vulnerability to infections.<sup>36</sup> Animal models demonstrated that aging alters the inflammatory response, and thus, increases mortality from viral infections. Moreover, neutrophils are excessively recruited, resulting in an enhancement of interleukin release.<sup>36</sup> Aging also alters alveolar macrophages' activity, decreases their number, inhibits phagocytosis, and reduces their neutrophil apoptosis capacity, making it difficult to resolve the inflammatory process.<sup>36</sup>

Acute phase reactants participate in the defence process against infections.<sup>37–39</sup> In our study, after age, inflammatory markers were found to have the highest predictive capacity for mortality from COVID-19, which is consistent with findings from other studies that reported a significant relationship between elevated values of LDH, CRP, or D-Dimer and a higher risk of mortality from COVID-19.<sup>3,23,29,30</sup> Cytokine storm is a complex process that can damage the respiratory epithelium, endothelial cell apoptosis, leading to alveolar edema, ARDS, multiple organ failure, and death.<sup>37,38</sup> It is induced by the activation of many leukocytes, including B cells, T cells, NK cells, macrophages, dendritic cells, neutrophils, monocytes, as well as by the activation of epithelial and endothelial cells that release large amounts of inflammatory cytokines.<sup>39</sup>

The male gender association with increased mortality risk from COVID-19 is consistent with data from earlier studies.<sup>1,22,24,26,32</sup> Some pathophysiological mechanisms were described and suggested to explain the decreased risk of mortality from COVID-19 among women. Sexual dimorphism in the immune response is well known. Animal models revealed that the influenza infection vaccine is more efficient in females as they have a better antibody response than males.<sup>40</sup> Another possible mechanism could be explained by the function of Angiotensin Converting Enzyme2 [ACE2], which serves as a receptor for the entry of SARS-CoV-2 into the cells and therefore plays a fundamental role in the pathogenesis of COVID-19. The ACE2 gene is located on the X chromosome, contributing to phenotypic differences between sexes.<sup>41</sup>

On the other hand, women have higher levels of estrogen hormone that regulates the expression of ACE2 and consequently promotes anti-inflammatory effects. Likewise, estrogen level inversely correlates with troponin release by cardiomyocytes that are exposed to ischemia or hypoxia.<sup>41</sup> Innate and adaptive immunity significantly differ between men and women. In the innate immune system, females have greater efficiency of

antigen-presenting cells, higher phagocytic capacity of neutrophils and macrophages, or greater expression of Toll-like receptors 7. In contrast, males have a more significant expression of Toll-like receptors 4 in macrophages or many natural killer cells. In the adaptive immune system, men have more CD8+ T cells, whereas women have a higher number of B cells, more enhanced production of antibodies, or cytotoxic T cells.<sup>42,43</sup> Another hypothesis is that the age or the comorbidity could vary between sexes, and this would justify, at least in part, the difference in mortality rate between genders. Nonetheless, our results do not support this hypothesis.<sup>44</sup>

Patients who died had a higher leukocyte and neutrophil counts, as well as lower eosinophil and lymphocyte counts than patients who survived. However, we did not observe any significant differences in terms of hemoglobin or platelets. The association of low counts of lymphocytes with greater disease severity is almost constant in the literature.<sup>1,3,22–32</sup> Zhou and colleagues analysed 191 patients and reported a significant relationship between mortality and lower blood lymphocyte counts in the univariate analysis.<sup>1</sup> However, the effect was not maintained in the multivariate analysis, given the greater weight of age, D-dimer, and SOFA score.<sup>1</sup> The mechanism behind this reduction in blood lymphocyte count is unclear. In SARS and MERS, viruses were shown to induce lymphocyte apoptosis in the lung and the liver, which could also be provoked by SARS-CoV-2.<sup>39</sup> The possibility of IL-1beta-induced pyroptosis was also proposed.<sup>45</sup> Another hypothesis concerning decreased blood lymphocyte count is that it could be due to the cytopathic effect produced by direct infection of T lymphocytes.<sup>39</sup> There could also be a relevant involvement of Treg cells. Likewise, cytokines released from SARS-CoV-2 infection attract lymphocytes from the blood into the infected site.<sup>45</sup> These cells play a fundamental role in the airway mucosa by suppressing effector cells and damaging tissue mechanisms, limiting the pulmonary immunopathological process. Decreased counts of Treg cells were observed in patients infected with SARS-CoV-2, especially in severe cases, which could also contribute to lymphopenia.<sup>46</sup>

