

## Trabajo Fin de Grado

Predicción de mortalidad en UCI de pacientes con COVID-19: un modelo predictivo supervisado con redes neuronales artificiales (RNA)

Prediction of mortality in ICU of patients with COVID-19: a predictive model supervised with artificial neural networks (ANN)

Autor

Fernando Benito Mariscal

Directores

Dr. Juan José Araiz Burdio

Dra. Begoña Zalba Etayo

Facultad de Medicina

2022



## RESUMEN

**Introducción.** Se han descrito muchos factores de riesgo de la COVID-19, pero no disponemos de un indicador de mortalidad en UCI para individualizar el pronóstico del paciente y establecer medidas de tratamiento más precoces y eficaces. Las Redes Neuronales Artificiales (RNA) son una forma de análisis multivariante alternativo.

**Objetivo.** Diseñar un indicador mediante RNA para predecir la mortalidad del paciente crítico ingresado en UCI por COVID-19, a partir de las variables presentes en las primeras 24 horas de ingreso en UCI, analizar las variables relacionadas y evaluar su capacidad predictiva.

**Material y Método.** Se ha realizado un estudio observacional, retrospectivo en el que se recogieron las variables presentes en las primeras 24 horas tras el ingreso en el Servicio de Medicina Intensiva del HCU "Lozano Blesa". Tras realizar un análisis estadístico se ha obtenido una Red MLP, siendo el indicador de mortalidad el valor de la neurona de la capa de salida que predice exitus. El software informático calcula la importancia de las variables (IV) incluidas y asigna cuales de ellas eran altamente predictivas.

**Resultados.** El valor medio del Indicador en el grupo que sobrevivió fue de  $0.322 + 0.215$ , frente a  $0.592 + 0.213$  en el grupo de falleció;  $p < 0.001$ . Los factores más importantes en la predicción de mortalidad fueron: la edad (IV normalizada 100%), la LDH (IV normalizada 67,1%) y la ferritina (IV normalizada 51,2%). El área ROC del indicador descrito mostró una capacidad predictiva aceptable: 0,808; 95% IC (0,736 y 0,879).

**Conclusiones.** Es posible crear un Indicador de Mortalidad del paciente ingresado en UCI por COVID-19 utilizando la metodología del aprendizaje automático de las redes neuronales artificiales. Las variables más importantes relacionadas con su mortalidad son la edad, la ferritina y la LDH. El Indicador propuesto muestra una capacidad predictiva aceptable y podría mejorarse aún más si fuese creado a partir de una base de datos multicéntrica.

## ABSTRACT

**Introduction.** Many risk factors for COVID-19 have been described, but we do not have an ICU mortality indicator to individualise patient prognosis and establish earlier and more effective treatment measures. Artificial Neural Networks (ANN) are an alternative form of multivariate analysis.

**Objective.** To design an ANN indicator to predict mortality in critically ill patients admitted to the ICU for COVID-19, based on the variables present in the first 24 hours of admission to the ICU, to analyse the related variables and to evaluate their predictive capacity.

Material and Method. An observational, retrospective study was carried out in which the variables present in the first 24 hours after admission to the Intensive Care Unit of the HCU "Lozano Blesa" were collected. After statistical analysis, a MLP Network was obtained, with the mortality indicator being the value of the neuron in the output layer that predicts exitus. The computer software calculates the importance of the variables (IV) included and assigns which of them were highly predictive.

Results. The mean Indicator value in the surviving group was  $0.322 \pm 0.215$ , versus  $0.592 \pm 0.213$  in the deceased group;  $p < 0.001$ . The most important factors in predicting mortality were: age (IV normalised 100%), LDH (IV normalised 67.1%) and ferritin (IV normalised 51.2%). The ROC area of the described indicator showed an acceptable predictive ability: 0.808; 95% CI (0.736 and 0.879).

Conclusions. It is possible to create a Mortality Indicator for patients admitted to the ICU by COVID-19 using the automatic learning methodology of artificial neural networks. The most important variables related to mortality are age, ferritin and LDH. The proposed Indicator shows an acceptable predictive capacity and could be further improved if created from a multicentre database.

## PALABRAS CLAVE

SARS-CoV-2, COVID-19, UCI, factores de riesgo, redes neuronales artificiales, pronóstico, indicador de mortalidad.

# ÍNDICE

1.	INTRODUCCIÓN .....	6
1.1.	El SARS-CoV-2 y la COVID-19.....	6
1.2.	Factores de riesgo .....	9
1.3.	Redes Neuronales Artificiales .....	10
2.	HIPÓTESIS Y OBJETIVOS .....	13
2.1.	HIPÓTESIS.....	13
2.2.	OBJETIVOS.....	13
3.	MATERIAL Y MÉTODOS .....	14
3.1.	Diseño .....	14
3.2.	Población.....	14
3.3.	Variables.....	14
3.4.	Estadística convencional .....	16
3.5.	Red Neuronal. ....	16
3.5.1.	Diseño: preprocesamiento y parámetros de la red .....	16
3.5.2.	Indicador .....	17
3.5.3.	Métricas de Rendimiento: <i>IV</i> y <i>área ROC</i> .....	17
3.6.	Programas .....	17
4.	RESULTADOS.....	18
4.1.	Resultados generales .....	18
4.2.	Antecedentes médicos.....	18
4.3.	Scores y datos analíticos .....	19
4.4.	Tratamiento y soporte respiratorio .....	20
4.5.	Red neuronal.....	22
4.5.1.	Estructura de la red .....	22
4.5.2.	Indicador .....	23
4.5.3.	Importancia de las variables .....	23
4.5.4.	Área ROC.....	24
5.	DISCUSIÓN .....	25
6.	CONCLUSIONES.....	27
7.	CONFLICTOS DE INTERESES, CONSIDERACIONES y AUTORIZACIONES .....	28
8.	AGRADECIMIENTOS .....	29
9.	ANEXOS.....	30
9.1.	Anexo 1 .....	30
9.2.	Anexo 2 .....	31
10.	BIBLIOGRAFÍA .....	32

# 1. INTRODUCCIÓN

## 1.1. El SARS-CoV-2 y la COVID-19

El SARS-CoV-2 es un patógeno emergente responsable de la enfermedad conocida como COVID-19 (*Coronavirus Disease 2019*). Se identificó por primera vez en la ciudad china de Wuhan, en diciembre de 2019, tras la comunicación de un brote de casos de neumonía bilateral cuya causa no había sido identificada.

Este virus se clasificó dentro de la familia *coronaviridae*, donde se agrupan distintos patógenos responsables tanto de infecciones respiratorias leves en humanos inmunocompetentes, como de gastroenteritis en animales<sup>(1)</sup>. Hasta hace relativamente poco, se le había otorgado poca atención a los coronavirus por la sintomatología leve que provocaba en humanos. Sin embargo, esto cambia a lo largo del siglo XXI, con dos ejemplares de esta familia que surgen de reservorios animales para provocar epidemias con mayor morbilidad y mortalidad, provocando una alarma sanitaria. Estos virus altamente patógenos son conocidos como SARS-CoV en 2003 y MERS-CoV en 2012<sup>(1)</sup>.

El SARS-CoV-2 es el tercer coronavirus con potencial pandémico en el último siglo. A medida que se dio la posibilidad de secuenciar su ADN, se estudiaron sus similitudes con los coronavirus de otros murciélagos, así como ya se había hecho con el anterior virus SARS-CoV. La hipótesis más aceptada hoy en día sobre el origen de este patógeno es que un coronavirus de murciélago haya evolucionado a través de hospedadores intermedios, a partir de los cuales ha alcanzado la mutación necesaria para presentar potencial virulento en el ser humano<sup>(2)</sup>. Lo más probable es que el pangolín, debido al comercio ilícito para su uso en medicina tradicional china, surgiera como vector definitivo para su transmisión en el ser humano<sup>(3)</sup>.

Este brote ha generado una pandemia a nivel mundial, que ha afectado a toda la población, independientemente del nivel social, económico o cultural. El día 31 de diciembre de 2021 las muertes reportadas oficialmente por el virus SARS-CoV-2 son de 5.940.000 fallecidos, no obstante, el número estimado es hasta tres veces mayor, alcanzando los 18.200.000 fallecimientos. Así, la tasa mundial de exceso de mortalidad por COVID-19 fue de 120,3 muertes por cada 100.000 habitantes<sup>(4)</sup>.

Estructuralmente, el virus SARS-CoV-2 tiene un ARN monocatenario. Su genoma codifica cuatro proteínas: la proteína S (*spike*), la proteína E (*envelope*), la proteína M (*membrane*) y la proteína N (*nucleocapsid*). La proteína S es la responsable del tropismo del virus, que permite la fusión de la membrana invasora del virus con la célula huésped, liberando así el genoma viral<sup>(5)</sup>.

En la **figura 1**<sup>(6)</sup> se representa cómo la proteína S (*spike*) interactúa con el receptor ACE2. Esta afinidad en la interacción de ACE2 con los receptores virales explica parcialmente la mayor transmisibilidad de SARS-CoV-2 en comparación con SARS-CoV<sup>(7)</sup>.

