



Universidad de Valladolid



**PROGRAMA DE DOCTORADO EN
INVESTIGACIÓN EN CIENCIAS DE LA SALUD**

TESIS DOCTORAL:

**DETECCIÓN DE PREDICTORES DE
MORTALIDAD MEDIANTE
INTELIGENCIA ARTIFICIAL EN
SITUACIONES DE EMERGENCIA
SANITARIA: ESTUDIO DE CASO EN
PACIENTES CON COVID-19**

Presentada por Mónica Queipo Riera para optar
al grado de

Doctor/a por la Universidad de Valladolid

Dirigida por:

Dra. Julia Barbado Ajo

Dr. Jorge Mateo Sotos

Dra. Margarita González-Vallinas Garrachón

*A mis Pilares,
por sostenerme*

AGRADECIMIENTOS

A mis directores, por vuestra dedicación, sabiduría, ánimo, paciencia, confianza y apoyo. Gracias por ser la guía en la dirección y enriquecimiento de esta investigación.

A los pacientes, motor de este proyecto y últimos destinatarios del trabajo realizado.

A Javi, mi red y soporte en cada situación, en cada mal día y aún más en los buenos. Gracias por apoyar y celebrar mi crecimiento.

A mi madre, lectora y correctora incansable, seguidora primaria de mi progreso en cada paso. Tú, y ella contigo, habéis moldeado y creado lo que soy. Gracias por hacerme brillar y brillar conmigo.

A mi padre, por animarme a subir más, a llegar más alto. Esta vez me faltarán tus flores.

A mi familia y mis amigos, por acompañarme y escuchar con emoción y una sonrisa cada nuevo proyecto, locura y novedad. Especialmente a Nieves, por suavizar el camino.

A la Facultad de Medicina de la Universidad de Valladolid y también a la de Ciencias, el final del camino y el principio.

A TODOS, GRACIAS DE TODO CORAZÓN

Índices

Índice de contenidos

1. RESUMEN.....	9
2. INTRODUCCIÓN.....	13
2.1. COVID-19.....	15
2.1.1. Antecedentes históricos.....	15
2.1.2. Origen de la enfermedad.....	16
2.1.3. Epidemiología.....	17
2.1.4. Etiopatogenia.....	19
2.1.5. Manifestaciones clínicas.....	25
2.1.6. Diagnóstico.....	26
2.1.7. Tratamiento.....	27
2.1.8. COVID persistente o long COVID.....	32
2.1.9. La pandemia: impacto social y económico.....	33
2.1.10. Las olas de la pandemia en España.....	35
2.2. LA INMUNIDAD DE GRUPO.....	37
2.3. INNOVACIÓN TECNOLÓGICA Y AVANCES EN INVESTIGACIÓN.....	39
2.3.1. Técnicas de diagnóstico.....	40
2.3.2. Vacunas de nueva generación.....	41
2.3.3. Nuevos tratamientos.....	42
2.3.4. Investigación con organoides.....	44
2.3.5. Tecnología organ-on-a-chip.....	45
2.4. INTELIGENCIA ARTIFICIAL.....	46
3. HIPÓTESIS Y OBJETIVOS.....	49
4. PACIENTES Y MÉTODOS.....	53
4.1. PACIENTES.....	55
4.1.1. Recogida de datos.....	55
4.1.2. Variables de estudio.....	55
4.1.3. Aspectos éticos.....	58
4.2. METODOLOGÍA.....	59
4.2.1. Modelos de aprendizaje automático (machine learning).....	59
4.2.2. Configuración y entrenamiento de los modelos.....	69
4.2.3. Evaluación del rendimiento de los modelos.....	71

5.	RESULTADOS.....	75
5.1.	PERFIL Y EVOLUCIÓN DE LOS PACIENTES	77
5.1.1.	<i>Pacientes de la primera cohorte (1º ola)</i>	<i>77</i>
5.1.2.	<i>Pacientes de la segunda cohorte (2ª a 4ª olas).....</i>	<i>79</i>
5.1.3.	<i>Pacientes de la tercera cohorte (5ª y 6ª olas).....</i>	<i>82</i>
5.2.	IDENTIFICACIÓN DE PREDICTORES DE MORTALIDAD	84
5.2.1.	<i>Comparación y selección de modelos de ML</i>	<i>84</i>
5.2.2.	<i>Estudio de pacientes en UCI: biomarcadores.....</i>	<i>93</i>
6.	DISCUSIÓN.....	95
7.	CONCLUSIONES.....	109
8.	ANEXOS	113
8.1.	VALORES DE REFERENCIA EN PARÁMETROS DE LABORATORIO	115
8.2.	TABLAS DE RESULTADOS	116
8.2.1.	<i>Resultados relacionados con el estado previo de los pacientes</i>	<i>116</i>
8.2.2.	<i>Resultados relacionados con la clínica de los pacientes al ingreso</i>	<i>118</i>
8.2.3.	<i>Resultados relacionados con la estancia hospitalaria.</i>	<i>122</i>
9.	RENDIMIENTO CIENTÍFICO	124
10.	REFERENCIAS	127

Índice de figuras

FIGURA 1. ESQUEMA DE LA TEORÍA MÁS ACEPTADA SOBRE LA POSIBLE TRANSMISIÓN ZONÓTICA DEL SARS-CoV-2	17
FIGURA 2. ESTRUCTURA ESQUEMÁTICA DEL SARS-CoV-2 Y MECANISMO DE INTRODUCCIÓN A LA CÉLULA HUÉSPED	20
FIGURA 3. ESQUEMA DE LA ACTUACIÓN DE LAS CÉLULAS INMUNES ANTE UNA INFECCIÓN POR SARS-CoV-2.	23
FIGURA 4. FILOGENIA DE LAS VARIANTES DEL SARS-CoV-2 DESDE EL INICIO DE LA PANDEMIA HASTA ENERO DE 2023	25
FIGURA 5. PROPORCIÓN DE POBLACIÓN VACUNADA CONTRA EL SARS-CoV-2 EN EL MUNDO HASTA 2023 (INCLUIDO).	31
FIGURA 6. EVOLUCIÓN DE LA PANDEMIA POR SARS-CoV-2 JUNTO CON LOS HITOS Y PROYECTOS PRINCIPALES EN ESPAÑA	36
FIGURA 7. ILUSTRACIÓN DEL FUNCIONAMIENTO DE LA INMUNIDAD DE GRUPO	39
FIGURA 8. ESQUEMA DEL MECANISMO DE ENTRENAMIENTO Y TEST DE LOS MODELOS DE APRENDIZAJE AUTOMÁTICO	71
FIGURA 9. CURVAS ROC PARA LOS SEIS MODELOS DE ML EVALUADOS EN EL GRUPO DE PACIENTES DE LA 1ª OLA.....	85
FIGURA 10. GRÁFICOS DE RADAR PARA LAS VARIABLES Y MODELOS DE ML UTILIZADOS EN EL GRUPO DE PACIENTES DE LA 1ª OLA.....	86
FIGURA 11. PARÁMETROS MÁS RELEVANTES ASOCIADOS A LA MORTALIDAD DE LOS PACIENTES HOSPITALIZADOS POR COVID-19 DURANTE LA PRIMERA OLA.....	87
FIGURA 12. CURVAS ROC PARA LOS SEIS MODELOS DE ML EVALUADOS EN EL GRUPO DE PACIENTES DE LAS OLAS 2ª A 4ª.	88
FIGURA 13. GRÁFICOS DE RADAR PARA LAS VARIABLES Y MODELOS DE ML UTILIZADOS EN EL GRUPO DE PACIENTES DE LAS OLAS 2ª A 4ª	89
FIGURA 14. PARÁMETROS MÁS RELEVANTES ASOCIADOS A LA MORTALIDAD DE LOS PACIENTES HOSPITALIZADOS POR COVID-19 DURANTE LAS OLAS 2ª A 4ª.....	90
FIGURA 15. CURVAS ROC PARA LOS SEIS MODELOS DE ML EVALUADOS EN EL GRUPO DE PACIENTES DE LAS OLAS 5ª Y 6ª	91
FIGURA 16. GRÁFICOS DE RADAR PARA LAS VARIABLES Y MODELOS DE ML UTILIZADOS EN EL GRUPO DE PACIENTES DE LAS OLAS 5ª Y 6ª	92
FIGURA 17. PARÁMETROS MÁS RELEVANTES ASOCIADOS A LA MORTALIDAD DE LOS PACIENTES HOSPITALIZADOS POR COVID-19 DURANTE LAS OLAS 5ª Y 6ª.....	93

Índice de tablas

TABLA 1. CITOQUINAS ESTUDIADAS EN LOS PACIENTES INGRESADOS EN UCI Y LOS RANGOS DE REFERENCIA ANALÍTICOS Y LA SENSIBILIDAD ANALÍTICA DEL CORRESPONDIENTE KIT DE ELISA UTILIZADO.	57
TABLA 2. RESULTADOS FINALES DE BALANCED ACCURACY (%), PRECISIÓN, MCC, F1 SCORE Y AUC PARA CADA MODELO DE ML EVALUADO EN EL ANÁLISIS DEL GRUPO DE PACIENTES DE LA PRIMERA OLA.	84
TABLA 3. RESULTADOS FINALES DE SENSIBILIDAD, ESPECIFICIDAD, KAPPA Y DYI PARA CADA MODELO DE ML EVALUADO EN EL ANÁLISIS DEL GRUPO DE PACIENTES DE LA PRIMERA OLA.	85
TABLA 4. RESULTADOS FINALES DE BALANCED ACCURACY (%), PRECISIÓN, MCC, F1 SCORE Y AUC PARA CADA MODELO DE ML EVALUADO EN EL ANÁLISIS DEL GRUPO DE PACIENTES DE LAS OLAS 2ª A 4ª.	87
TABLA 5. RESULTADOS FINALES DE SENSIBILIDAD, ESPECIFICIDAD, KAPPA Y DYI PARA CADA MODELO DE ML EVALUADO EN EL ANÁLISIS DEL GRUPO DE PACIENTES DE LAS OLAS 2ª A 4ª.	87
TABLA 6. RESULTADOS FINALES DE BALANCED ACCURACY (%), PRECISIÓN, MCC, F1 SCORE Y AUC PARA CADA MODELO DE ML EVALUADO EN EL ANÁLISIS DEL GRUPO DE PACIENTES DE LAS OLAS 5ª Y 6ª.	90
TABLA 7. RESULTADOS FINALES DE SENSIBILIDAD, ESPECIFICIDAD, KAPPA Y DYI PARA CADA MODELO DE ML EVALUADO EN EL ANÁLISIS DEL GRUPO DE PACIENTES DE LAS OLAS 5ª Y 6ª.	90
TABLA 8. CONCENTRACIÓN DE CITOQUINAS EN PACIENTES DE LA PRIMERA OLA TRAS 15 DÍAS CON VMI EN LA UCI.	93
TABLA A1. CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS ASOCIADAS AL ESTADO PREVIO DE LOS PACIENTES. VARIABLES NUMÉRICAS.	116
TABLA A2. CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS ASOCIADAS AL ESTADO PREVIO DE LOS PACIENTES. VARIABLES CATEGÓRICAS.	117
TABLA A3. CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS RELACIONADAS CON LA SINTOMATOLOGÍA Y DIAGNÓSTICO CLÍNICO Y POR IMAGEN DE LOS PACIENTES CON COVID-19 AL INGRESO HOSPITALARIO.	118
TABLA A4. CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS DE LOS PACIENTES DE LA 1ª OLA AL INGRESO HOSPITALARIO. VARIABLES NUMÉRICAS.	119