Eosinophils have antiviral activity and participate in adaptive immunity.<sup>39</sup> Eosinopenia is also a common finding in patients admitted for COVID-19 and is significantly related to higher mortality rate in these patients.<sup>1,3,22–32</sup> Multiple factors can cause it. Eosinopenia can be due to the high migration rate of eosinophils from the peripheral blood to the infected organ. It could also be the consequence of an increased Th1 response that would also

antagonize the Th2 response. Other possible causes of eosinopenia that have been considered include the blockage of eosinophilopoiesis in the bone marrow and the apoptosis of eosinophils induced by inflammatory mediators released during acute infection, such as type 1 interferon.<sup>39</sup>

The neutrophil count relationship with mortality from COVID-19 is not consistent in the literature, since while some studies associated neutrophil count with a higher mortality rate,<sup>25,27,31</sup> others did not find any association.<sup>1,32</sup> Neutrophils represent a critical point in the immune response and form part of the innate immune defence. However, their role in viral infections remains unclear. It is known that neutrophils participate in various mechanisms involved in the pathophysiology of COVID-19. SARS-CoV-2 enters cells together with angiotensin-converting enzyme II [ACE2], reduces its expression, and favours neutrophilic infiltration.<sup>47</sup> The infiltration and the degranulation of neutrophils at the infection site provoke cytokine release and contribute to the cytokine storm. Neutrophils also seem to be involved in the formation of thrombi in the blood vessels.<sup>47</sup> Likewise, hyperglycemia, another factor associated with a worse prognosis in patients with COVID-19, favors the release of inflammatory mediators by neutrophils.<sup>47</sup>

Given the high number of patients who have developed this infection it seems necessary to protocolize their follow-up with clear criteria for referring patients to the corresponding specialized units.<sup>48</sup>

Furthermore, considering that vaccines are safe and effective, an active involvement of healthcare professionals in the vaccination strategy is a key point to improve the vaccination coverage.<sup>49,50</sup>

Our study has some limitations, most of which are inherent to retrospective studies. On the one hand, we did not consider the effect of comorbidities as this diagnosis was not available. Furthermore, we did not evaluate the effect of treatments performed during admission, which might modify some patients' evolution. Similarly, there was a lack of survival signs-related data for some patients, such as cardiac troponin I or N-terminal pro-brain natriuretic peptide. Furthermore, all patients were recruited from Spanish private healthcare centers. Therefore, our study may not represent the general population, which would limit our findings' external validity.

Our study also has some strengths. The use of machine learning techniques of the state-of-the-art has the potential ability to extract new information on the biological mechanisms involved between the studied biomarkers and the risk of death in patients with COVID-19. This information is obtained while accounting for the possible non-linear relationship between the biomarkers and the response variables. Undoubtedly, simpler models such as logistic regression cannot achieve this goal in many scenarios. On the other hand, explainable artificial intelligence techniques such as optimal classification trees enable a simple explanation of the findings from a set of easily interpretable clinical rules that can be easily used in real clinical situations. Our study's other strengths are its multicenter nature, the inclusion of a large sample size, and the analysis of clinical variables widely available in all healthcare systems.

In conclusion, age seems to be the primary predictor of mortality in patients with SARS-CoV-2 infection, while the impact of acute phase reactants and blood cellularity is also highly relevant.

## Role of the funding source

There was no financial support for the conduct of the research and/or preparation of the article.