Se pueden visualizar en la figura los mecanismos fisiopatológicos:

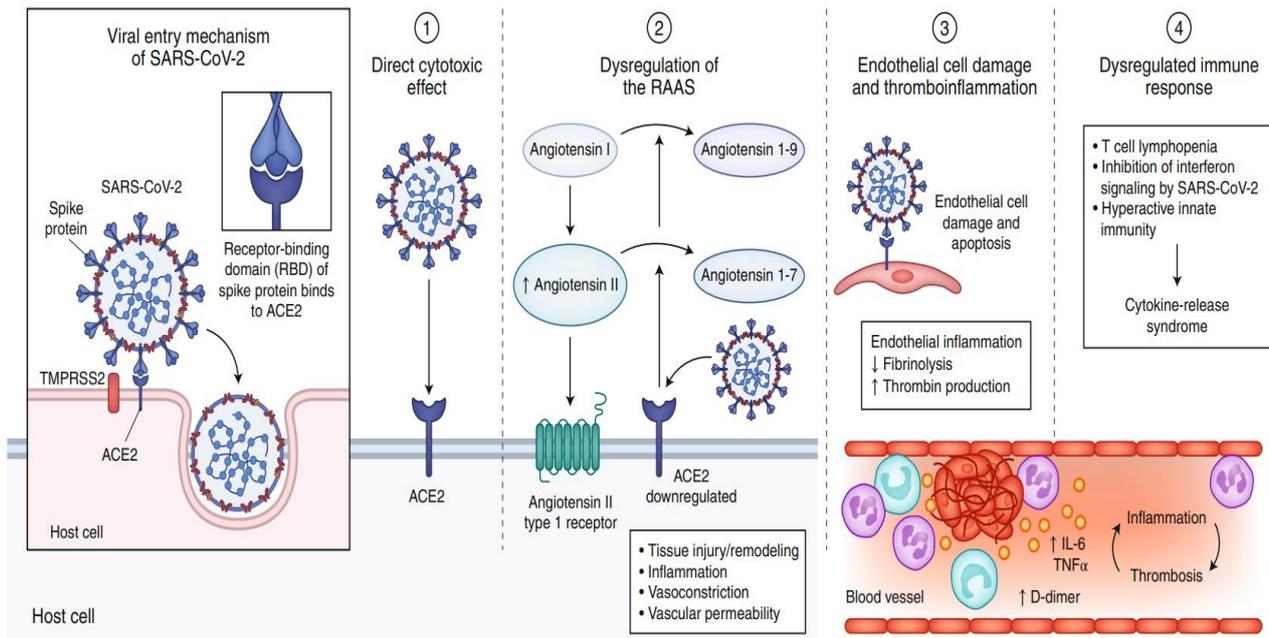


Figura 1. Fisiopatología del COVID19

1. El daño citotóxico directo causado por la carga viral tiene tropismo por las células que están expuestas en la vía aérea, incluido aquellas del tracto respiratorio agudo. En función de cómo avance la enfermedad en cada paciente, la replicación viral provocará neumonía o SDRA (Síndrome de Distrés Respiratorio Agudo) en el parénquima alveolar. Dicho tropismo no es independiente de este parénquima, si no que se encuentra en otros muchos tejidos corporales<sup>(8)</sup>.
2. La disregulación del SRAA (Sistema Renina Angiotensina Aldosterona) debido a un descenso del ACE2 provoca una disminución de la formación de angiotensina I y angiotensina II, que tienen propiedades vasodilatadoras, antifibróticas y antiproliferativas, afectando a la permeabilidad vascular.
3. El daño de las células endoteliales mediada por ACE2 y la inflamación consecuente genera una atmósfera protrombótica en el organismo. La producción excesiva de trombina y la disfunción microvascular crearán un estado protrombótico, como está demostrado en exámenes histopatológicos de pacientes con COVID-19, donde se observan endotelitis en los vasos sanguíneos<sup>(9)</sup>.

4. Por último, existe evidencia firme de que la sobreactivación de la respuesta inmune provoca la conocida tormenta de citoquinas, observada sobre todo en casos de infección grave. Dicha tormenta genera una reacción inmune descontrolada que provoca alteraciones en la coagulación, daño multiorgánico y en ocasiones la muerte del paciente.

El Comité Internacional de Taxonomía de Virus, le ha bautizado como SARS-CoV-2, en referencia al principal órgano donde genera patología: “*Severe Acute Respiratory Syndrome*” (SARS-CoV-2)<sup>(10)</sup>. Sin embargo, también genera complicaciones en otros sistemas (cardiovascular, neurológico, renal, etc.) provocando un fallo multiorgánico en los casos más graves mediante mecanismos que se explicarán a continuación.

Como refleja la **figura 2**<sup>(11)</sup>, el modelo patológico de la COVID-19 se puede subdividir en tres subconjuntos distintos pero unidos entre sí:

1. **Infección temprana:** tras la inoculación de la enfermedad y el periodo de incubación, se observa un periodo asociado a síntomas leves inespecíficos como fiebre y tos seca, además de malestar general. En la analítica sanguínea se puede encontrar linfopenia y neutrofilia.
2. **Afectación pulmonar:** se establece la enfermedad pulmonar localizada, existiendo multiplicación viral. En esta etapa el paciente puede desarrollar hipoxia ( $PaO_2/FiO_2 < 300$  mmHg), y a nivel radiológico se revelan infiltrados bilaterales y opacidades en vidrio esmerilado como signos de neumonía bilateral. Los análisis pueden mostrar elevación de las transaminasas y empeoramiento de la linfopenia.
3. **Hiperrespuesta inflamatoria sistémica:** una parte de aquellos pacientes con afectación grave pasará a esta etapa de la enfermedad, donde la afectación ya es también extrapulmonar: además del SDRA, puede entrar en estado de shock e insuficiencia cardíaca severa. Así, los marcadores de inflamación que se nombrarán a continuación variarán en función estado del paciente. En esta etapa crítica, el pronóstico y la recuperación empeoran. La línea negra de la gráfica corresponde a la gravedad de la enfermedad.

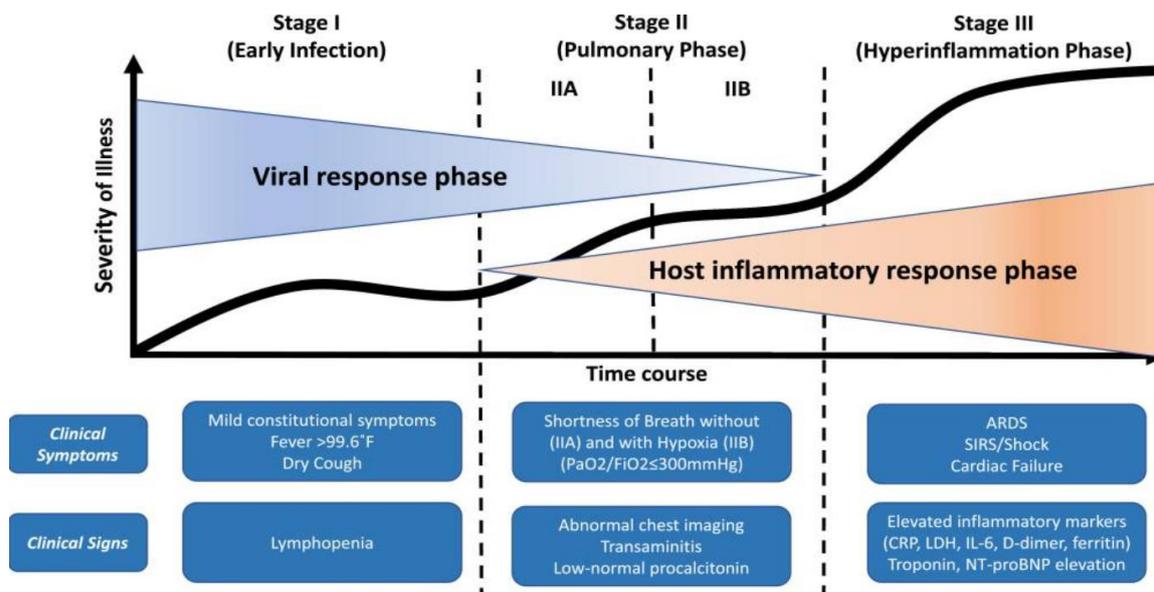


Figura 2 Clasificación de los estados de enfermedad de la COVID-19

## 1.2. Factores de riesgo

Se han descrito muchos factores de riesgo de la enfermedad que podríamos dividir en dos grandes grupos:

- Factores biológicos.** Es necesario destacar cómo en este fenómeno de hiperreactividad inmune provocado por el COVID-19 se da la elevación característica de algunos marcadores de inflamación. Como parámetros séricos se encuentran la proteína C reactiva (PCR), la ferritina, el dímero D, el fibrinógeno y el lactato deshidrogenasa (LDH).  
Existen otros parámetros sobre los que se han hecho estudios<sup>(12)</sup> con objetivo diagnóstico de infección severa, como son la procalcitonina (PCT) y la interleucina 6 (IL-6). Este último marcador también se está registrando con fines terapéuticos.  
Por otra parte, la variación del número total de linfocitos, neutrófilos y leucocitos se considera marcador de inmunidad celular alterada, por lo que sería interesante conocer su asociación pronóstica con la mortalidad de nuestros pacientes a estudio.
- Factores clínicos.** Por otro lado, es importante analizar si aquellos factores de riesgo que empeoran el pronóstico en varias enfermedades, también lo hacen en ésta. Estos son la hipertensión arterial (HTA), la diabetes mellitus (DM) o el tabaquismo, así como aquellos datos que estudian el perfil cardiovascular del paciente.