TABLA A5. CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS DE LOS PACIENTES DE LAS OLAS 2ª A 4ª AL INGRESO HOSPITALARIO. VARIABLES NUMÉRICAS.	120
TABLA A6. CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS DE LOS PACIENTES DE LAS OLAS 5ª Y 6ª AL INGRESO HOSPITALARIO. VARIABLES NUMÉRICAS.	121
TABLA A7. CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS RELACIONADAS CON LA ESTANCIA HOSPITALARIA DE LOS PACIENTES	122
TABLA A8. DÍAS DE ESTANCIA HOSPITALARIA EN CADA GRUPO DE PACIENTES ANALIZADO. .	123

Abreviaturas

2DNMs	<i>Two-Dimensional Nanomaterials</i>
3CLpro	Proteasa 3C-like
3D	3 dimensiones
ACE2	Enzima convertidora de angiotensina 2
AEMPS	Agencia Española de Medicamentos y Productos Sanitarios
ALT/GPT	Alanina aminotransferasa/Glutamato-piruvato transaminasa
APACHE-II	<i>Acute physiology and chronic health disease classification system II</i>
ARN	Ácido Ribonucleico
ARNm	ARN mensajero
AST/GOT	Aspartato aminotransferasa/ Glutamato oxalacetato transaminasa
ATMPs	Advanced Therapies Medical Products
AUC	<i>Area under the curve</i>
BLDA	<i>Bayesian Linear Discriminant Analysis</i>
CARMEN	<i>Combinatorial Arrayed Reactions for Multiplexed Evaluation of Nucleic acids</i>
CAR-T	<i>Chimeric antigen receptor T</i>
CKI	<i>Cohen's Kappa Index</i>
CK-MB	Creatina quinasa-MB
CNAF	Cánula nasal de alto flujo
CNB	Centro Nacional de Biotecnología
COVID-19	<i>Coronavirus Disease 2019</i>
CRISPR	<i>Clustered Regularly Interspaced Short Palindromic Repeats</i>
CSIC	Consejo Superior de Investigaciones Científicas
CTL	<i>Cytotoxic T lymphocyte</i>
CURB-65	<i>Scales for the assessment of severity in community-acquired pneumonia</i>
CXCL	<i>C-X-C Motif Chemokine Ligand</i>
DETECTR	<i>DNA Endonuclease-Targeted CRISPR Trans Reporter</i>
DL	<i>Deep learning</i>
DT	<i>Decission Tree</i>
DYI	<i>Degenerated Youden's Index</i>
EBV	<i>Epstein-Bar virus</i>
ECMO	Extracorporeal membrane oxygenation
ELISA	<i>Enzyme-linked immunosorbent assay</i>
EPIC	Enfermedad pulmonar intersticial crónica

EPOC	Enfermedad pulmonar obstructiva crónica
FasL	Ligando Fas (CD95L)
FC	Frecuencia cardiaca
FDA	<i>Food and Drug Administration</i>
FR	Frecuencia respiratoria
GCS	<i>Glasgow coma scale</i>
G-CSF	<i>Granulocyte Colony-Stimulating Factor</i>
GGT	Gamma glutamil transferasa
GISAID	<i>Global Initiative on Sharing All Influenza Data</i>
GM-CSF	<i>Granulocyte-Macrophage Colony-Stimulating Factor</i>
GNB	<i>Gaussean Naïve Bayes</i>
Gradiente Aa	Gradiente alveolo-arterial
HCoV-229E	Human Coronavirus 229E
HCoV-HKU1	Human Coronavirus HKU1
HCoV-NL63	Human Coronavirus NL63
HCoV-OC43	Human Coronavirus OC43
HIT	<i>Herd immunity threshold</i>
HTA	Hipertensión arterial
INR.	Índice Internacional Normalizado
IA	Inteligencia artificial
IFNAR2	<i>Interferon-α/β receptor subunit 2</i>
IFN-I	Interferón I
IFN- γ	Interferón γ
IgA	Inmunoglobulina A
IL-X	Interleuquina-X
IP	<i>Interferon gamma-induced protein</i>
JAK-STAT	<i>Janus kinase - Signal transducer and Activator of Transmission</i>
KNN	<i>k-Nearest neighbors</i>
LbL	<i>Layer-by-layer</i>
LIF	<i>Leukemia inhibitory factor</i>
LPP	Lipopoliplex
MCC	<i>Mathew's correlation coeficient</i>
MCP	<i>Monocyte Chemoattractant Protein</i>
MERS	<i>Middle East Respiratory Syndrome</i>
MHC	<i>Major Histocompatibility Complex</i>
MIP	<i>Macrophage Inflammatory Protein</i>

ML	<i>Machine learning</i>
MSC	Mesenchymal Stem Cells
NF- κ B	<i>Nuclear factor κB</i>
NHC	β -D-N4-hidroxicitidina
NK	<i>Natural killer</i>
NRP-1	Neuropilina-1
NTD	<i>N-terminal domain</i>
OMS	Organización Mundial de la Salud
PAD	Presión arterial diastólica
PAS	Presión arterial sistólica
PASC	<i>Post-acute sequelae of SARS-CoV-2</i>
pCO ₂	Presión parcial de dióxido de carbono
PCR	Polymerase chain reaction
pDCs	Células dendríticas plasmacitoides
Proteína E	Proteína de la envoltura
Proteína M	Proteína de la membrana
Proteína N	Proteína de la nucleocápside
Proteína S	Proteína spike
PRRAR	Secuencia de aminoácidos: prolina (P), arginina (R), arginina (R), alanina (A) y arginina (R)
PRRs	<i>Pattern Recognition Receptor</i>
RBD	<i>Receptor binding domain</i>
RBF	<i>Radial basis function</i>
RF	<i>Random Forest</i>
RIG-I	<i>Retinoic acid-inducible gene I</i>
ROC	<i>Receiver operating characteristic</i>
SaO ₂	Saturación arterial de oxígeno
SARS	<i>Severe Actute Respiratory Syndrome</i>
SARS-CoV	<i>Severe Actute Respiratory Syndrome Coronavirus</i>
SARS-CoV-2	<i>Severe Actute Respiratory Syndrome Coronavirus 2</i>
SCVAM-1	<i>Soluble Vascular Cell Adhesion Molecule-1</i>
SDRA	Síndrome de distrés respiratorio agudo
SHERLOCK	<i>Specific High-sensitivity Enzymatic Reporter unLOCKing</i>
SOFA	<i>Sequential organ failure assessment</i>
SpO ₂	Saturación de oxígeno periférico
SVM	<i>Support Vector Machine</i>

TC	Tomografía computarizada
Tfh, T _{FH}	<i>T follicular helper cell</i>
TFGe	Tasa de Filtrado Glomerular estimada
TGF	<i>Transforming Growth Factor</i>
Th, T _H	<i>T-helper cell</i>
TLRs	<i>Toll like receptors</i>
TMPRSS2	Proteasa serina transmembrana 2
TNF- α	Factor de necrosis tumoral α
TP	Actividad de protrombina
Treg	<i>Regulatory T cell</i>
TTPA	Tiempo de tromboplastina parcial activada
TYK	<i>Tyrosine kinase 2</i>
UCI	Unidad de cuidados intensivos
VCM	Volumen corpuscular medio
VIH	Virus de la Inmunodeficiencia Humana
VLP	<i>Virus-like particles</i>
VMI	Ventilación mecánica invasiva
VMNI	Ventilación mecánica no invasiva
VOC	<i>Variant of concern</i>
VOI	<i>Variant of interest</i>
VPP	Valor predictivo positivo
VUM	<i>Variant under monitoring</i>
XGB	<i>Extreme Gradient Boosting</i>

1. Resumen

Desde los primeros casos identificados en diciembre de 2019, la pandemia de la COVID-19 ha supuesto una crisis sanitaria global que ha puesto en jaque a los sistemas sanitarios de todo el mundo. El cuadro clínico de la infección abarca desde casos asintomáticos hasta pacientes muy críticos, pudiendo provocar la muerte. Esta variabilidad se debe a factores que influyen en la capacidad del sistema inmune para responder al proceso inflamatorio y la “tormenta de citoquinas” que se puede producir tras la infección. Su alta transmisibilidad, las nuevas variantes emergentes y las características de cada población han evidenciado la necesidad de implementar protocolos de atención y diagnóstico adaptados y personalizados.

La inteligencia artificial y, en concreto, los algoritmos de aprendizaje automático o “*machine learning*” (ML), han adquirido importancia durante la pandemia como herramienta útil para el análisis estadístico de datos clínicos debido a su capacidad para identificar patrones en grandes conjuntos de datos rápidamente. Este estudio propone estos sistemas para la determinación de predictores de mortalidad en pacientes hospitalizados por COVID-19.

Se tomaron como muestra tres cohortes de pacientes hospitalizados por COVID-19 en el Hospital Universitario Río Hortega de Valladolid en diferentes periodos de la pandemia. Tras analizar sus características demográficas, clínicas y analíticas en el momento del ingreso y durante su estancia hospitalaria, se concluyó que era una muestra sólida para el estudio.

En la identificación de predictores de mortalidad de la COVID-19 destacó la edad avanzada en todas las cohortes, y otros parámetros como la escala CURB-65 y el recuento de linfocitos en algunos periodos. Se concluyó que el análisis de citoquinas no era una herramienta fiable para la identificación de biomarcadores de mortalidad en este grupo de pacientes. Se observaron cambios en los predictores de mortalidad a lo largo de la pandemia, evidenciando la evolución tanto de la enfermedad, como de la respuesta inmunológica gracias al aumento del porcentaje de población inmunizada de forma natural y a través de la vacunación.

Se seleccionaron e implementaron los modelos *EXtreme Gradient Boosting* (XGB) y *Random Forest* (RF) por su alto rendimiento y precisión, superando en, al menos, un 5 % al siguiente modelo, y mostrando un área bajo la curva muy elevado ($AUC > 90 \%$) en cada caso.

El presente estudio expone cómo los sistemas de ML son una herramienta útil para la predicción de mortalidad por COVID-19 utilizando una muestra pequeña de pacientes, evidenciando su efectividad como herramienta analítica para futuras emergencias sanitarias similares. Dada la heterogeneidad de las poblaciones y de este tipo de enfermedades, estas técnicas podrían llegar a incorporarse a la práctica clínica habitual como herramienta de apoyo al diagnóstico, permitiendo una atención médica personalizada y adaptada, mejorando así la calidad de los sistemas sanitarios y la optimización de recursos.