## Conflict of interest

- Francisco-Javier Gonzalez-Barcala has received honoraria for consultancy, projects or presentations from Chiesi, Menarini, Rovi, Bial, GlaxoSmithKline, Laboratorios Esteve, Teva, GebroPharma, ALK, Roxall, Stallergenes-Greer, Boehringer Ingelheim, Mundipharma and Novartis.
- Francisco-Javier González-Barcala, associate editor of Open Respiratory Archives, has not participated in the editorial process of this article.
- Irene Nieto Codesido has received honoraria for consultancy, projects or presentations from Chiesi, GlaxoSmithKline, Laboratorios Esteve, Bial and Mundipharma.

## Acknowledgements

HM Hospitals has provided the data for this article.

## References

1. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395:1054–62.
2. Du Y, Tu L, Zhu P, Mu M, Wang R, Yang P, et al. Clinical features of 85 fatal cases of COVID-19 from Wuhan. A retrospective observational study. *Am J Respir Crit Care Med*. 2020;201:1372–9.
3. Zhang J, Wang X, Jia X, Li J, Hu K, Chen G, et al. Risk factors for disease severity, unimprovement, and mortality in COVID-19 patients in Wuhan, China. *Clin Microbiol Infect*. 2020;26:767–72.
4. WHO Virtual press conference on COVID. [https://www.who.int/docs/default-source/coronavirus/transcripts/who-audio-emergencies-coronavirus-press-conference-full-and-final-11mar2020.pdf?sfvrsn=cb432bb3\\_2](https://www.who.int/docs/default-source/coronavirus/transcripts/who-audio-emergencies-coronavirus-press-conference-full-and-final-11mar2020.pdf?sfvrsn=cb432bb3_2) [accessed 11.3.20].
5. WHO Coronavirus (COVID-19) Dashboard. <https://covid19.who.int/> [accessed 31.10.21].
6. Déniz-Armengol C, Ramos-Izquierdo R, Ureña-Lliveras A. Patient with pneumonia caused by COVID-19: surgical findings and radiological correlation of lung lesions. *Arch Bronconeumol*. 2020;56(S2):33.
7. Perez-Girbes A. Acute pulmonary embolism and Covid-19: a common association in seriously ill patients? *Arch Bronconeumol*. 2020;56(S2):34.
8. Bao J, Li C, Zhang K, Kang H, Chen W, Gu B. Comparative analysis of laboratory indexes of severe and non-severe patients infected with COVID-19. *Clin Chim Acta*. 2020;509:180–94.
9. Montserrat-Canal JM, Suarez-Giron M, Egea C, Embid C, Matute-Villaciés M, de Manuel Martínez L, et al. Spanish Society of Pulmonology and Thoracic Surgery positioning on the use of telemedicine in sleep-disordered breathing and mechanical ventilation. *Arch Bronconeumol*. 2020;S0300-2896(20)30190-3.
10. Isea de la Viña J, Mayol J, Ortega AL, Alcazar-Navarrete B. Lung cancer patients on the waiting list in the midst of the COVID-19 crisis: what do we do now? *Arch Bronconeumol*. 2020. <http://dx.doi.org/10.1016/j.arbres.2020.05.004>.
11. Casan-Clara P, Martínez-Gonzalez C. Post COVID-19 pneumology. *Arch Bronconeumol*. 2020;56 Suppl. 2:3–4.
12. Izquierdo-Alonso JL, Rodríguez-González-Moro JM. Pneumology in the days of COVID-19. *Arch Bronconeumol*. 2020;56 Suppl. 2:1–2.
13. Barreiro E, Jiménez C, García de Pedro J, Ramírez Prieto MT. COVID-19 and XXI century pulmonology: challenge or opportunity? *Arch Bronconeumol*. 2020;56:411–2.
14. Cordovilla R, Álvarez S, Llanos L, Núñez Arce A, Cases Viedma E, Díaz-Pérez D, et al. SEPAR and AEER consensus recommendations on the use of bronchoscopy and airway sampling in patients with suspected or confirmed COVID-19 infection. *Arch Bronconeumol*. 2020;56 Suppl. 2:19–26.
15. Fraile-Olivero CA, Jarabo-Sarceda JR, Cerdán-Santacruz C, Santos Capa P, Arribas Manzanal PD, Barriuso R, et al. When the need is greater than feasible. *Arch Bronconeumol*. 2020. S0300-2896(20)30174-5.
16. Ramírez-Prieto MT, Moreno-Zabaleta R, García-Romero-de-Tejada JA. COVID-19 Neumonia. What does ventilatory support offer? Review of a cohort of patients treated in a second level hospital. *Open Respir Arch*. 2020;2:297–305.
17. Duarte de Araújo A, Correia-de-Sousa J. COPD: will there be room for nebulisers after the current COVID-19 pandemic? *Open Respir Arch*. 2020;2:265–6.
18. Corbacho-Abelaira MD, Corbacho-Abelaira F, Ruano-Ravina A, Fernández-Villar A. Use of conventional chest imaging and artificial intelligence in COVID-19 infection. A review of the literature. *Open Respir Arch*. 2021;3:100078.
19. Lopez-Padilla D, Teran-Tinedo JR, Lopez-Martin S, Caballero Segura FJ, Gallo González V, Recio Moreno B, et al. All roads lead to Rome: results of non-invasive respiratory therapies applied in a tertiary-care hospital without an intermediate care unit during the COVID-19 pandemic. *Open Respir Arch*. 2021;3:100081.
20. Segrelles-Calvo G, de Granda-Orive JI, López-Padilla D, Zamora-García E. Therapeutic limitation in elderly patients: reflections regarding COVID19. *Arch Bronconeumol*. 2020;S0300-2896(20)30194-0.