En la literatura actual, está en auge el desarrollo de modelos predictivos en relación con la muerte en hospitalización o ingreso en una Unidad de Cuidados Intensivos (UCI) por COVID-19. Esto es debido al auge de la pandemia en el que nos encontramos, pero también por la falta de predictores exactos de mortalidad. Por ejemplo, según algunos estudios el SOFA Score

(*Sequential Organ Failure Assessment Score*), que valora la situación global analizando el funcionamiento de los órganos corporales, no ha alcanzado la suficiente precisión como para predecir la gravedad de aquellos pacientes con neumonía por COVID-19 que requerían ventilación mecánica<sup>(13)</sup>.

Uno de los estudios que se ha convertido en referencia internacional dentro de este tema es un estudio multicéntrico español<sup>(14)</sup>, donde se determinan tres fenotipos (*clusters*) de enfermedad según las características clínicas de los pacientes. Pero, en este estudio no se determina claramente el *cluster* donde se encontraría un paciente concreto.

Parece importante pues, disponer cuanto antes de **un indicador de mortalidad en UCI**, para individualizar el pronóstico del paciente y establecer medidas de tratamiento más precoces y eficaces, sobre todo en el contexto epidemiológico en el que nos encontramos y sabiendo que esta es una cuestión todavía sin resolver.

### 1.3. Redes Neuronales Artificiales

En la **Figura 3** se representa una neurona artificial, que es en esencia una **función matemática**. Consiste en un procesador, donde la información -procedente del exterior o de otras neuronas-, a partir de un vector de entrada, es sometida a una serie de operaciones por las cuales calcula una única respuesta.

De la misma manera que una neurona biológica tiene un potencial de acción por el cual se dispara el estímulo nervioso, este sistema artificial debe alcanzar cierto umbral en la respuesta de salida para que la neurona envíe información a las otras neuronas conectadas<sup>(15),(16)</sup>. El esquema representa  $n$  entradas desde la neurona presináptica ( $x_1, x_2, x_n$ ), a los que la neurona le otorga distintos pesos ( $w_1, w_2, w_n$ ), generando un único valor  $y$  como salida. En la función de cada neurona, existen ciertos coeficientes, como son los pesos ( $w$ ) o los sesgos ( $\theta$ ).

En cada una de las neuronas, los pesos ( $w$ ) representan proporcionalmente la influencia de la entrada, aumentando o reduciendo el estado de activación de la neurona. El sesgo ( $\theta$ ) de la función tiene la capacidad de controlar lo predispuesta que está la neurona para llegar al umbral de respuesta, y es independiente de los pesos. Un valor bajo del sesgo hará que la neurona requiera un valor de entrada bajo para generar una respuesta de salida, mientras que un valor alto del sesgo provocará un aumento del umbral<sup>(15)</sup>.

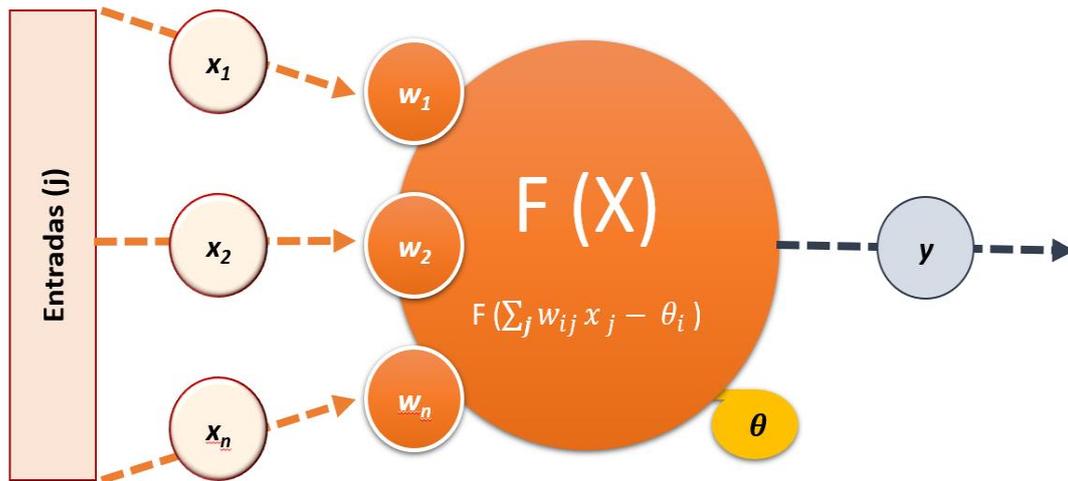


Figura 3. Esquema de una Neurona Artificial

Las Redes Neuronales Artificiales (RNA) son un novedoso método de aprendizaje automático. **Se considera una forma de análisis multivariante alternativo** que simula la estructura neuronal del sistema nervioso. Las neuronas artificiales están conectadas y transmiten señales entre sí: los valores de entrada pasan al soma de éstas, donde son procesados y generan otro valor que se transmite a otras neuronas.

En la **Figura 4** se observa cómo está constituida por una serie de capas (cada de entrada, capas ocultas y capa de salida) constituidas por un número determinado de neuronas artificiales, interconectadas entre sí. La capa de entrada es la que recibe las variables del exterior que queremos analizar. La información atraviesa un número indeterminado de capas centrales, y en la última capa se obtendrán unos valores de salida.

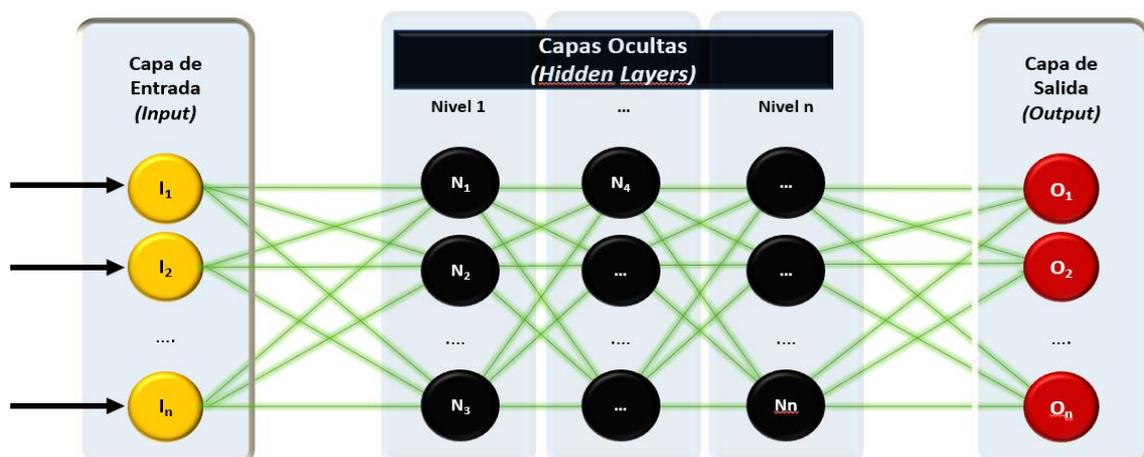


Figura 4. Esquema de una Red Neuronal Artificial

De la misma manera que el cerebro humano aprende de las experiencias vividas, las RNA funcionan gracias a la capacidad de identificar ciertos patrones, con el objetivo de predecir relaciones no reconocibles a simple vista. A diferencia de los sistemas convencionales, las redes neuronales tienen un sistema de aprendizaje automático.

Mediante diversos sistemas la red neuronal puede comparar información a través de las distintas capas, comparando los valores de las variables de entrada y salida de cada dato, con el objetivo de ajustar cada uno de los pesos. El aprendizaje de la RNA<sup>(17)</sup> consiste básicamente en la modificación de estos pesos en respuesta a la información que recibe, mediante un entrenamiento repetido, basándose en un mecanismo de probabilidades. Así, cambia las conexiones dentro de la asociación neuronal de tal manera que luego puede clasificar correctamente los datos desconocidos en función de las reglas aprendidas.

En función de las capas que presente una RNA, se modificará la complejidad de ésta, identificando patrones entre una cantidad inmensa de datos que son desapercibidos por el investigador. Una red profunda puede llegar a descubrir relaciones muy complejas, he aquí la gran utilidad de método de análisis multivariante, llegando a conclusiones que están fuera del alcance del cerebro humano y de la estadística convencional<sup>(18)</sup>.

Dentro del uso de las RNA en medicina intensivista, existe bibliografía acerca de la comparación de las nuevas redes con los sistemas de análisis tradicionales<sup>(19)</sup> Hasta ahora se han utilizado con objetivos como el de predecir la duración media de la estancia dentro de los cuidados intensivos<sup>(20)</sup>, pero con la nueva situación epidemiológica desde el inicio de la pandemia en 2020 se están creando modelos de predicción de mortalidad lo más precisos posibles que ayuden a desglosar las características más relevantes.

## 2. HIPÓTESIS Y OBJETIVOS

### 2.1. HIPÓTESIS

Debido a que no disponemos de un indicador individualizado que prediga la mortalidad en pacientes COVID-19 ingresados en UCI, se plantea este proyecto con los siguientes objetivos.