2. Introducción

2.1. COVID-19

La enfermedad causada por el virus SARS-CoV-2 (*Severe Acute Respiratory Syndrome Coronavirus 2*), la COVID-19 (*Coronavirus Disease 2019*), emergió a finales de 2019 en Wuhan, China, y se propagó rápidamente a nivel global, desencadenando una pandemia sin precedentes en la historia moderna. Este virus pertenece a la familia de los coronavirus, conocidos por causar enfermedades respiratorias en humanos y animales. Las manifestaciones clínicas varían desde síntomas leves o incluso casos asintomáticos, hasta cuadros clínicos graves que pueden llevar a la hospitalización o la muerte. La respuesta mundial a esta pandemia ha incluido medidas de salud pública para mitigar la rápida expansión de la enfermedad, así como tratamientos sintomáticos y un rápido desarrollo de vacunas, con el objetivo de reducir el impacto de la pandemia en la salud y el bienestar de la población.

2.1.1. Antecedentes históricos

El SARS-CoV-2 fue identificado por primera vez el 12 de noviembre de 2019 en Wuhan (China) tras el diagnóstico de los primeros pacientes con insuficiencia respiratoria aguda debido a una neumonía atípica de causa desconocida. Inicialmente el virus se denominó 2019-nCoV (nuevo coronavirus de 2019) y también HCoV-19 (coronavirus humano 2019) (1).

Desde el primer coronavirus identificado en 1960 (2) se han descubierto un total de siete que afectan a humanos. Cuatro de ellos (HCoV-229E, HCoV-OC43, HCoV-NL63 y HCoV-HKU1) causan enfermedades respiratorias leves similares a un resfriado (3) y los tres restantes han tenido un impacto social considerable, siendo el SARS-CoV-2 el más devastador.

El SARS-CoV, causó un brote de Síndrome Respiratorio Agudo Severo (SARS, por sus siglas en inglés) entre los años 2002 y 2003. Surgió también en China y se propagó a nivel global, afectando a más de 8.000 personas. La transmisión de este coronavirus se producía, principalmente, a través de gotas respiratorias y contacto cercano. La tasa de mortalidad fue de un 10 %, afectando a 29 países antes de erradicarse. Las medidas de control de salud pública, como la identificación de contactos y la implantación de cuarentenas, permitieron controlar el brote (4).

El MERS-CoV, por su parte, causó el brote de Síndrome Respiratorio de Oriente Medio (MERS, por sus siglas en inglés) en 2012, año en el que fue identificado por primera vez en Arabia Saudita. Se han reportado unos 2.500 casos y 850 muertes, aproximadamente, y se ha expandido por 27 países. Se transmite también por gotas respiratorias y contacto cercano. La tasa de mortalidad de esta enfermedad fue más alta, aproximadamente de un 35 %, y la población más afectada fue aquella que tenía comorbilidades previas (5).

Aunque la naturaleza de estos brotes sea distinta a la de la COVID-19, se consideran antecedentes históricos importantes del mismo, ya que mostraron la capacidad de los coronavirus para causar brotes significativos. Además, las experiencias y lecciones aprendidas han sido fundamentales en la respuesta global a la pandemia de 2019.

2.1.2. Origen de la enfermedad

Un estudio llevado a cabo por la Organización Mundial de la Salud (OMS) en colaboración con diferentes países demostró que más del 32 % de los primeros casos reportados en Wuhan estuvieron asociados con el mercado de mariscos *Huanan Wholesale Seafood Market* en dicha ciudad (6). En este mercado se vendían animales vivos como murciélagos, zorros, visones o mapaches, todos ellos susceptibles a infecciones por sarbecovirus, subgénero de la familia de los betacoronavirus que incluye el SARS-CoV y el SARS-CoV-2.

La transmisión de un patógeno desde un huésped animal a un nuevo huésped de otra especie, provocando su infección, ocurre gracias al fenómeno de “salto evolutivo” o “*spillover*”. Este fenómeno se produce cuando un huésped susceptible se expone al patógeno o cuando en el patógeno aparece un cambio genómico clave tal que permite infectar al nuevo huésped (7).

Cuando esto ocurre, se puede producir una zoonosis, definida por la OMS como “*cualquier infección natural transmisible de animales vertebrados a humanos* (8)”. Un ejemplo claro de zoonosis es la enfermedad de la rabia, que se transmite a través de la mordida de un perro infectado o, en casos excepcionales, de otros animales salvajes (9). Para calificar una zoonosis como tal, deben cumplirse varios criterios. En primer lugar, la transmisión debe ocurrir desde un animal vertebrado a un humano de forma natural. Debe identificarse un reservorio animal en el que el patógeno pueda vivir y reproducirse sin causar enfermedad en el mismo. El patógeno debe ser capaz de infectar a los humanos y causar enfermedad, es decir, debe poder replicarse y provocar síntomas clínicos. Por último, debe haber evidencia epidemiológica que muestre patrones de infección que apoyen la teoría de la transmisión zoonótica, incluyendo brotes en humanos que coincidan con la exposición a los animales reservorios (10).

Ha habido numerosas hipótesis sobre la cadena de transmisión del SARS-CoV-2. El coronavirus detectado en humanos comparte una similitud genética del 87,6 % con el coronavirus detectado en murciélagos de la especie *Rhinolophus pusillus* en China en 2015 (ZXC21/ZC45) (11) y del 96,2 % con el coronavirus identificado en la especie *Rhinolophus affinis* en 2013 (RaTG13), también en China (12). Por otro lado, el coronavirus detectado en los pangolines (MP789), tiene una importante similitud genética con el SARS-CoV-2 en su envoltura (100 %), en las proteínas M y N de su nucleocápside (98,6 % y 97,8 %, respectivamente) y en su proteína spike (90,7 %). A raíz de estos descubrimientos, se ha teorizado que, o bien los pangolines pueden

haber sido el huésped reservorio intermedio, o bien el SARS-CoV-2 podría derivar de una recombinación entre los coronavirus detectados en el pangolín y en el murciélago (13), siendo la primera opción la más aceptada (14). Esta teoría queda esquematizada en la Figura 1.

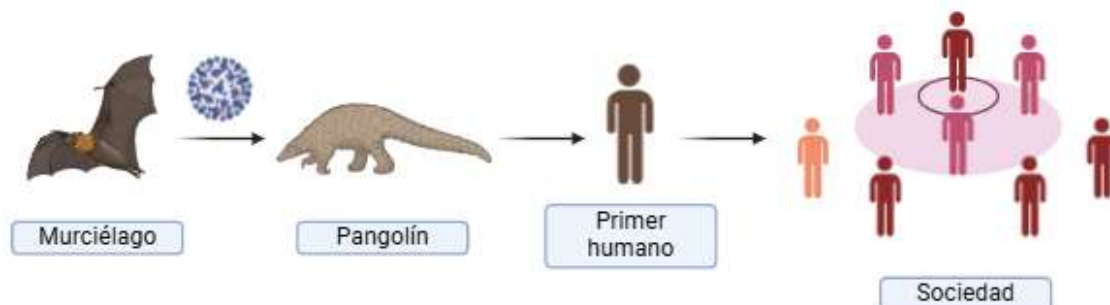


Figura 1. Esquema de la teoría más aceptada sobre la posible transmisión zoonótica del SARS-CoV-2. Imagen creada en BioRender.com

Cabe mencionar una de las teorías más controvertidas y rápidamente descartada por la OMS: la posibilidad de que el virus hubiera sido fabricado en el Instituto de Virología de Wuhan y, debido a un escape accidental, se hubiera propagado posteriormente al resto del mundo (15). Por una parte, los estudios genómicos de los primeros casos y su agrupación cerca del mercado de mariscos de Huanan en Wuhan sugieren un origen natural. Además, no se han encontrado evidencias que avalen la existencia de experimentos o propuestas para aumentar la transmisibilidad del virus en el laboratorio. Por último, los patrones de transmisión del virus son consistentes con brotes zoonóticos previos registrados históricamente (16).

2.1.3. Epidemiología

La COVID-19 se ha propagado rápidamente a nivel mundial, llevando a la OMS a declarar una pandemia el 11 de marzo de 2020 (17). A fecha de redacción de este trabajo (mayo de 2025), se han reportado 777.720.205 casos y más de 7 millones de muertes en todo el mundo (18).

La alta movilidad humana y la interconexión global facilitaron la expansión del virus, con brotes más significativos en países densamente poblados con altos niveles de tráfico internacional. Las tasas de infección y mortalidad fueron considerablemente diferentes entre regiones, probablemente influenciadas por factores como la respuesta de los gobiernos, la infraestructura y preparación sanitaria, la densidad de población y las características demográficas. Estados Unidos y Reino Unido, por ejemplo, tuvieron una incidencia mayor debido a la falta de coordinación, la respuesta inicial lenta y la alta movilidad de la población (19). Otros países como India o Brasil se vieron desbordados como consecuencia de la alta densidad poblacional, las desigualdades socioeconómicas y la falta de infraestructura sanitaria adecuada (20,21). Una situación similar se vivió en algunos países africanos, a los que se le añadió la alta prevalencia de VIH y tuberculosis, que complicaron enormemente la respuesta al virus (22,23). Zonas como China o Australia

llevaron a cabo estrictas medidas de confinamiento y control iniciales que fueron efectivas, pero los brotes posteriores y la fatiga pandémica ocasionaron otros desafíos (24–26). En Europa, Italia fue uno de los primeros países gravemente afectados dada la alta proporción de población por encima de los 65 años y la falta de preparación del sistema sanitario para este tipo de emergencias, que provocaron la saturación temprana del mismo (27,28).

El SARS-CoV-2 se transmite, principalmente, a través de gotas respiratorias o aerosoles cuando una persona infectada tose, estornuda o habla (29). También puede propagarse al tocar superficies contaminadas antes de tocarse la cara. Aunque se ha demostrado que esta no es la vía principal de contagio, se sabe que el virus puede permanecer viable en diferentes superficies durante horas o días, y por ello se mantuvo la importancia de la higiene de manos y la desinfección de superficies (30).

Las medidas de control más importantes que se adoptaron para disminuir la propagación del virus incluían el uso de mascarillas, el distanciamiento social, la cuarentena y la vacunación masiva. Las estrictas medidas de distanciamiento social y las cuarentenas demostraron ser altamente efectivas para reducir la transmisión del virus (31). Por su parte, los métodos de cribado basados en diagnóstico fueron eficaces en la contención de la propagación ya que identificaban a los viajeros infectados, permitiendo una detección temprana y un mejor manejo de los casos (32).

En el caso de las infecciones graves, se ha llevado a cabo un amplio debate acerca de ciertos factores de riesgo, tales como la edad avanzada, el género masculino y la presencia de comorbilidades, debido a su impacto en la morbilidad y mortalidad, así como en el pronóstico de los pacientes.

Las personas con una edad superior a los 65 años presentan un riesgo significativamente mayor de complicaciones graves, atribuible al deterioro del sistema inmunológico para combatir infecciones. Además, estas personas tienden a padecer con mayor frecuencia enfermedades crónicas (33). La hipertensión arterial se ha identificado como una de las comorbilidades más prevalentes en pacientes graves con el diagnóstico de COVID-19. Sin embargo, no se ha demostrado una correlación entre un mayor riesgo de mortalidad o gravedad por el nuevo patógeno y la hipertensión por sí misma, a menos que estuviera asociada a la edad avanzada (34). En el caso de, la diabetes se ha identificado como un factor de riesgo significativo, particularmente, de nuevo, en pacientes de edad avanzada (34,35). La obesidad, por su parte, está relacionada con inflamación crónica y podría exacerbar la respuesta inflamatoria al virus, ocasionando problemas respiratorios en algunos pacientes, lo que disminuye la capacidad del organismo para enfrentarse a la infección (36). Las enfermedades autoinmunes previas también desempeñan un papel importante debido a la disregulación del sistema inmune, especialmente si

estos pacientes están bajo el efecto de terapias inmunosupresoras o tienen otras patologías crónicas previas (37).