21. Matabuena M, Padilla OHM, Gonzalez-Barcala FJ. Statistical and mathematical modeling in the Coronavirus epidemic: some considerations to minimize biases in the results. *Arch Bronconeumol.* 2020;56:601–2, <http://dx.doi.org/10.1016/j.arbres.2020.04.022>.
22. Qin L, Li X, Shi J, Yu M, Wang K, Tao Y, et al. Gendered effects on inflammation reaction and outcome of COVID-19 patients in Wuhan. *J Med Virol.* 2020;10, 1002/jmv.26137.
23. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet.* 2020;395:507–13.
24. Argenziano MG, Bruce SL, Slater CL, Tiao JR, Baldwin MR, Barr RG, et al. Characterization and clinical course of 1000 patients with coronavirus disease 2019 in New York: retrospective case series. *BMJ.* 2020;369, m1996.
25. Hu L, Chen S, Fu Y, Gao Z, Long H, Ren HW, et al. Risk factors associated with clinical outcomes in 323 COVID-19 hospitalized patients in Wuhan, China. *Clin Infect Dis.* 2020, ciaa539.
26. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'Donnell L, Chernyak Y, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ.* 2020;369, m1966.
27. Pan F, Yang L, Li Y, Liang B, Li L, Ye T, et al. Factors associated with death outcome in patients with severe coronavirus disease-19 (COVID-19): a case-control study. *Int J Med Sci.* 2020;17:1281–92.
28. Mo P, Xing Y, Xiao Y, Deng L, Zhao Q, Wang H, et al. Clinical characteristics of refractory COVID-19 pneumonia in Wuhan, China. *Clin Infect Dis.* 2020, ciaa270.
29. Chen T, Wu D, Chen H, Yan W, Yang D, Chen G, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *BMJ.* 2020;368, m1091.
30. Xie J, Hungerford D, Chen H, Abrams ST, Li S, Wang G, et al. Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19. *Lancet.* <https://doi.org/10.2139/ssrn.3562456> [preprint].
31. Liu Y, Du X, Chen J, Jin Y, Peng L, Wang HHX, et al. Neutrophil-to-lymphocyte ratio as an independent risk factor for mortality in hospitalized patients with COVID-19. *J Infect.* 2020;81:e6–12.
32. Du RH, Liang LR, Yang CQ, Wang W, Cao TZ, Li M, et al. Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: a prospective cohort study. *Eur Respir J.* 2020;55:2000524.
33. Szekely GJ, Rizzo ML. Partial distance correlation with methods for dissimilarities. *Ann Stat.* 2014;42:2382–412.
34. Shen C, Vogelstein JT. Decision forests induce characteristic kernels; 2018.
35. França G, Rizzo ML, Vogelstein JT. Kernel k-groups via Hartigan's method. *IEEE Trans Pattern Anal Mach Intell.* 2020, <http://dx.doi.org/10.1109/TPAMI.2020.2998120>. PP.
36. Chen J, Kelley WJ, Goldstein DR. Role of aging and the immune response to respiratory viral infections: potential implications for COVID-19. *J Immunol.* 2020;205:313–20.