### 2.2. OBJETIVOS

Objetivo principal:

- **Diseñar un indicador** para predecir la mortalidad del paciente crítico ingresado en UCI por COVID-19, a partir de las variables presentes en las primeras 24 horas de ingreso en UCI.

Objetivos secundarios:

- **Analizar la importancia relativa** de las variables incluidas en el indicador.
- **Evaluar la capacidad predictiva** del indicador.

### 3. MATERIAL Y MÉTODOS

#### 3.1. Diseño

Consiste en un estudio observacional, descriptivo y retrospectivo en el que se incluyeron de manera consecutiva a todos los pacientes ingresados en UCI por la infección del virus SARS-CoV-2.

#### 3.2. Población

Se ha incluido en el estudio a los pacientes que ingresaron por infección SARS-CoV-2 en el Servicio de Medicina Intensiva (SMI) del Hospital Clínico Universitario “Lozano Blesa” (HCU) entre el 3 marzo de 2020 y el 16 de diciembre de 2021.

Se han recogido los valores de las variables analizadas presentes en las primeras 24 horas tras el ingreso en UCI.

- **Criterios de inclusión:** pacientes con diagnóstico de neumonía bilateral por SARS-CoV2, ingresados por un período superior a 24h en la SMI del HCU Lozano Blesa.
- **Criterios de exclusión:**
  - i. Pacientes ingresados con el diagnóstico de infección por SARS-CoV-2 con estancia inferior a 24 horas.
  - ii. Pérdida de seguimiento de los pacientes por falta de datos clínicos o analíticos.

#### 3.3. Variables

Se han recogido los siguientes datos:

- **Datos demográficos:**
  1. Género (Varón/Mujer)
  2. Edad (el punto de corte SI/NO es la cifra de 65 años)
- **Antecedentes:**
  3. Tabaquismo (activo, exfumador, nunca).
  4. HTA (SI/NO)
  5. Dislipemia (SI/NO)
  6. Diabetes mellitus (SI/NO)
  7. Síndrome metabólico (SI/NO)
  8. EPOC (SI/NO)
  9. Asma (SI/NO)
  10. Otras neumopatías -SAOS, enf. intersticiales, fibrosis pulmonar idiopática- (SI/NO)
  11. Tabaquismo (activo, exfumador, nunca)
  12. Cardiopatía isquémica (SI/NO).
  13. Hepatopatía crónica (SI/NO).
  14. Insuficiencia renal crónica (SI/NO).
  15. Accidente cerebrovascular (SI/NO).
  16. Tratamiento inmunomodulador (SI/NO)
  17. Inmunosupresión (SI/NO)
  18. Cáncer activo (SI/NO).
  19. Enfermedad hematológica (SI/NO)
  20. Embarazo actual (SI/NO)
  21. Enfermedad neurológica (SI/NO)

- **Scores:**
  - 22. Escala APACHE II (*Para la RNA el punto de corte SI/NO será un valor de 15*)
  - 23. Escala SOFA (*Para la RNA el punto de corte SI/NO será un valor de 5*)
- **Datos radiológicos:**
  - 24. Neumonía bilateral (SI/NO)
- **Soporte respiratorio:**
  - 25. Ventilación mecánica invasiva (VMI) (SI/NO)
  - 26. Oxigenación alto flujo (OAF) (SI/NO)
  - 27. Prono (SI/NO)
- **Soporte hemodinámico:**
  - 28. Inestabilidad hemodinámica (hipotensión arterial, shock hemodinámico) (SI/NO)
- **Datos analíticos:**
  - 29. Fracaso renal agudo (FRA) - comparación de la creatinina sérica con cifras pasadas - (SI/NO en base a la escala AKI – *Acute Kidney Injury* -)
  - 30. CK (creatina kinasa)
  - 31. LDH (lactato deshidrogenasa) (*Para la RNA el punto de corte SI/NO será un valor de 500 U/L*)
  - 32. Leucocitos totales
  - 33. Neutrófilos
  - 34. Linfocitos
  - 35. Dímero D (DD) (*Para la RNA el punto de corte SI/NO será un valor de 2000 ng/L*)
  - 36. Lactato
  - 37. Proteína C Reactiva (PCR)
  - 38. Procalcitonina (PCT)
  - 39. IL-6
  - 40. Ferritina (*Para la RNA el punto de corte SI/NO será un valor de 1800 ng/ml*)
  - 41. PaFi (relación entre la presión arterial de oxígeno -PaO<sub>2</sub>- y la fracción inspirada de oxígeno -FiO<sub>2</sub>-)
- **Tratamientos preingreso o a las 24 horas:**
  - 42. Corticoides (SI/NO)
  - 43. Antibióticos (SI/NO)
  - 44. Tocilizumab (SI/NO)
  - 45. Remdesivir (SI/NO)
  - 46. Kaletra (SI/NO)
  - 47. PHI (Plasma hiperinmune) (SI/NO)
  - 48. Inmunoglobulinas (SI/NO)
  - 49. Anakinra (Anti IL-1) (SI/NO)
  - 50. Hidroxicloroquina (SI/NO)
  - 51. Otros AC monoclonales (SI/NO)
  - 52. Otros tratamientos (SI/NO)
- **Evolución**
  - 53. Exitus (SI/NO)
  - 54. Estancia en SMI (días)

### 3.4. Estadística convencional

El análisis estadístico será descriptivo e inferencial:

- Para las variables cualitativas se calculará la distribución de frecuencias de las categorías; para variables cuantitativas se calcularán medidas de tendencia central (*media*) y de dispersión (*desviación estándar*).
- Se utilizarán las siguientes pruebas en el Contraste de Hipótesis: Chi-cuadrado de Pearson derivado de tablas de contingencia para variables cualitativas; “t” de Student o Test No Paramétricos para variables cuantitativas, en función de su distribución de normalidad.

Se considerará estadísticamente significativa una  $p < 0.05$ .

### 3.5. Red Neuronal.

#### 3.5.1. Diseño: preprocesamiento y parámetros de la red

**Preprocesamiento.** La Ingeniería de Datos consiste en el proceso de transformación de las variables independientes de los datos sin procesar a entradas (vectores de atributos). Entre las posibilidades que contempla están: la recodificación de variables y la creación de nuevas variables mediante el agrupamiento, operadores o combinación de atributos.

**Parámetros de la Red.** Usaremos una red del tipo **Perceptrón Multicapa (MLP)**. La partición utilizada será un 70% de los casos para el conjunto de entrenamiento y el 30% restante para la validación.

El MLP es una red neuronal formada por diferentes capas:

- **Capa de entrada.** Estará constituida por tantas neuronas como variables sean introducidas. Serán introducidas en la red aquellas variables que demostraron significación estadística: en el estudio multicéntrico anteriormente mencionado<sup>(14)</sup> y/o en nuestra serie de pacientes.

**La función de activación de las neuronas de la capa de entrada será la función de identidad ( $y=x$ ).**

- **Capas ocultas.** El número de capas ocultas y de neuronas es determinado automáticamente por el software.

**La función de activación de las neuronas de las capas ocultas será la tangente hiperbólica ( $y=tgh(x)$ ).**

- **Capa de salida.** Estará formada por 2 neuronas (NO Evento vs SI Evento) y usarán la función Softmax transformando los valores de entrada en un valor de salida entre 0 y 1 (0=supervivencia, 1=exitus), y ambos suman 1.

**La función de activación de las neuronas de la capa de salida será la función Softmax ( $y(ck)=\frac{\exp(ck)}{\sum_j \exp(cj)}$ ).**

### 3.5.2. Indicador

Se considera como valor del **Indicador de Mortalidad** el valor de la neurona de la capa de salida que predice exitus. Como la red descrita trabaja con la función softmax en la capa de salida este puede ser interpretado fácilmente como una probabilidad de exitus: *así un valor 0.025 indica que la probabilidad de exitus es del 2.5%*.

### 3.5.3. Métricas de Rendimiento: IV y área ROC

- **La Importancia de las Variables (IV)** corresponde a un análisis de sensibilidad que incorpora el propio software informático. A cada variable se le asigna un valor en el análisis de la red neuronal, siendo más importante cuanto más alto sea. Además, se calcula la **importancia normalizada (IV<sub>N</sub>)**. Se considerarán altamente predictivas aquellas variables con una IV normalizada mayor al 20%, parcialmente predictivas entre el 5% y el 20% y nada predictivas por debajo del 5%.
- **Área ROC**. Para validar el rendimiento de este indicador se ha utilizado el área ROC (*Receiver Operating Characteristic*), en el que el valor 1 significa una predicción perfecta, y un valor de 0,5 una predicción no superior al azar.

## 3.6. Programas

Para el análisis estadístico del conjunto de datos se utilizó el software SPSS versión 26.

## 4. RESULTADOS

### 4.1. Resultados generales

Los resultados globales de la población y las diferencias registradas entre el grupo que sobrevive y el que muere están resumidos en la **Tabla 1**. La edad media de los pacientes que fallecieron fue significativamente mayor (68.5 vs 59.9;  $p < 0.0001$ ).