Por otra parte, se ha visto que, aunque la prevalencia de la enfermedad en hombres y mujeres es similar, el riesgo de mortalidad es mayor en hombres (38). Esto puede ser debido a diferencias biológicas entre sexos, ya que hormonas como los estrógenos, más abundantes en las mujeres, podrían tener un efecto protector contra infecciones virales. Las diferencias sociales relacionadas con comportamientos desfavorables para la salud, más frecuentes en hombres, también pueden contribuir a peores resultados clínicos (39).

Los niños y adolescentes sanos presentan, por norma general, sintomatología más leve y mejor pronóstico en comparación con los adultos, desarrollando una clínica grave en menos del 5 % de los casos (40,41).

Desde su inicio hasta 2023, año en el que la OMS declaró el fin de la pandemia (42), a nivel global, la prevalencia de la enfermedad, es decir, el número total de casos que se han dado en una población y periodo determinados, aumentó de 1,1 % a 9,8 %. Por el contrario, la tasa de mortalidad descendió de 2,4 % a 0,9 %. La tasa de infección acumulada fue mayor en los países de altos ingresos (35,3 %) que en los de bajos ingresos (0,6 %). Sin embargo, las cifras se invierten cuando se trata de la tasa de letalidad, siendo de 0,6 % en los países industrializados, frente a 1,6 % en las zonas en vías de desarrollo (43).

2.1.4. Etiopatogenia

El SARS-CoV-2 es un virus esférico pleomórfico de ARN de cadena simple y positiva, que consta de 29.903 bases que codifican 9.860 residuos de aminoácidos (44). Su estructura consta de una envoltura o cápside cuyo tamaño ronda entre los 50 nm y los 140 nm. Su forma recuerda a la de una corona que se forma a partir de peplómeros largos (20 nm) en forma de maza (45). Su estructura incluye varias proteínas clave.

La proteína Spike (S) facilita la entrada del virus en las células humanas al unirse a su receptor, la enzima convertidora de angiotensina 2 (ACE2), lo que facilita su replicación. Esta proteína se ensambla en trímeros que forman espículas en la superficie del virus y tiene dos subunidades principales. La subunidad S1 contiene el dominio de unión al receptor (RBD, por sus siglas en inglés), que es crucial para la unión al receptor ACE2. Esta subunidad también contiene el dominio N-terminal (NTD, por sus siglas en inglés), que contribuye a la estabilidad estructural de la proteína y favorece la interacción con receptores alternativos como la neuropilina-1 (NRP-1), que puede influir en el tropismo del virus. El NTD también favorece la evasión del sistema inmunitario al unirse a moléculas como la bilirrubina y la biliverdina, lo que impide la unión de este receptor a anticuerpos que activarían la respuesta inmunitaria. Por su

parte, la subunidad S2 facilita la fusión de la membrana viral con la membrana celular. Entre estas dos subunidades está el llamado sitio de hendidura polibásico (PRRAR), una secuencia de aminoácidos (prolina, arginina, arginina, alanina, arginina) que es escindida por la furina, una enzima que se encuentra en muchas células humanas. Este corte por furina activa la proteína S para su entrada en la célula huésped. Este sitio PRRAR no está presente en los betacoronavirus MERS-CoV y SARS-CoV, lo que explica que su infectividad sea menor que la del SARS-CoV-2, aunque sí lo tienen otros coronavirus humanos como el HCoV-HKU1 y el HCoV-OC43 (46). Después de su unión a ACE2, la proteasa serina transmembrana 2 (TMPRSS2, por sus siglas en inglés) activa de nuevo la proteína S escindiéndola en la subunidad S2 por proteólisis, promoviendo así la fusión de la membrana viral con la membrana de la célula huésped, tal como muestra la Figura 2.

La proteína de la envoltura (E) y la proteína de la membrana (M) son esenciales para la formación y estabilidad del virus. La proteína E, además de estabilizar la estructura del virus, es crucial para la liberación de nuevas partículas virales, ya que interactúa con la membrana celular del huésped formando vesículas que ayudan en la difusión de nuevos viriones. La proteína M es la más abundante en la envoltura del virus, e interactúa con otras proteínas virales como la S o la E, facilitando el ensamblaje de nuevas partículas virales. Además, esta proteína M interactúa con la proteína de la nucleocápside (N), responsable de la formación del complejo ribonucleoproteico, ayudando a organizar y proteger el ARN dentro del virión (47).

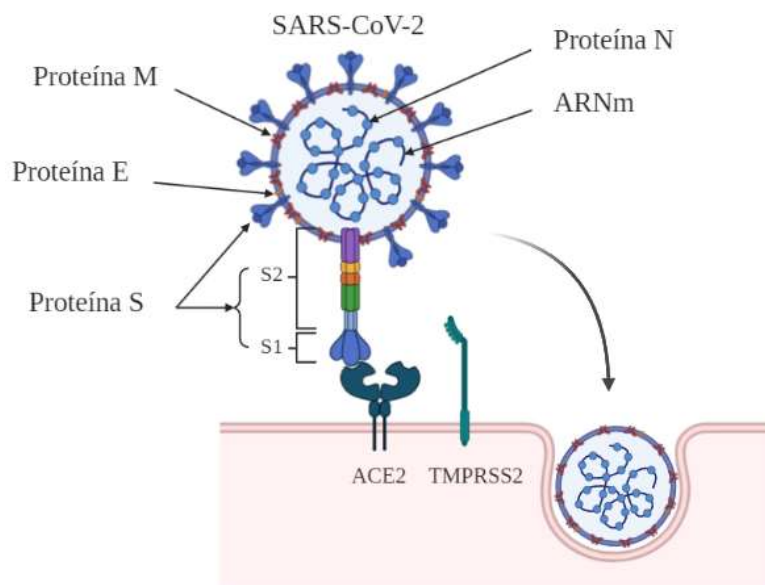


Figura 2. Estructura esquemática del SARS-CoV-2 y mecanismo de introducción a la célula huésped. Imagen creada en BioRender.com

Una vez el virus ha entrado en la célula del huésped, libera su ARN en su citoplasma y utiliza la maquinaria celular para replicarlo. Las nuevas partículas virales se ensamblan en el retículo endoplásmico y el aparato de Golgi, y luego se liberan para infectar otras células (48,49).

Ante esto, la célula huésped detecta el patógeno y activa una respuesta inmune apropiada. La respuesta inmune al SARS-CoV-2 involucra tanto a la inmunidad innata como a la adaptativa.

La respuesta inmune innata es la primera línea de defensa contra el SARS-CoV-2. Cuando el virus se replica en células epiteliales del tracto respiratorio alto y los pulmones es detectado por las células dendríticas plasmacitoides (pDCs) y los receptores de reconocimiento de patrones (PRRs), como los receptores tipo Toll endosomales (TLRs) o los receptores tipo RIG-I (gen inducible por ácido retinoico I). Esto da lugar a la activación de la respuesta por la vía del interferón I (IFN-I) y del factor nuclear κ B (NF- κ B), resultando en una liberación de citoquinas pro-inflamatorias. Estas citoquinas ayudan a reclutar células inmunitarias al sitio de la infección, sin embargo, en casos graves de COVID-19, esta respuesta puede volverse descontrolada, provocando una "tormenta de citoquinas" y, como consecuencia, un importante daño tisular.

Esta tormenta de citoquinas consiste en la liberación excesiva de citoquinas proinflamatorias como la interleuquina-6 (IL-6), la interleuquina-1- β (IL-1 β), el interferón- γ (IFN- γ) o el factor de necrosis tumoral- α (TNF- α). La IL-6, además, es un potente activador del sistema del complemento, crucial en la respuesta innata, que contribuye a la opsonización y lisis de las partículas virales y a la activación de otras células inmunitarias. No obstante, una activación excesiva del complemento puede llevar a una inflamación exacerbada y daño orgánico, como ocurre en casos graves de COVID-19.

La comprensión de los mecanismos que involucran a estas citoquinas puede permitir la utilización de sus niveles como biomarcadores de factores de riesgo y/o biomarcadores pronósticos, con la posibilidad de crear guías terapéuticas que mejoren la atención sanitaria. De hecho, se ha demostrado la correlación entre la severidad de la COVID-19 y ciertos biomarcadores inflamatorios como la mencionada IL-6, el factor inhibidor de leucemia (LIF) (50), la familia IL-1 (IL-1b, IL-1RA, IL-18, IL-33), las citoquinas asociadas a las células T-helper-1 (Th1) (IL-2, IL-2R, IL-12, IL15, IL-27, TNF- α , sTNFR1), T-helper-2 (Th2) (IL-4, IL-5, IL-10, IL-13) y T-helper-17 (Th17) (IL-17A, IL-22), algunas quimiocinas (IL-8, MCP-1, MCP-3, MIP-1 α , MIP-1 β , MIP-3 β , CXCL9), proteínas y factores de crecimiento relacionadas con el IFN- γ , IP-10 (IL-7, TGF- β , GM-CSF, G-CSF) (51) o la molécula de adhesión celular vascular-1 (sVCAM-1) (52).

La respuesta inmune adaptativa se activa posteriormente y es más específica. Los linfocitos T CD4⁺ y CD8⁺ son esenciales en esta fase. Los linfocitos T CD4⁺ ayudan a coordinar la respuesta inmune activando y reclutando otros tipos de células inmunitarias. Los linfocitos T CD4⁺ foliculares (T_{FH}) activan a los linfocitos B fomentando la producción de anticuerpos. Las células T CD4⁺-helper (T_H) secretan citoquinas para mantener la respuesta inflamatoria. Los linfocitos T CD4⁺ reguladores (T_{reg}), por su parte, suprimen la activación y proliferación de otras

células inmunitarias para mantener la tolerancia inmunológica y prevenir respuestas autoinmunes, produciendo citoquinas inmunosupresoras como TGF- β o IL-10. En cuanto a los linfocitos T CD8+, conocidos como células T citotóxicas (CTL, por sus siglas en inglés), son responsables de destruir las células infectadas por el virus, mediante el reconocimiento de los péptidos virales presentados por el complejo mayor de histocompatibilidad clase I (MHC-I, por sus siglas en inglés) en la superficie de células propias infectadas. La eliminación de estas células ocurre gracias a la liberación y actuación conjunta de granzimas y perforinas que provocan, en última instancia, la apoptosis de la célula infectada.

Respecto a la producción de anticuerpos, los linfocitos B activados se diferencian en células plasmáticas que producen anticuerpos selectivos capaces de neutralizar al virus uniéndose específicamente a sus proteínas. De este modo bloquean su capacidad para unirse a los receptores celulares y facilitan su disponibilidad para la eliminación por otras células inmunitarias (53).