37. Sun X, Wang T, Cai D, Hu Z, Chen J, Liao H, et al. Cytokine storm intervention in the early stages of COVID-19 pneumonia. *Cytokine Growth Factor Rev.* 2020;53:38–42.
38. Zhang C, Wu Z, Li JW, Zhao H, Wang GQ. The cytokine release syndrome (CRS) of severe COVID-19 and Interleukin-6 receptor (IL-6R) antagonist Tocilizumab may be the key to reduce the mortality. *Int J Antimicrob Agents.* 2020;55:105954.
39. Azkur AK, Akdis M, Azkur D, Sokolowska M, van de Veen W, Brüggemann MC, et al. Immune response to SARS-CoV-2 and mechanisms of immunopathological changes in COVID-19. *Allergy.* 2020;75:1564–70.
40. Fink AL, Engle K, Ursin RL, Tang WY, Klein SL. Biological sex affects vaccine efficacy and protection against influenza in mice. *Proc Natl Acad Sci USA.* 2018;115:12477–82.
41. Li Y, Jerkic M, Slutsky AS, Zhang H. Molecular mechanisms of sex bias differences in COVID-19 mortality. *Crit Care.* 2020;24:405.
42. Klein SL, Flanagan KL. Sex differences in immune responses. *Nat Rev Immunol.* 2016;16:626–38.
43. Takahashi T, Wong P, Ellingson M, Lucas C, Klein J, Israelow B, et al. Sex differences in immune responses to SARS-CoV-2 that underlie disease outcomes. *MedRxiv.* 2020, <http://dx.doi.org/10.1101/2020.06.06.20123414>. Preprint.
44. Kragholm K, Porsborg Andersen M, Gerds TA, Butt JH, Østergaard L, Polcwiartek C, et al. Association between male sex and outcomes of Coronavirus Disease 2019 (Covid-19) – a Danish nationwide, register-based study. *Clin Infect Dis.* 2020, ciaa924.
45. Tay MZ, Poh CM, Rénia L, MacAry PA, Ng LFP. The trinity of COVID-19: immunity, inflammation and intervention. *Nat Rev Immunol.* 2020;20:363–74.
46. Qin C, Zhou L, Hu Z, Zhang S, Yang S, Tao Y, et al. Dysregulation of immune response in patients with coronavirus 2019 (COVID-19) in Wuhan, China. *Clin Infect Dis.* 2020;71:762–8.
47. Tomar B, Anders HJ, Desai J, Mulay SR. Neutrophils and neutrophil extracellular traps drive necroinflammation in COVID-19. *Cells.* 2020;9:1383.
48. Sibila O, Molina-Molina M, Valenzuela C, Ríos Cortés A, Arbillaga-Etxarri A, Torralba García Y, et al. Spanish Society of Pulmonology and Thoracic Surgery (SEPAR) consensus for post-COVID-19 clinical follow-up. *Open Respir Arch.* 2020;2:278–83.
49. Villar-Alvarez F, Martinez-Garcia MA, Jimenez D, Fariñas-Guerrero F, Ortiz de Lejarazu-Leonardo R, López-Campos JL, et al. SEPAR recommendations for COVID-19 vaccination in patients with respiratory diseases. *Open Respir Arch.* 2021;3:100097.
50. Ortiz de Lejarazu-Leonardo R. Future of vaccination against SARS-CoV-2 infection. *Open Respir Arch.* 2021;3:100117.