	Global N = 221	Supervivencia n = 127	Exitus n = 94	P VALUE
<b>Demográficos:</b>				
Sexo: varón/mujer	141/80 (63.8/36.2)	82/45 (64.6/35.4)	59/35 (62.8/37.2)	0.7830
Edad	63.6 ± 11.6	59.9 ± 12.5	68.5 ± 8.0	< 0.0001
<b>Evolución:</b>				
Exitus	94 (42.7)			
Estancia en SMI	21.5 ± 19.7	20.4 ± 20.4	23.0 ± 18.8	0.3455

Tabla 1. Descripción de los grupos según supervivencia o éxitos

### 4.2. Antecedentes médicos

Se pueden visualizar en la **Tabla 2** los antecedentes médicos de la población registrada. El diagnóstico de EPOC fue el antecedente con diferencias más significativas entre ambos grupos (4% vs. 13,8%;  $p = 0,0083$ ).

	Global N = 221	Supervivencia n = 127	Exitus n = 94	P VALUE
HTA	129 (58.6)	66 (52.4)	63 (67.0)	0.0292
DLP (dislipemia)	105 (47.7)	56 (44.4)	49 (52.1)	0.2590
DM (diabetes mellitus)	66 (30.0)	29 (23.0)	37 (39.4)	0.0089
Síndrome metabólico	42 (19.1)	19 (15.1)	23 (24.5)	0.0796
EPOC	18 (8.2)	5 (4.0)	13 (13.8)	0.0083
Asma	11 (5.0)	5 (4.0)	6 (6.4)	0.4163
Otras neumopatías	18 (8.2)	10 (7.9)	8 (8.5)	0.8779
Tabaco	53 (24.1)	23 (18.3)	30 (32.0)	0.0191
Cardiopatía isquémica	19 (8.6)	10 (7.9)	9 (9.6)	0.6688
Hepatopatía crónica	9 (4.1)	6 (4.8)	3 (3.2)	0.5608
Nefropatía crónica	24 (10.9)	13 (10.3)	11 (11.7)	0.7445
ACV (Accidente cerebrovascular)	15 (6.8)	5 (4.0)	10 (10.6)	0.0522
Tratamiento inmunosupresor	12 (5.5)	7 (5.6)	5 (5.3)	0.9391
Inmunodeficiencias	6 (2.7)	2 (1.6)	4 (4.3)	0.2294

Cáncer activo	12 (5.5)	7 (5.6)	5 (5.3)	0.9391
Enfermedad hematológica	12 (5.5)	3 (2.4)	9 (9.6)	0.0201

Tabla 2. Antecedentes médicos

Quedan reflejados en la **Figura 5** los principales factores de riesgo cardiovascular en la población estudiada: 129 (58,6%) eran hipertensos, 105 (47,7%) estaban diagnosticados de dislipemia, 66 (30%) tenían diabetes mellitus y 42 (19%) presentaban un diagnóstico de síndrome metabólico.

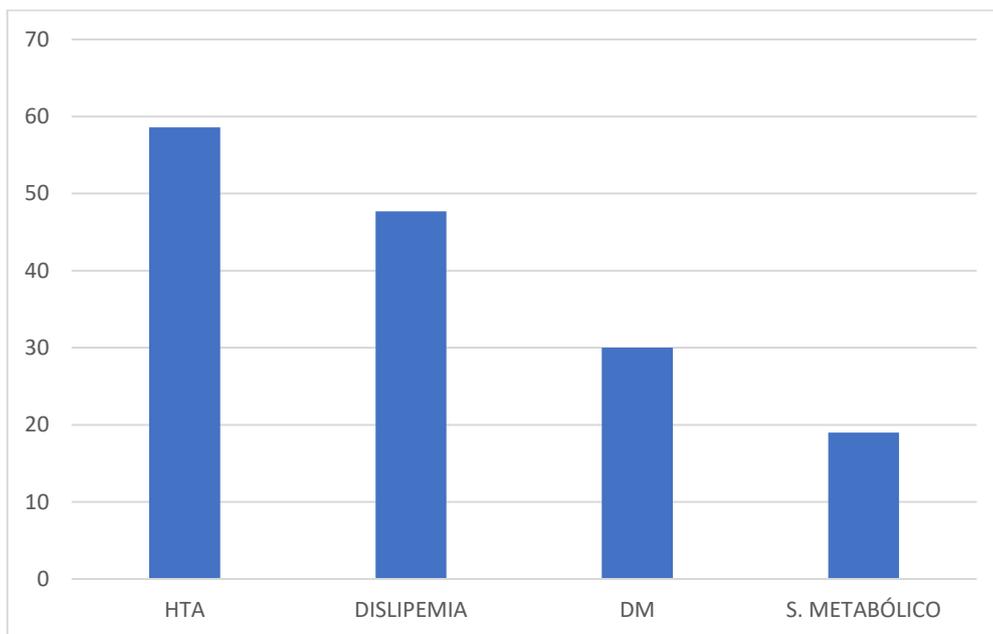


Figura 5. Factores de riesgo cardiovascular

Además del diagnóstico de EPOC, los antecedentes médicos que mostraron diferencias estadísticamente significativas en el grupo que sobrevivió y en el que no fueron: hipertensión arterial (52,4% vs. 67%;  $p=0,0292$ ), diabetes mellitus (23% vs. 39,4%;  $p=0,0089$ ), tabaquismo (18,3% vs. 32%;  $p=0,0191$ ), accidente cerebrovascular (4% vs. 10,6%;  $p=0,0522$ ) y enfermedad hematológica (2,4% vs. 9,6%;  $p=0,0201$ ).

#### 4.3. Scores y datos analíticos

En la **Tabla 3** se pueden estudiar los scores diagnósticos y los cambios analíticos de la infección severa por la COVID-19, sobre todo aquellos que se observan en la hiperrespuesta inmune sistémica ya mencionada. Tuvieron diferencias estadísticamente significativas el score APACHE II (11,3 vs. 14,2;  $p=0,0001$ ) y la ferritina (1351 vs. 2448;  $p=0,0283$ ).

	Global N = 221	Supervivencia n = 127	Exitus n = 94	P VALUE
<b>Scores:</b>				
APACHE II	12.5 ± 5.6	11.3 ± 5.1	14.2 ± 5.9	<b>0.0001</b>
SOFA	4.2 ± 2.1	4.0 ± 2.0	4.5 ± 2.2	0.0915
<b>Datos Analíticos:</b>				
FRA (Fracaso renal agudo)	15 (6.8)	8 (6.3)	7 (7.4)	0.7374
CK (Creatina kinasa)	291 ± 897	234 ± 639	375 ± 1185	0.2994
LDH (Lactato deshidrogenasa)	582 ± 985	595 ± 1265	563 ± 266	0.8268
Leucocitos	9731 ± 4458	9483 ± 4184	10063 ± 4806	0.3435
Neutrófilos	8753 ± 3974	8631 ± 3641	8917 ± 439	0.6013
Linfocitos	745 ± 1297	681 ± 532	830 ± 1886	0.4046
Dímero D	5154 ± 1344	5330 ± 16201	4901 ± 8004	0.8240
Lactato	2 ± 1	2 ± 1.2	2 ± 0.8	0.9929
PCR (Proteína C reactiva)	129 ± 101	134 ± 103	121 ± 98	0.3352
PCT (Procalcitonina)	3.7 ± 16.1	3.1 ± 15.1	4.5 ± 17.6	0.1345
IL-6 (Interleukina 6)	36.9 ± 17	37.2 ± 16.8	36.3 ± 18.4	0.7948
PaFi (PaO <sub>2</sub> /FiO <sub>2</sub> )	98.8 ± 44	102 ± 46	95 ± 42	0.2328
Ferritina	1829 ± 2968	1351 ± 917	2448 ± 4311	<b>0.0283</b>
<b>Datos Radiológicos:</b>				
Neumonía bilateral	216 (98.2)	123 (97.6)	93 (98.9)	0.4695

Tabla 3. Scores y datos analíticos

#### 4.4. Tratamiento y soporte respiratorio

Se muestra en la **Tabla 4** aquellas estadísticas relacionadas con el cuidado de los pacientes en la planta UCI. No se hallaron diferencias importantes con respecto al tratamiento farmacológico administrado ni con la terapia respiratoria.

	Global N = 221	Supervivencia n = 127	Exitus n = 94	P VALUE
<b>Soporte respiratorio:</b>				
OAF	170 (77.6)	97 (77.6)	73 (77.7)	0.9916
Prono al ingreso ó 24 h	101 (46.1)	52 (41.6)	49 (52.1)	0.1219
VMI	159 (72.6)	86 (68.8)	73 (77.7)	0.1457

<b>Soporte hemodinámico:</b>				
Inestabilidad hemodinámica	<b>44 (20.1)</b>	20 (16.0)	24 (25.5)	0.0814
<b>Tratamientos preingreso o a las 24 h:</b>				
Corticoides	<b>194 (88.6)</b>	112 (89.6)	82 (87.2)	0.5858
Antibióticos	<b>189 (86.3)</b>	106 (84.8)	83 (88.3)	0.4562
Tocilizumab	<b>68 (31.1)</b>	45 (36.0)	23 (24.5)	0.0679
Remdesivir	<b>59 (26.9)</b>	30 (24.0)	29 (30.6)	0.2580
Kaletra	<b>36 (16.4)</b>	25 (20.0)	11 (11.7)	0.1010
Inmunoglobulinas	<b>1 (0.5)</b>	0	1 (1.1)	0.2478
PHI (Plasma hiperinmune)	<b>46 (21.0)</b>	19 (15.2)	27 (28.7)	<b>0.0150</b>
Anakinra	<b>1 (0.5)</b>	0	1 (1.1)	0.2478
Hidroxiclороquina	<b>32 (14.6)</b>	23 (18.4)	9 (9.6)	0.0672
Otros anticuerpos monoclonales	<b>3 (1.4)</b>	0	3 (3.2)	<b>0.0443</b>
Otros tratamientos	<b>22 (10.0)</b>	9 (7.2)	13 (13.8)	0.1062

Tabla 4. Tratamiento y soporte

## 4.5. Red neuronal

### 4.5.1. Estructura de la red

Siguiendo la metodología descrita, finalmente se obtuvo una **Red 14:5:2**, representada en la **Figura 6**. En dicha gráfica la intensidad de las líneas marca la ponderación de cada entrada.