Tras la infección inicial, algunos linfocitos T y B se convierten en células de memoria. Estas células persisten en el organismo y proporcionan una respuesta rápida y robusta en caso de una reinfección. La memoria inmunológica es un componente clave de la respuesta adaptativa, es fundamental para la protección a largo plazo y es la base de la eficacia de las vacunas. La investigación ha demostrado que, incluso cuando los niveles de anticuerpos disminuyen con el tiempo, la presencia de células B y T de memoria puede proporcionar una protección duradera (54). El mencionado proceso de actuación y especialización de las células inmunes frente a la infección del SARS-CoV-2 queda reflejado esquemáticamente en la Figura 3.

Se cree que la variabilidad genética humana también juega un papel en la respuesta inmune al SARS-CoV-2, y que esto podría explicar aquellos casos más raros de pacientes graves sin factores de riesgo asociados, o lo contrario, pacientes con factores de riesgo que presentan una enfermedad leve. Se han identificado variantes genéticas que confieren protección contra la infección, y otras que, por el contrario, aumentan el riesgo de desarrollar la enfermedad grave. Dichas variantes se localizan en genes relacionados con interferones tipo I, anticuerpos contra interferones tipo I, genes que regulan la “tormenta de citoquinas”, o genes relacionados con la coagulación sanguínea. Sin embargo, los estudios genéticos de este tipo son complejos y están limitados por la heterogeneidad de las muestras, así como por la dificultad de establecer causalidad entre genotipo y fenotipo. Su validación requiere una metodología rigurosa que considere los mecanismos moleculares, celulares y tisulares de todo el organismo, lo cual obliga a realizar estudios exhaustivos a nivel bioquímico e inmunológico (55).

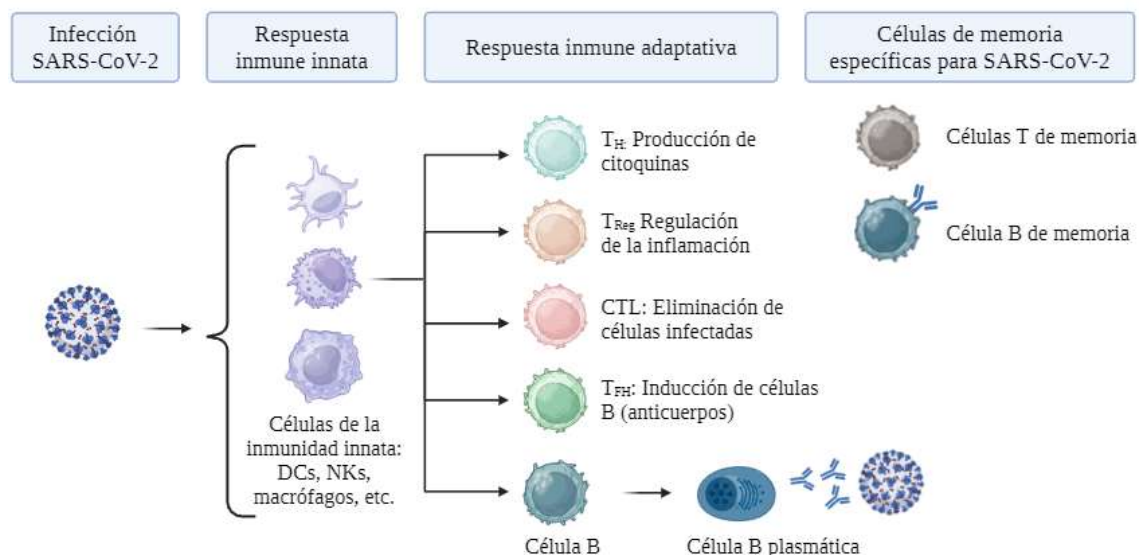


Figura 3. Esquema de la actuación de las células inmunes ante una infección por SARS-CoV-2. La infección lleva al desarrollo de una respuesta inmune innata (DCs, NKs, macrófagos, etc.), que conduce a la inducción de la respuesta adaptativa mediada por células T y B. Estas se diferencian para dar una respuesta específica para el virus. Finalmente se obtienen células B y T de memoria, que permanecen en el organismo. Imagen creada en BioRender.com

2.1.4.1. Variantes y su impacto epidemiológico

De forma similar a otros virus de ARN, el SARS-CoV-2 experimenta una evolución rápida gracias a su alta transmisibilidad. Las mutaciones que ocurren durante la replicación de su material genético en cada célula huésped pueden dar lugar a variantes que confieran al virus una ventaja evolutiva, como la mayor capacidad infectiva o una mejor evasión inmunológica. Estas mutaciones con ventajas selectivas se acumulan y tienden a propagarse más rápidamente en la población, generando una diversificación del virus en múltiples variantes.

La OMS utiliza una clasificación de estas variantes según su interés, véase “variantes de preocupación” (VOC, por sus siglas en inglés), “variantes de interés” (VOI) y “variantes bajo monitorización” (VUM).

La primera secuencia génica del SARS-CoV-2, Wuhan-Hu-1, fue obtenida el 24 de diciembre de 2019 (56). Sin embargo, es la secuencia WIV04/2019, recolectada el 30 de diciembre de 2019, la que se considera “original” o “de referencia” por el GISAID (Global Initiative on Sharing All Influenza Data) (57), ya que es más representativa de las secuencias del primer brote y cuenta con una alta calidad genómica (58). Tras esta, las calificadas como VOC por la OMS fueron las variantes Alfa, Beta, Gamma, Delta, Kappa y Omicron. La nomenclatura de letras griegas para VOCs y VOIs fue anunciada de forma oficial por la OMS el 31 de mayo de 2021, con el objetivo de evitar el estigma asociado al uso de nombres de países (59).

La variante Alfa (B.1.1.7), conocida inicialmente como variante británica, fue la primera VOC identificada tras la cepa inicial. Se identificó por primera vez en Reino Unido en septiembre de 2020. Esta variante tiene 17 mutaciones, de las cuales 8 afectan a la región de la proteína S

(delección $\Delta 144$, D1118H, N501Y, P681H, delección $\Delta 69-70$, T716I, S982A, A570D). En concreto, la mutación N501Y aumenta la afinidad de la proteína S por el receptor ACE2, mejorando la integración e infiltración viral. Se ha demostrado que esta variante aumentó la gravedad de la enfermedad (60,61).

La variante Beta (B.1.351) se identificó por primera vez en Sudáfrica en septiembre de 2020. Esta incluye nueve mutaciones, de las cuales 3 (N501Y, E484K y K417N) afectan al RBD, aumentando también la capacidad de la proteína S de unirse al receptor ACE2 (60,61).

La variante Gamma (P.1), una subvariante de B.1.1.28 identificada en diciembre de 2020, tiene 10 mutaciones solo en la proteína S, de las cuales 3 (E484K, L18F, and K417N) afectan al RBD, de igual manera que en la variante Beta (60,61).

Las variantes Delta (B.1.617.2) y Kappa (B.1.617.1) pertenecen al mismo linaje y ambas fueron identificadas por primera vez en diciembre de 2020 en la India. La primera tiene 10 mutaciones en la proteína S, y la variante Kappa fue investigada especialmente por tener alteraciones únicas que diferían de las previas (60–62).

La variante Omicron, identificada por primera vez en noviembre de 2021 en Sudáfrica, ha sido de gran interés dada su gran cantidad de mutaciones en diferentes partes de su genoma, como se muestra en la Figura 4, dando lugar a una nueva ola de casos. Se han identificado gran cantidad de subvariantes de Omicron, siendo la más comunes las correspondientes al inicio de su expansión BA.2, BA.1.1, y BA.1, acumulando 28 mutaciones en el caso de BA.2 (4 en la proteína S), y hasta 60 mutaciones en el caso de BA.1 (32 en la proteína S). Es por esta elevada cantidad de mutaciones genéticas por lo que esta variante ha sido más contagiosa que otras. De hecho, en abril de 2022 se identificaron en Sudáfrica las subvariantes BA.5 y BA.4, que evadían eficazmente la respuesta inmune gracias a las mutaciones L452R, F486V y Q493 en la proteína S. Estas subvariantes causaron una nueva ola de casos a nivel global, al igual que las subvariantes XBB.1.16 y XBB.1.9.2 (60,61).

Aunque el final de la pandemia se declarase en 2023, el virus sigue mutando y se siguen identificando nuevas variantes y subvariantes. La evolución del virus representa una amenaza para el control de la enfermedad, ya que las alteraciones virales podrían disminuir el alcance de la inmunidad de grupo ante variantes nuevas (63). Por ello, es necesario mantener la investigación del virus con el objetivo de adaptar tratamientos, técnicas de diagnóstico y estrategias de vacunación para las posibles nuevas variantes. No obstante, se espera que estas subvariantes deriven en una infección más leve que permita al virus mantenerse en circulación, como podría ser la subvariante de Omicron EG.5, que ha demostrado ser de bajo riesgo y su clasificación fue actualizada, pasando de VOI a VUM en agosto de 2023 (60,61).

Desde diciembre de 2023, ninguna variante ha sido declarada como VOC. A fecha de realización de este trabajo (mayo de 2025), solamente hay dos variantes VOI (Omicron BA.2.86 y Omicron KP.3) y dos variantes VUM (Omicron XEC y Omicron LP.8.1). El resto de variantes, previas a estas o de identificación reciente, han sido reclasificadas o desescaladas a variantes de menor preocupación, ya sea porque no están circulando, porque circulan sin ningún impacto epidemiológico o porque hay evidencia científica de que no están asociadas a nuevas características o propiedades de gravedad clínica o epidemiológica (61).

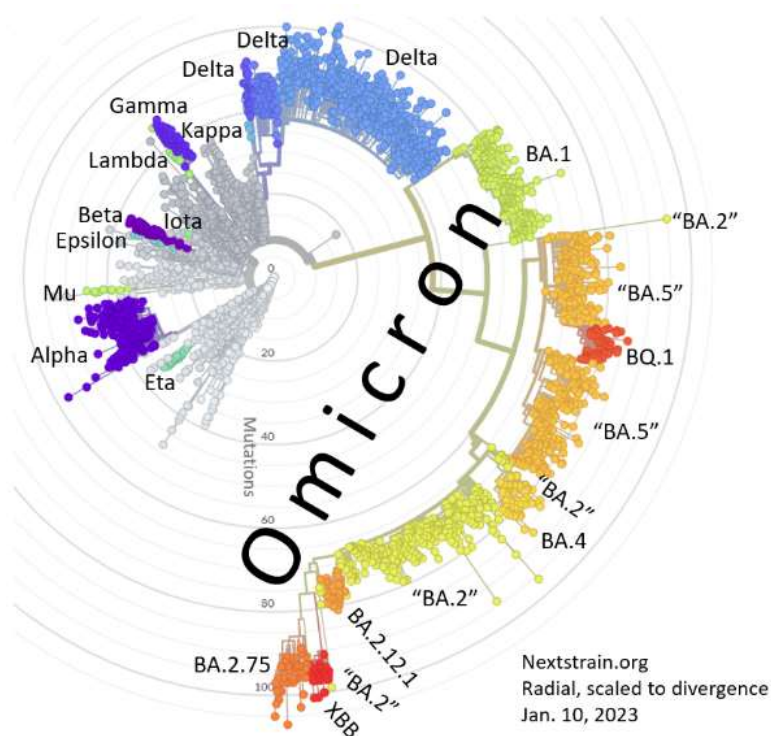


Figura 4. Filogenia de las variantes del SARS-CoV-2 desde el inicio de la pandemia hasta enero de 2023. Se representa como un formato radial escalado según divergencia (número de mutaciones), observándose la diversidad de las variantes de Omicron en comparación con otras ramas. Fuente: Nextstrain.org (64).

2.1.5. Manifestaciones clínicas

La COVID-19 puede variar desde una infección leve o incluso asintomática, hasta una enfermedad grave que puede ocasionar el ingreso en la unidad de cuidados intensivos (UCI).

La sintomatología ha ido variando a lo largo de la pandemia, siendo las manifestaciones más habituales la fiebre (78,3 %), la tos (60,2 %), la fatiga (45,6 %) y, en algunas variantes, la pérdida de gusto y olfato (38,7 %) (65). Síntomas menos comunes son los dolores musculares y de garganta, la congestión nasal, el dolor de cabeza, la diarrea, las náuseas y los vómitos. Por último, la sintomatología más grave engloba la dificultad respiratoria, el dolor o presión en el pecho y la pérdida de habla o movimiento. Las complicaciones graves pueden derivar en neumonía, síndrome de distrés respiratorio agudo (SDRA), sepsis y/o fallo multiorgánico. Los pronósticos más críticos son más comunes en personas con factores de riesgo.

La inflamación excesiva del organismo puede provocar otro tipo de complicaciones no respiratorias como problemas cardiovasculares (miocarditis, trombosis), pudiendo utilizar como biomarcador para su pronóstico un nivel elevado de dímero D y troponinas cardíacas (66). También se han dado patologías neurológicas que afectan al sistema nervioso (encefalitis, Síndrome de Guillain-Barré)(67). La COVID-19 puede afectar de forma severa a otros sistemas generando complicaciones renales (insuficiencia renal aguda) (68), gastrointestinales al alterar la microbiota intestinal (hepatitis, pancreatitis) (69), hematológicas (coagulopatías) (70) o dermatológicas (lesiones cutáneas) (71).

Se toma como “tiempo de incubación de la enfermedad” los primeros 14 días desde la exposición al virus, aunque la mayor parte de los pacientes muestran síntomas entre los días 4 y 5 tras la exposición. Durante la primera semana de la enfermedad, los síntomas son similares a los de la gripe y, en el caso de desarrollar la enfermedad severa, esto ocurre en la segunda semana tras el contagio.

En cuanto al espectro de la enfermedad, se estima que el 30-40 % de los casos son sujetos asintomáticos. De los sintomáticos, el 80 % son casos leves (sin neumonía o con neumonía leve), menos del 10 % de los casos son severos (disnea, hipoxia, y afectación pulmonar en 24-48h) y menos del 5 % son críticos (fallo respiratorio, shock, fallo multiorgánico). La tasa global de mortalidad se mantiene actualmente por debajo del 1 % (72).

2.1.6. Diagnóstico

El diagnóstico de COVID-19 combina evaluación clínica, pruebas de laboratorio y estudios de imagen. La evaluación clínica incluye la revisión de la clínica del paciente, el historial de exposición y la situación epidemiológica.

Las pruebas de laboratorio más comunes para el diagnóstico de COVID-19 son: la Reacción en Cadena de la Polimerasa (PCR, por sus siglas en inglés), que se diseña para detectar el material genético del virus y es considerada el estándar para el diagnóstico; las pruebas de antígenos, que detectan proteínas virales y son útiles para el diagnóstico rápido aunque menos precisas; y, por último, las pruebas de anticuerpos, que detectan la presencia de anticuerpos contra el SARS-CoV-2, indicando una infección pasada.

En casos de enfermedad grave se utilizan estudios de imagen como radiografías de tórax y tomografías computarizadas (TC) con el objetivo de evaluar el daño pulmonar y diferenciar la enfermedad de otras afecciones respiratorias como la gripe (73).

Además de la analítica y las pruebas de imagen rutinarias, se han llevado a cabo numerosos estudios sobre biomarcadores para el diagnóstico del COVID-19 grave. Aunque los resultados son muy heterogéneos dependiendo de la población analizada y el momento de la pandemia en el

que se realiza el estudio, los biomarcadores más frecuentemente observados relativos al diagnóstico son las citoquinas IL-6, IL-8, IL-18 y TNF (74), y algunos valores elevados en la analítica rutinaria como leucocitos, plaquetas, procalcitonina, ferritina y proteína C reactiva (75). En cuanto a los biomarcadores de severidad y pronóstico de la enfermedad, los más frecuentemente registrados son la saturación de oxígeno periférico (SpO₂), la saturación arterial de oxígeno (SaO₂), y los marcadores de lesión cardiaca como las troponinas I y T, la creatina quinasa-MB (CK-MB) o la mioglobina (76,77).

2.1.7. Tratamiento

2.1.7.1. Tratamiento sintomático

El tratamiento sintomático de la COVID-19 se centra en aliviar los síntomas. Las medidas han ido cambiando a lo largo de la pandemia, siendo las más comunes la hidratación, el reposo, y la toma de fármacos antipiréticos y analgésicos, como el paracetamol o el ibuprofeno, para reducir la fiebre y aliviar el dolor, y antitusivos que alivien la congestión nasal y la tos.

Ante la falta de tratamientos efectivos y la situación social crítica, se utilizaron varios medicamentos de manera experimental. La hidroxycloroquina fue un tratamiento muy extendido al inicio de la pandemia, aunque posteriormente se demostró que era ineficaz. Además, diferentes estudios demostraron que aquellos pacientes que recibieron tratamiento con hidroxycloroquina presentaron toxicidad, con aumento del intervalo QT, especialmente cuando se administró junto con azitromicina (78). La combinación de antirretrovirales lopinavir/ritonavir, eficaces en otras infecciones virales, fueron utilizados y, aunque redujeron la carga viral, no se demostró eficacia clínica. Finalmente, el antiviral remdesivir demostró reducir significativamente la sintomatología respiratoria y fue uno de los primeros fármacos en recibir la autorización para su uso de emergencia.

Otras estrategias innovadoras se fueron implementando en el tratamiento de los casos severos, como el uso de tocilizumab, un anticuerpo monoclonal humanizado que se une al receptor de la IL-6 impidiendo su interacción y, por tanto, reduciendo la señalización inflamatoria. Por otra parte, se estudió como tratamiento potencial contra la COVID-19 el plasma de convaleciente. Consiste en la utilización de plasma de pacientes que ya han superado la infección y que, por tanto, contiene anticuerpos específicos contra el SARS-CoV-2. Esta estrategia ha demostrado ser poco efectiva en población general, aunque sí ha dado buenos resultados en pacientes inmunocomprometidos (79).

Estos tratamientos se utilizaron en combinación con anticoagulantes, con el fin de reducir el riesgo de eventos tromboembólicos, y corticosteroides, (dexametasona, metilprednisolona, prednisona) utilizados en la práctica habitual como antiinflamatorios. La administración de

antibióticos, aunque muy extendida en los pacientes hospitalizados de la primera ola, se restringió posteriormente a los casos con alta sospecha de co-infecciones o infecciones nosocomiales, especialmente en la UCI. Su utilización indiscriminada se desaconseja ya que el uso innecesario de antibióticos contribuye al problema de la resistencia bacteriana (80).

El soporte ventilatorio ha sido una estrategia clave en el manejo de la COVID-19. En los casos de neumonía con insuficiencia respiratoria se utilizó la oxigenoterapia suplementaria. Por su parte, en los casos más severos, particularmente aquellos pacientes que sufrieron síndrome de distrés respiratorio agudo (SDRA), las estrategias más utilizadas fueron la ventilación mecánica invasiva (VMI) y no invasiva (VMNI), la cánula nasal de alto flujo (CNAF) y la oxigenación por membrana extracorpórea (ECMO, por sus siglas en inglés). Sin embargo, su uso prolongado promovió la aparición de infecciones nosocomiales que contribuyen a empeorar el estado de salud del paciente y a extender su estancia hospitalaria. En los periodos más críticos de la emergencia global, además, la falta de equipos de ventilación y personal capacitado fue un desafío significativo.

Actualmente, entre los tratamientos estándar, se mantienen el remdesivir, los corticosteroides y los anticuerpos monoclonales, aunque el uso de estos últimos ha disminuido debido a la variabilidad genética de las nuevas variantes. Se ha añadido el tratamiento con la combinación de los antivirales nirmatrelvir y ritonavir (Paxlovid) para pacientes con una enfermedad leve o moderada con alto riesgo de hospitalización. El nirmatrelvir es un inhibidor de la proteasa principal del SARS-CoV-2 Mpro, también conocida como proteasa 3C-like (3CLpro) o proteasa nsp5, crucial para la replicación viral. El ritonavir inhibe el metabolismo de nirmatrelvir, mediado por la enzima CYP3A, aumentando sus concentraciones plasmáticas. Este tratamiento se autorizó por la Unión Europea en enero de 2022 (81), con ciertas limitaciones en pacientes inmunocomprometidos debido al riesgo de efectos adversos, toxicidad e infecciones causado por la acumulación de fármacos habitualmente metabolizados por CYP3A.

2.1.7.2. Vacunas

Las vacunas contra la COVID-19 han sido fundamentales para controlar la pandemia. A finales del 2020 se habían desarrollado rápidamente varias vacunas en al menos 43 ensayos clínicos (82). Las vacunas de Pfizer-BioNTech y Moderna comenzaron a utilizarse en diciembre de 2020, priorizando inicialmente la vacunación de poblaciones vulnerables como ancianos residentes en centros de cuidados y personal sanitario.

Existen varios tipos de vacunas desarrolladas y aprobadas para su uso, incluyendo:

- **Vacunas de ARN mensajero (ARNm).** Estas vacunas, a pesar de ser las más populares y aceptadas durante la pandemia, no se habían utilizado en humanos hasta la

llegada del SARS-CoV-2. En ese momento se estaba investigando a nivel preclínico su uso en otras enfermedades (83). En este grupo se incluyen las de Pfizer-BioNTech (BNT162b2) y Moderna (ARNm-1273), que utilizan nanopartículas lipídicas para transportar el ARNm. Estas partículas, además de proteger el ARNm de su degradación prematura, facilitan su entrada en las células para que pueda ser traducido por los ribosomas. De esta manera las células producen la proteína del virus codificada por el material genético introducido desencadenando una respuesta inmune. Ambas han mostrado tasas elevadas de eficacia de, aproximadamente, el 95 % y 94 %, respectivamente, contra la COVID-19 sintomática.

- **Vacunas de virus inactivados.** Ejemplos de este tipo serían la CoronaVac de Sinovac, o la Sinopharm's BBIBP-CorV, que utilizan el virus inactivado para provocar una respuesta inmune. El virus se inactiva por métodos de calor, productos químicos o radiación, destruyendo su capacidad para replicarse y causar enfermedad. El virus inactivado no puede infectar células ni replicarse, pero las proteínas virales presentes en la vacuna mantienen intactas las estructuras que el sistema inmunitario puede reconocer para generar una respuesta. Se consideran más seguras y producen menos efectos secundarios que las de ARNm, pero su eficacia es menor, en un rango del 50 % al 80 %.
- **Vacunas de vectores virales.** La vacuna de AstraZeneca Vaxzevria (J07BX03), la de Johnson & Johnson (Ad26.COV2.S) y la Sputnik V pertenecen a este grupo. Estas vacunas utilizan un virus modificado que transporta material genético del SARS-CoV-2 al interior de las células para activar la respuesta inmune. Son efectivas, pero pueden causar más efectos secundarios, generalmente fiebre y dolor en la zona de la inyección, en comparación con las de virus inactivados. Sus tasas de eficacia rondan del 58 % al 92 %.
- **Vacunas de subunidades proteicas.** Estas vacunas, como la de Novavax, utilizan proteínas del virus, generalmente subunidades proteicas, para estimular la respuesta inmune. Son seguras y se han utilizado en varias poblaciones (84,85).