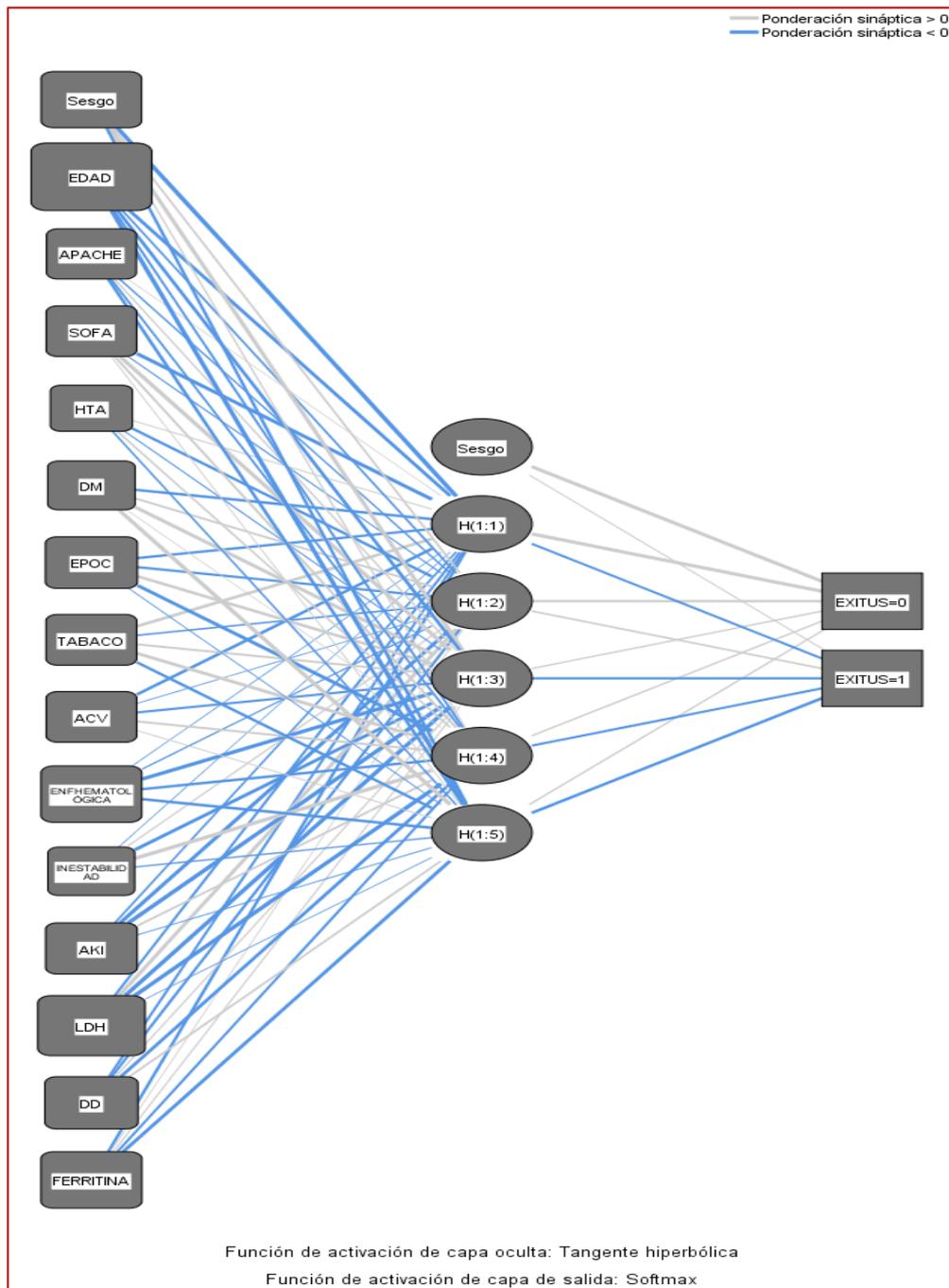


Figura 6. Red Neuronal

En el **Anexo 1** se adjuntan los pesos y los sesgos de cada una de las neuronas.

#### 4.5.2. Indicador

El valor medio del Indicador de Mortalidad en el grupo que sobrevivió fue de  $0.322 \pm 0.215$ , frente a  $0.592 \pm 0.213$  en el grupo de falleció;  $p < 0.001$  (**Figura 7**).

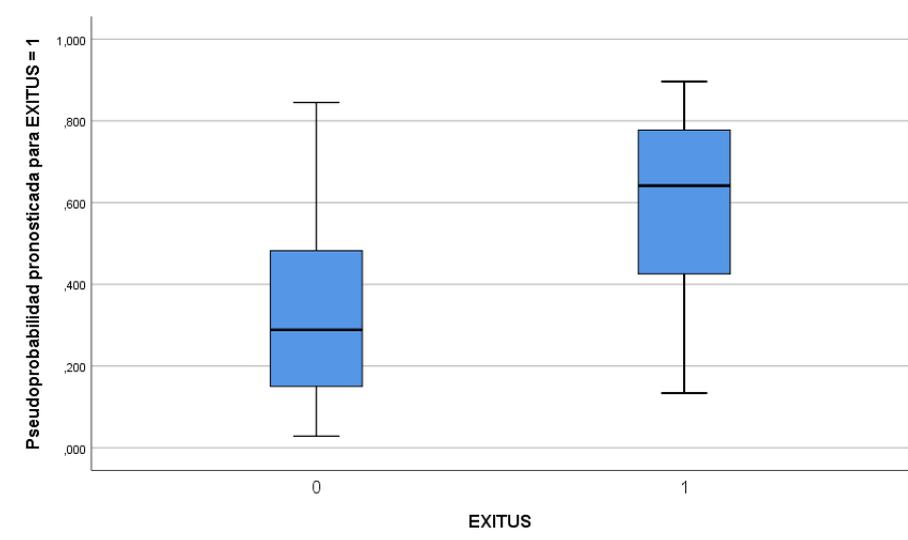


Figura 7. Indicador

#### 4.5.3. Importancia de las variables

La importancia de las variables en dicho indicador se refleja en la **Tabla 5**, así como gráficamente en la **Figura 8**.

Importancia de las variables independientes		
	Importancia	Importancia normalizada
EDAD	,202	100,0%
APACHE	,047	23,1%
SOFA	,051	25,1%
HTA	,009	4,5%
DM	,036	17,6%
EPOC	,058	28,8%
TABACO	,052	25,5%
ACV	,051	25,0%
ENF HEMATOLÓGICA	,103	51,1%
INESTABILIDAD	,032	16,0%
AKI	,056	27,8%
LDH	,136	67,1%
DD	,064	31,7%
FERRITINA	,104	51,2%

Tabla 5. Importancia de las variables

Como puede apreciarse, los factores más importantes con la predicción del evento fueron la **edad** (100% importancia normalizada), seguida de los datos analíticos de la **LDH** (67,1%) y la **ferritina** (51,2%). Otras variables que también se consideran altamente predictivas son el presentar una **enfermedad hematológica**, el **dímero D**, antecedentes de **EPOC**, **AKI** (fracaso renal agudo), **tabaquismo**, antecedentes de **ACV** y valores elevados en las escalas **SOFA** y **APACHE**.

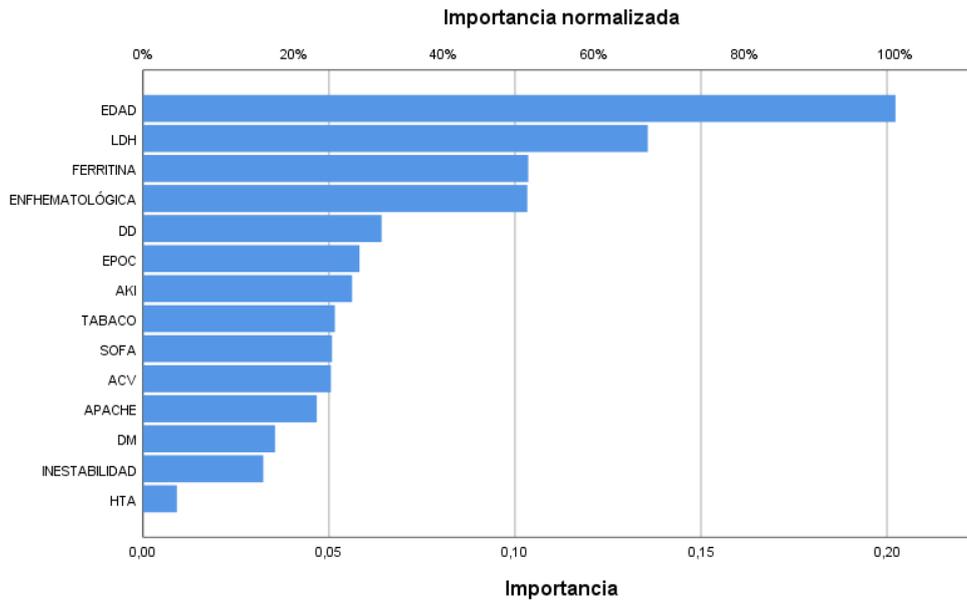


Figura 8. Importancia normalizada (%)

#### 4.5.4. Área ROC

El indicador de mortalidad ha obtenido un área ROC de 0,808 con un intervalo de confianza del 95% (IC 95%) comprendido entre 0,736 y 0,879, lo que significa en líneas generales un buen rendimiento (**Figura 9**).

Con un punto de corte del indicador de 0,13, este tendría una sensibilidad es del 100%. Por lo tanto, sobrevivirán todos los pacientes con una cifra inferior (punto rojo del área ROC).

Con un punto de corte del indicador de 0,85 este tendría una especificidad del 100%, asumiendo que todos aquellos pacientes cuya cifra del indicador sea superior, morirán (punto azul del área ROC).

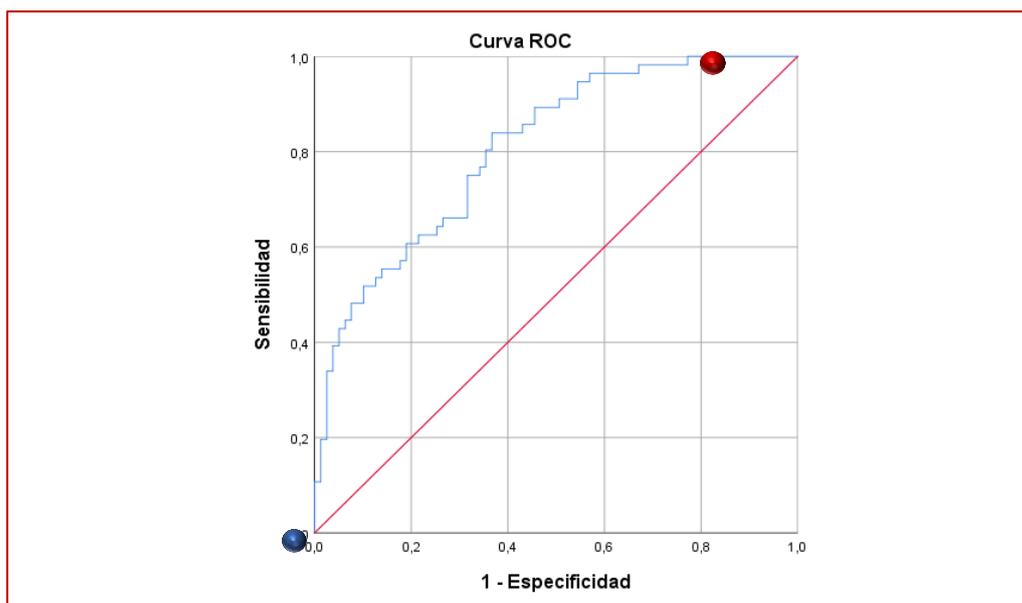


Figura 9. Área ROC del indicador

## 5. DISCUSIÓN

Respecto a la importancia de las variables predictoras de mortalidad, coincidimos con el estudio multicéntrico de referencia basado en la separación de los pacientes por clusters<sup>(14)</sup> en que las variables de más importancia son la edad, la LDH y la ferritina. El factor de riesgo más determinante ha sido la edad avanzada tanto en nuestro estudio (59,9 vs. 68,5 años;  $p < 0,0001$ ) como en el estudio mencionado, donde un 80% de los pacientes ingresados con más de 80 años terminaron falleciendo. La ferritina, analizada por ser un útil marcador de respuesta inflamatoria en las fases más graves de la enfermedad, resultó estadísticamente significativa en nuestro estudio (1351 vs. 2448;  $p = 0,028$ ) y también en el de referencia (1530 vs. 1900;  $p < 0,001$ ).

En contra de la opinión del trabajo de referencia mencionado<sup>(14)</sup>, en nuestro estudio la hipertensión es pobremente predictiva, pues se ha obtenido una  $IV_N$  del 4,5% al comparar las variables estudiadas. Sin embargo, en dicho estudio existía un porcentaje mayor de hipertensos en el grupo de pacientes que falleció que en el que no lo hizo (41,2% vs. 61,9%;  $p < 0,001$ ).

Respecto al análisis de algunos trastornos sistémicos como posibles factores pronósticos, estamos de acuerdo con otro estudio de referencia<sup>(21)</sup> en que la diabetes mellitus (23% vs. 39,4%;  $p = 0,008$ ; importancia normalizada 17,6%), el tabaquismo (18,3% vs. 32%;  $p = 0,019$ ;  $IV_N$  25,5%) y la EPOC (4% vs. 13,8%;  $p = 0,008$ ;  $IV_N$  28,8%) están directamente relacionados con la mortalidad.

También hemos comparado los resultados de este trabajo con un estudio de predicción de mortalidad mediante algoritmos de aprendizaje automático<sup>(22)</sup>. En el recuento total de células de defensa, nuestro estudio no encuentra diferencias significativas en el número de linfocitos (681 vs. 830;  $p = 0,404$ ) o de neutrófilos (8631 vs. 8917;  $p = 0,601$ ). Sin embargo, en dicho estudio el descenso de los linfocitos (1087 vs. 805;  $p < 0,001$ ) y el aumento del porcentaje total de neutrófilos (72% vs. 81,4%;  $p < 0,001$ ) fueron clasificados como datos de mal pronóstico.

Es de interés destacar que, si bien no existe bibliografía acerca de las enfermedades hematológicas como factor predictor de mal pronóstico, en nuestro estudio se ha obtenido significación estadística (2,4% vs. 9,6%;  $p = 0,020$ ). Sucede una algo parecido con aquellos pacientes que ingresaron en la UCI en situación de inestabilidad hemodinámica (16% vs. 25,5%;  $p = 0,0814$  y una  $IV_N$  de 16%). Estos resultados obtenidos podrían ser objeto de posteriores investigaciones.

Uno de los objetivos por los que se ha llevado a cabo este trabajo ha sido la falta de indicadores eficaces para el paciente enfermo por COVID-19, debido en parte a que los clásicos scores como el APACHE II o el SOFA no eran lo suficientemente específicos. A pesar de todo ello, tanto en nuestro trabajo (APACHE 11.3 vs. 14.2;  $p < 0.0001$  con una  $IV_N$  del 23,1% y SOFA con una  $IV_N$  DEL 25,1%) como en el artículo basado en la separación por clusters<sup>(14)</sup> (APACHE 13 vs. 17;  $p < 0.001$  y SOFA 5 vs. 7;  $p < 0.001$ ) una cifra superior a 15 en APACHE II o a 5 en SOFA se considera dato de mal pronóstico.

Como cualquier estudio, este también tiene sus limitaciones. En comparación con otros trabajos de referencia en los que se ha basado este proyecto, nuestro estudio ha podido acceder a un número de pacientes más limitado. Sin embargo, la gravedad que ha provocado la COVID-19 se traduce en un número significativo de eventos, lo que permite comparar dos grupos muy potentes de pacientes.

Por otra parte, esta investigación es unicéntrica, recogiendo los datos únicamente del SMI del HCU Lozano Blesa, en comparación con otros trabajos de referencia más heterogéneos que han realizado un estudio multicéntrico.

Otra limitación que destacar es la recogida de algunos de los datos clínicos. Es el caso de la PaFi (relación entre la presión arterial de  $O_2$  - $PaO_2$ - y la fracción inspirada de  $O_2$  - $FiO_2$ -), donde existen limitaciones para reconocer la hora exacta de la toma del dato y cuánto tiempo había pasado desde la intubación orotraqueal o incluso si aún no se había llevado a cabo este procedimiento. Sería interesante el desarrollo de otros estudios con el objetivo común de crear un indicador de mortalidad en SMI lo más exacto posible.

Por último, sería interesante incluir en futuras investigaciones otros datos multimodales que definan el estado del paciente, así como el tipo de vacuna administrada o el tiempo pasado desde la inoculación de dicha vacuna.