Las vacunas de virus atenuado, que se han utilizado para combatir otras infecciones virales, no se han utilizado en el caso del SARS-CoV-2 por varias razones. Este tipo de vacunas contienen virus debilitados que sí pueden replicarse en el organismo, pero que no poseen capacidad para provocar la enfermedad. Aunque son seguras para la mayoría, pueden no ser adecuadas para individuos con el sistema inmune comprometido. Adicionalmente, existe la posibilidad de que el virus atenuado pueda revertir a una forma más virulenta, aunque el riesgo es muy bajo. Precisamente por la necesidad de realizar comprobaciones más exhaustivas, su desarrollo requiere un proceso largo y complejo. Con otras alternativas eficaces, y ante la necesidad urgente de una

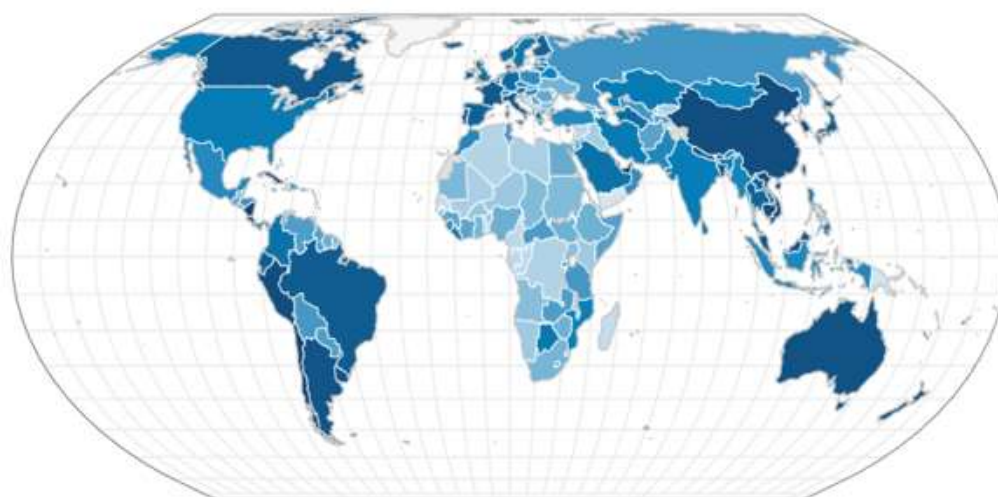
respuesta rápida y segura a la pandemia, las tecnologías mencionadas anteriormente fueron las preferidas (86,87).

Las vacunas han demostrado ser efectivas para reducir la gravedad de la enfermedad, las hospitalizaciones y las muertes. Las investigaciones al respecto han sido muy rápidas debido a la emergencia global y gracias al esfuerzo de la comunidad científica ante una situación sin precedentes. Debido a las diferentes variantes emergentes se ha necesitado rediseñar vacunas que mantuvieran su eficacia. De hecho, se siguen realizando esfuerzos por crear vacunas de nueva generación para la COVID-19 que mejoren la eficacia y seguridad de las de primera generación utilizando, por ejemplo, vías de administración menos agresivas como la inhalación o la administración oral (89). Por otra parte, se están llevando a cabo estudios para abordar la seguridad y eficacia de las vacunas contra la COVID-19 a largo plazo, especialmente en poblaciones con patologías previas (90).

La reticencia a las vacunas ha sido un importante desafío en el esfuerzo global para combatir la pandemia. A pesar de su eficacia y seguridad comprobadas, una parte sustancial de la población ha sido y sigue siendo reacia a recibirlas. Esta reticencia está influenciada por una variedad de factores, incluyendo la desinformación, los problemas de confianza y las variables demográficas. La cantidad de información que no proviene de fuentes fiables y que circula por las redes sociales es un factor importante, amplificando las dudas sobre la seguridad de las vacunas. Se ha demostrado que factores como la edad, la etnia, el nivel de estudios o el nivel de ingresos afectan a la aceptación de la vacuna (91). Es por ello que la confianza en los profesionales sanitarios y las instituciones juegan un papel fundamental. Las estrategias de divulgación que proporcionan información clara, consistente y basada en evidencias pueden ayudar a aumentar la confianza de la población y enfrentar la desinformación.

En diciembre de 2023 se habían administrado más de 13 billones de vacunas en dosis completa. A nivel global, se había llegado a un 32 % de población vacunada con dos dosis de la vacuna contra la COVID-19, y del 67 % en el caso de personas que habían recibido una sola dosis. Como se puede apreciar en la Figura 5, la proporción de población con la dosis completa de vacunación varía mucho entre diferentes zonas. En la mencionada fecha era de un 65 % en Europa, un 33 % en el continente africano, un 52 % en la zona del Mediterráneo oriental, un 72 % en América, un 70 % en el sudeste asiático y un 86 % en la región del Pacífico occidental (92).

Es complicado estimar estas cifras a día de hoy debido a que la OMS dejó de actualizar los datos en sus plataformas a finales de 2023, cuando se declaró finalizada la pandemia. Además, existen importantes discrepancias entre los datos clínicos y administrativos, y entre los datos nacionales y estatales o locales en las diferentes zonas (93).



% de población total vacunada de SARS-CoV-2



Figura 5. **Proporción de población vacunada contra el SARS-CoV-2 en el mundo hasta 2023 (incluido).** Mapamundi que muestra un sistema de color para indicar las diferencias en el porcentaje de población vacunada con, al menos, dos dosis de vacuna contra la COVID-19 en las diferentes zonas a fecha 31 de diciembre de 2023. Datos obtenidos de la OMS (92).

2.1.7.3. Terapias avanzadas

La Asociación Española de Medicamentos y Productos Sanitarios (AEMPS) define los “medicamentos de terapia avanzada” (ATMPs), por sus siglas en inglés, como “medicamentos de uso humano basados en genes (terapia génica¹), células (terapia celular²) o tejidos (ingeniería tisular³) que incluyen productos de origen autólogo, alogénico o xenogénico”. El marco legal para estos tratamientos lo constituyen el Reglamento (CE) N°1394/2007, el Real Decreto Legislativo 1/2015, de 24 de julio y el Real Decreto 477/2014, de 13 de junio, por el que se regula la autorización de ATMPs de fabricación no industrial.

¹ **Medicamento de terapia génica:** medicamento que incluye un principio activo que contiene un ácido nucleico recombinante, o está constituido por él, utilizado en seres humanos, o administrado a los mismos, con objeto de regular, reparar, sustituir, añadir o eliminar una secuencia génica (94).

² **Medicamento de terapia celular:** medicamento que contiene células o tejidos, o está constituido por ellos, que han sido objeto de manipulación sustancial de modo que se hayan alterado sus características biológicas, funciones fisiológicas o propiedades estructurales pertinentes para el uso clínico previsto, o por células o tejidos que no se pretende destinar a la misma función esencial en el receptor y en el donante (94).

³ **Producto de ingeniería tisular:** aquel que contiene o está formado por células o tejidos manipulados por ingeniería y del que se alega que tiene propiedades, se administra a las personas para regenerar, restaurar o reemplazar un tejido humano”. Se considera que son “manipulados por ingeniería” si “las células o tejidos han sido sometidos a manipulación sustancial, de modo que se logren las características biológicas, funciones fisiológicas o propiedades estructurales pertinentes para la regeneración, reparación o sustitución pretendidas” o bien si “las células o tejidos no están destinados a emplearse para la misma función o funciones esenciales en el receptor y en el donante (94).

Este tipo de terapias han sido exploradas para abordar diversas enfermedades, incluyendo la COVID-19. En su contexto, se han investigado varios ATMPs como tratamiento con el objetivo de mejorar los resultados clínicos de los pacientes. Las terapias génicas y tisulares, aunque pueden ser estrategias prometedoras para inducir la respuesta inmunitaria o reparar tejidos dañados, se encuentran aún en fase de investigación para confirmar su eficacia y seguridad.

Las terapias celulares, por su parte, se han utilizado en el tratamiento de pacientes durante la pandemia, permitiendo regular la respuesta inmune y reducir la tormenta de citoquinas. En particular, las células madre mesenquimales (MSC, por sus siglas en inglés), han sido las más estudiadas como tratamiento para la COVID-19. Estas células poseen propiedades inmunomoduladoras y antiinflamatorias, lo que hace que sea una terapia prometedora sobre todo en los casos severos y críticos. El tratamiento con MSCs ha demostrado reducir la mortalidad y la tasa de estancia hospitalaria, ha revelado propiedades regenerativas y se ha asociado con una disminución de la necesidad de ventilación mecánica invasiva (95). Otras terapias celulares, como las que utilizan células *natural killer* (NK), también han obtenido resultados preliminares prometedores debido a su rol en la prevención de la inflamación gracias a su capacidad para eliminar eficazmente células infectadas (96).

2.1.8. COVID persistente o *long* COVID

El término "*long* COVID" o "COVID persistente" se refiere a una continuación o desarrollo de nuevos síntomas pasados 3 meses de la infección inicial por SARS-CoV-2, prolongándose estos más de dos meses sin que haya un diagnóstico alternativo. Esta condición también se conoce como PASC (*post-acute sequelae of SARS-CoV-2*) (97).

Las personas con *long* COVID pueden experimentar una amplia variedad de síntomas que difieren en severidad y duración y que afectan a múltiples órganos y sistemas, incluyendo no solo el respiratorio, sino también el cardiovascular, el neurológico y el gastrointestinal. Uno de los síntomas más comunes es la fatiga extrema, que provoca incluso incapacidad para realizar ejercicio físico. En niños se ha visto que esta clínica puede aparecer y/o mantenerse hasta 12 meses después de la infección inicial (98). A nivel respiratorio también se suelen dar tos y dificultad respiratoria. Se han dado síntomas neurológicos como dolores de cabeza, anosmia y "niebla mental", siendo estos más frecuentes en mujeres. Por último, se han reportado también síntomas psicológicos como la depresión, la ansiedad e insomnio, aunque estos últimos se han asociado a su vez a la persistencia de los síntomas (99). De hecho, el *long* COVID puede tener un impacto significativo en la calidad de vida de los pacientes, afectando a su capacidad para trabajar, realizar actividades diarias y mantener relaciones sociales, debido a la continuidad del cuadro clínico y al agotamiento mental y físico.

Cualquier persona que haya tenido una infección por SARS-CoV-2 puede desarrollar *long COVID*, estimándose una prevalencia del 7,5 % al 41 % de las personas que han padecido COVID-19. Se han identificado algunos factores de riesgo que incluyen ser mujer, la edad avanzada, patologías previas, la falta de vacunación y una enfermedad inicial de COVID-19 grave, especialmente si esta incluyó la necesidad de ingreso en la UCI y de soporte ventilatorio (100). Por otra parte, la vacunación también parece reducir la prevalencia del *long COVID*, aunque no elimina el riesgo. Se ha observado un aumento de *long COVID* en pacientes no vacunados en comparación con los que han completado la pauta de vacunación, ya sea antes o después de la infección inicial (101).