## 6. CONCLUSIONES

A la vista de este TFG y de sus resultados, podríamos extraer algunas **conclusiones**:

- **Primera. Es posible crear un Indicador de Mortalidad del paciente ingresado en UCI por COVID-19 utilizando la metodología del aprendizaje automático de las redes neuronales artificiales.**
- **Segunda. Las variables más importantes del Indicador propuesto y por tanto relacionados con la mortalidad de la COVID-19 son: *la edad, la ferritina y la LDH.***
- **Tercera. El Indicador propuesto muestra una capacidad predictiva aceptable y que podría mejorarse aún más si fuese creado a partir de base de datos multicéntricas.**

## 7. CONFLICTOS DE INTERESES, CONSIDERACIONES y AUTORIZACIONES

**Los autores declaran no tener ningún conflicto de interés en relación con el contenido del presente proyecto.**

**Los autores declaran no han recibido ningún soporte económico para la realización de este proyecto.**

El trabajo ha sido dirigido por la **Dra. Begoña Zalba Etayo** y el **Dr. Juan Jose Araiz Burdio, Jefa de Sección y Jefe de Servicio**, respectivamente, del Servicio de Medicina Intensiva del Hospital Universitario “Lozano Blesa”, **Profesores Asociados** de Medicina e **Investigadores del GIE de Críticos** (GIIS-075) del Instituto de Investigación Sanitaria Aragón.

El TFG expuesto cuenta con la aprobación del Comité de Ética de la Investigación de la Comunidad Autónoma de Aragón (CEICA). La solicitud se envió el **22/04/2022** y finalmente se resolvió, tras algunas correcciones, en la reunión del **20/05/2022 (Anexo 2)**.

## 8. AGRADECIMIENTOS

Mi más sincero agradecimiento al **Dr. Juan José Araiz Burdio** y a la **Dra. Begoña Zalba Etayo**, tutores de proyecto, por su continua disponibilidad tanto en el proceso de toma de datos como en la redacción de este trabajo.

Al **Dr. Íñigo Isern** y a la **Dra. María Montes**, médicos residentes del Servicio de Medicina Intensiva, por haber sido tan cercanos y haber solventado mis dudas en el inicio de este proyecto.

A todo el personal del Servicio de Medicina Intensiva del Hospital Universitario “Lozano Blesa” por su amabilidad y por haberme permitido hacer uso de sus instalaciones.

A la Facultad de Medicina de la Universidad de Zaragoza por contribuir a mi formación como médico.

A mis amigos, por haber sido un pilar fundamental a lo largo de esta carrera y por todos los momentos inolvidables que hemos vivido.

A mi familia, porque no habría llegado hasta aquí sin su apoyo incondicional y en especial a mi madre, por su gran corazón e interminable paciencia.

## 9. ANEXOS

### 9.1. Anexo 1

Predictor		Estimaciones de parámetro						
		Pronosticado					Capa de salida	
		Capa oculta 1					[EXITUS=0]	[EXITUS=1]
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)		
Capa de entrada	(Sesgo)	-,915	,248	,542	,144	-,342		
	EDAD	-,524	-,260	-,433	-,280	-,676		
	APACHE	,003	-,202	,153	-,268	-,400		
	SOFA	-,629	-,139	,634	,282	,042		
	HTA	,085	-,401	-,118	,238	-,223		
	DM	-,446	,310	,578	,010	,764		
	EPOC	-,394	-,370	,511	-,588	-,050		
	TABACO	,537	-,271	,319	,553	-,404		
	ACV	-,482	-,046	-,330	,321	,069		
	ENFHEMATOLÓGICA	-,033	-,186	-,159	-,409	-,488		
	INESTABILIDAD	-,021	,152	-,544	1,381	-,195		
	AKI	-,245	-,568	-,964	,281	-,085		
	LDH	-,316	,535	-,873	-,919	-,037		
	DD	-,311	-,328	,140	-,490	,276		
	FERRITINA	-,354	,031	,121	-,317	-,524		
Capa oculta 1	(Sesgo)						,705	,098
	H(1:1)						,787	-,325
	H(1:2)						,378	,273
	H(1:3)						,208	-,382
	H(1:4)						,211	-,334
	H(1:5)						,208	-,459

## 9.2. Anexo 2



### Informe Dictamen Favorable Trabajos académicos

C.I. PI22/254

18 de mayo de 2022

Dña. María González Hinjos, Secretaria del CEIC Aragón (CEICA)

#### CERTIFICA

**1º.** Que el CEIC Aragón (CEICA) en su reunión del día 18/05/2022, Acta Nº 10/2022 ha evaluado la propuesta del Trabajo:

**Título: Predicción de mortalidad en UCI de pacientes con COVID-19: un modelo predictivo supervisado con redes neuronales artificiales**

**Alumno: Fernando Benito Mariscal**

**Tutores: Juan José Araiz Burdio y Begoña Zalba Etayo**

**Versión protocolo: 22-04-2022**

**2º.** Considera que

- El proyecto se plantea siguiendo los requisitos de la Ley 14/2007, de 3 de julio, de Investigación Biomédica y los principios éticos aplicables.
- El Tutor/Director garantiza la confidencialidad de la información, la cesión de los datos seudonimizados y el adecuado tratamiento de los datos, en cumplimiento de la legislación vigente y la correcta utilización de los recursos materiales necesarios para su realización.

**3º.** Por lo que este CEIC emite **DICTAMEN FAVORABLE a la realización del proyecto.**

Lo que firmo en Zaragoza

GONZALEZ  
HINJOS MARIA -  
DNI 03857456B

Firmado digitalmente  
por GONZALEZ HINJOS  
MARIA - DNI 03857456B  
Fecha: 2022.05.20  
13:46:14 +02'00'

María González Hinjos  
Secretaria del CEIC Aragón (CEICA)

## 10. BIBLIOGRAFÍA

1. Paules CI, Marston HD, Fauci AS. Coronavirus Infections-More Than Just the Common Cold. *Jama*. 2020;323(8):707-8.
2. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. 2020;395(10224):565-74.
3. Cyranoski D. Did pangolins spread the China coronavirus to people?. 2020.
4. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020-21. *Lancet*. 2022;399(10334):1513-36.
5. Información microbiológica acerca de SARS-CoV-2. In: y CdCdA, Sanitarias E, editors. 2021. p. 10.
6. Gupta A, Madhavan MV, Sehgal K, Nair N, Mahajan S, Sehrawat TS, et al. Extrapulmonary manifestations of COVID-19. *Nat Med*. 2020;26(7):1017-32.
7. Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell*. 2020;181(4):894-904.e9.
8. Puelles VG, Lütgehetmann M, Lindenmeyer MT, Sperhake JP, Wong MN, Allweiss L, et al. Multiorgan and Renal Tropism of SARS-CoV-2. *N Engl J Med*. 2020;383(6):590-2.
9. Varga Z, Flammer AJ, Steiger P, Haberecker M, Andermatt R, Zinkernagel AS, et al. Endothelial cell infection and endotheliitis in COVID-19. *Lancet*. 2020;395(10234):1417-8.
10. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol*. 2020;5(4):536-44.
11. Siddiqi HK, Mehra MR. COVID-19 illness in native and immunosuppressed states: A clinical-therapeutic staging proposal. *J Heart Lung Transplant*. 2020;39(5):405-7.
12. Mudatsir M, Fajar JK, Wulandari L, Soegiarto G, Ilmawan M, Purnamasari Y, et al. Predictors of COVID-19 severity: a systematic review and meta-analysis. *F1000Res*. 2020;9:1107.
13. Raschke RA, Agarwal S, Rangan P, Heise CW, Curry SC. Discriminant Accuracy of the SOFA Score for Determining the Probable Mortality of Patients With COVID-19 Pneumonia Requiring Mechanical Ventilation. *Jama*. 2021;325(14):1469-70.
14. Rodríguez A, Ruiz-Botella M, Martín-Loeches I, Jimenez Herrera M, Solé-Violan J, Gómez J, et al. Deploying unsupervised clustering analysis to derive clinical phenotypes and risk factors associated with mortality risk in 2022 critically ill patients with COVID-19 in Spain. *Crit Care*. 2021;25(1):63.
15. Ivan Nunes da Silva DHS, Rogerio Andrade Flauzino, Luisa Helena Bartocci Liboni, Silas Franco dos Reis Alves. *Artificial Neural Networks*: Springer Cham; 2017.
16. Nielsen M. *Natural Networks and Deep Learning*. England: Determination Press (2015); Dec 2019.
17. Núñez Reiz A, Armengol de la Hoz MA, Sánchez García M. *Big Data Analysis and Machine Learning in Intensive Care Units*. *Med Intensiva (Engl Ed)*. 2019;43(7):416-26.

18. Gholipour K, Asghari-Jafarabadi M, Iezadi S, Jannati A, Keshavarz S. Modelling the prevalence of diabetes mellitus risk factors based on artificial neural network and multiple regression. *East Mediterr Health J.* 2018;24(8):770-7.
19. Traeger M, Eberhart A, Geldner G, Morin AM, Putzke C, Wulf H, et al. [Artificial neural networks. Theory and applications in anesthesia, intensive care and emergency medicine]. *Anaesthesist.* 2003;52(11):1055-61.
20. Tu JV, Guerriere MR. Use of a neural network as a predictive instrument for length of stay in the intensive care unit following cardiac surgery. *Proc Annu Symp Comput Appl Med Care.* 1992:666-72.
21. Yadaw AS, Li YC, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. *Lancet Digit Health.* 2020;2(10):e516-e25.
22. Aznar-Gimeno R, Esteban LM, Labata-Lezaun G, Del-Hoyo-Alonso R, Abadia-Gallego D, Paño-Pardo JR, et al. A Clinical Decision Web to Predict ICU Admission or Death for Patients Hospitalised with COVID-19 Using Machine Learning Algorithms. *Int J Environ Res Public Health.* 2021;18(16).