Aunque los mecanismos exactos de *long COVID* no se comprenden completamente, se sabe que, de forma similar a la infección inicial, consiste en una desregulación del sistema inmune, especialmente el sistema del complemento y valores elevados de células T CD4+ y T CD8+, que incluye una inflamación persistente y una liberación de citoquinas que no cesa y puede llevar a síntomas crónicos. Esto contribuye, además, al agotamiento celular, impidiendo que el sistema no se recupere por completo (102,103).

Algunos estudios sugieren que la persistencia de fragmentos de ARN del SARS-CoV-2 en algunos órganos puede contribuir a los síntomas del *long COVID* (104). Adicionalmente, se ha observado la reactivación reciente de virus latentes, como el Virus de Epstein-Bar (EBV, por sus siglas en inglés), en una proporción significativa de pacientes con *long COVID*. Se cree que los niveles de anticuerpos en sangre contra el EBV están correlacionados con ciertas citoquinas asociadas a la patología del COVID-19. Esto sugiere que la reactivación no sería un evento aislado, sino que podría estar relacionada con la activación persistente del sistema inmune de los pacientes con *long COVID* (105).

Cabe destacar que los estudios que se están llevando a cabo son, a menudo, heterogéneos, impidiendo la estandarización de resultados y conclusiones. Por eso, a pesar de los progresos que se han realizado en la comprensión del *long COVID*, aún quedan muchas dudas al respecto de su fisiopatología, prevalencia y manejo.

2.1.9. La pandemia: impacto social y económico

La pandemia de COVID-19 ha tenido un impacto profundo y multifacético en la sociedad y la economía globales. La pandemia conllevó un colapso económico debido a los confinamientos y la supresión de muchas actividades, desencadenando una profunda recesión que afectó hasta a las economías más poderosas como China, Estados Unidos, la Unión Europea o Reino Unido. Sectores como la pesca, la silvicultura, el turismo o la hostelería vivieron un impacto significativo que derivó en pérdidas de empleo y reducciones operacionales. Esto fomentó el aumento de la

deuda pública a nivel mundial, que alcanzó niveles históricamente altos en países desarrollados (106).

La situación agravó las desigualdades salariales, especialmente en países como Estados Unidos o Brasil, donde las inequidades económicas se han intensificado en los últimos años. Los países de ingresos medios-bajos enfrentaron las reducciones más notables, afectando desproporcionadamente a grupos vulnerables (107). De hecho, muchos jóvenes con ingresos fluctuantes enfrentaron desafíos económicos, recurriendo a la deuda para gestionar sus gastos. Esto llevó a la toma de decisiones drásticas como abandonar la educación o asumir trabajos de alto riesgo.

A nivel social, la pandemia no solo interrumpió las interacciones sociales, sino también los sistemas de apoyo, lo que llevó a un aumento de la sensación de soledad y de aislamiento social. Así mismo, el estrés psicológico generado por el pánico, la ansiedad y la depresión asociadas a la posibilidad de contagio (conocido como “coronafobia”), contribuyó a la incertidumbre y a la expansión de la desinformación. Esto llevó a un periodo guiado por la confusión, la desconfianza en las autoridades sanitarias y la desobediencia de las guías y recomendaciones sanitarias, afectando al impacto del virus. La desinformación causó reticencia al uso de vacunas, mascarillas o tratamientos, lo que aumentó la morbilidad del virus y debilitó las medidas de seguridad pública. Además, la expansión de información falsa causó miedo y estrés psicológico y exacerbó tensiones sociales, políticas y religiosas (108).

Las dificultades económicas, el estrés emocional y la falta de interacciones sociales fuera del entorno familiar aumentaron la violencia doméstica a nivel global. Los problemas a los que las víctimas se enfrentaron van más allá de la violencia en sí misma. La imposibilidad de acceder a sistemas de apoyo, la incapacidad de alejarse de la persona que ejerce violencia, la falta de apoyo familiar directo y la imposibilidad de huir debido a las restricciones de movilidad, agravaron su situación. Estas circunstancias propiciaron el aislamiento y la vulnerabilidad de las víctimas, limitando sus opciones de búsqueda de ayuda y protección. De hecho, se observó un descenso significativo del uso de servicios de emergencia para este tipo de casos, lo que dificultó los trabajos de detección y actuación (109).

En el ámbito académico, los niños y adolescentes, así como los estudiantes de niveles superiores, enfrentaron desafíos a nivel de salud mental debido al cierre de centros académicos y la pérdida de interacciones sociales, tan necesarias en esa etapa (110). Además, los métodos docentes tuvieron que reinventarse para adaptarse a las plataformas online, lo cual dio lugar a resultados mixtos, desde alumnos que mejoraron su rendimiento hasta otros que tuvieron dificultades para afrontar la falta de presencia física e interacción personal. Lo que es cierto es

que la pandemia ha fomentado el desarrollo de metodologías docentes innovadoras, derivando en la posibilidad de plantear una reforma educativa y reasignar prioridades.

De forma similar, este nuevo formato ha llevado a un cambio en los sistemas de trabajo de muchas organizaciones y empresas, que han adoptado modelos totalmente remotos o híbridos para sus trabajadores. Esto presenta ventajas en cuanto a flexibilidad, ahorro de tiempo y comodidad, aunque conlleva desafíos como una mayor dificultad en la supervisión de empleados o la falta de comunicación en persona. Aun así, el aumento de puestos de trabajo remotos se ha asociado con un incremento en la satisfacción laboral y el rendimiento en algunos casos, así como la mejora en la confianza dentro de la organización (111).

En relación al medio ambiente, la pandemia ha tenido un impacto profundo dando lugar a cambios positivos y negativos. Por una parte, se produjo una mejora significativa en la calidad del aire y el agua gracias a la reducción de actividades industriales y de transporte durante las etapas de confinamiento. Por la misma razón, mejoró también la contaminación acústica. Por otro lado, aumentó la generación de residuos sólidos, especialmente desechos médicos y plásticos, a causa del uso masivo de equipos de protección personal y las interrupciones en las tareas de reciclaje (112).

Para asegurar que los beneficios ambientales no sean únicamente temporales, se han propuesto estrategias como reducir la dependencia de los combustibles fósiles o incrementar las políticas de gestión de residuos. Las investigaciones que estudian las relaciones entre ciencia ambiental y salud pública también han ganado interés, y se insta a estos grupos a desarrollar estrategias que equilibren la recuperación económica manteniendo la sostenibilidad ambiental (113).

2.1.10. Las olas de la pandemia en España

La pandemia ha experimentado diferentes olas, tal como muestra la Figura 6, caracterizadas por diferentes variantes y perfiles de pacientes. Estas olas han modificado las manifestaciones clínicas, la gravedad y la recuperación de los pacientes hospitalizados.

En cuanto a los parámetros analíticos, en los pacientes hospitalizados en la segunda ola disminuyeron los niveles de IL-6, proteína C reactiva y dímero D en sangre, y aumentó la tasa de subpoblaciones linfocitarias en etapas avanzadas de la enfermedad, sugiriendo una respuesta inmune menos severa en comparación con la primera ola (114).

A nivel pulmonar, los pacientes de las últimas olas mostraron mejor capacidad de recuperación, sugiriendo un cambio tanto en las características del paciente como en la eficacia del tratamiento a lo largo del tiempo. Hacia el final de la pandemia solamente el 3 % de los pacientes hospitalizados presentaba neumonía bilateral y se redujo significativamente la necesidad de soporte ventilatorio (115).



Figura 6. *Evolución de la pandemia por SARS-CoV-2 junto con los hitos y proyectos principales en España. Principales hitos (en marrón) y proyectos desarrollados por del ISCIII (en azul) relacionados con la pandemia, mostrados de forma cronológica según avanzaba la pandemia (nº de casos diarios) (infografía realizada por Lara Mazagatos) (116).*

La primera ola fue la de mayor severidad, asociándose problemas respiratorios graves sin una media de edad particularmente alta, mientras que en las últimas olas los pacientes acudían con un perfil de más edad, unos 79 años de media, y mayor cantidad de patologías previas pero sintomatología respiratoria más leve. Esto sugiere que los pacientes de más edad con patologías previas tenían más probabilidad de ser hospitalizados. En la primera ola, además, la tasa de mortalidad fue la más alta de toda la pandemia con una mayor incidencia de insuficiencia respiratoria grave e inflamación (115).

Las diferencias observadas entre pacientes hospitalizados entre el inicio y el final de la pandemia sugieren que la enfermedad ha evolucionado, que tiene unas características distintas y que, por tanto, requiere un enfoque diferente.

2.2. La inmunidad de grupo

La inmunidad de grupo, también conocida como inmunidad colectiva, es un concepto crucial en la epidemiología y la salud pública. Se refiere a la protección indirecta contra una enfermedad infecciosa que se logra cuando una proporción significativa de la población se vuelve inmune, ya sea a través de la vacunación o de haber superado la infección. Este fenómeno reduce la probabilidad de transmisión del patógeno, protegiendo así a los individuos no inmunizados. Cabe destacar que alcanzar la inmunidad de grupo no implica que la transmisión desaparezca.

En ocasiones, se denomina también inmunidad de grupo a la proporción de población inmunizada, independientemente de si esta proporción ha sobrepasado el umbral necesario para que esta inmunidad de grupo sea efectiva para frenar la expansión del patógeno (117). En este trabajo se tratará el concepto “inmunidad de grupo” como la primera definición mencionada: fenómeno que se alcanza cuando la proporción adecuada de población se inmuniza ante un patógeno, permitiendo así frenar su expansión.

La inmunidad de grupo se basa en la interrupción de las cadenas de transmisión del patógeno. Su expansión está limitada por la disponibilidad de huéspedes susceptibles. Por tanto, cuando una cantidad suficiente de personas en una comunidad es inmune, el patógeno encuentra menos individuos susceptibles a infectar y su propagación disminuye. Este principio ha sido fundamental en la erradicación de enfermedades como la viruela en 1980 y en el control de otras como el sarampión y la poliomielitis en la década de 1990, para las cuales aún existen programas de vacunación con el objetivo de mantener la inmunidad colectiva.

La inmunidad de grupo no solo protege a los individuos inmunizados, sino también a aquellos que no pueden vacunarse por diversas causas, véase condiciones médicas como alergias severas a los componentes de la vacuna, un sistema inmunitario comprometido o pertenecer a grupos minoritarios para los cuales la vacuna pueda no haber sido evaluada de forma específica.

Para calcular el umbral necesario para alcanzar la inmunidad de grupo (*Herd Immunity Threshold* o HIT), la fórmula utilizada es la siguiente:

$$HIT = 1 - \frac{1}{R_0} \quad [1]$$

donde R_0 es el número básico de reproducción. Si el valor de R_0 es mayor a uno, el número de casos aumentará y la enfermedad puede causar una epidemia, mientras que si es menor que uno, el número de casos descenderá. Este valor se calcula utilizando diferentes métodos según los datos disponibles:

- Modelo de la tasa de contacto: $R_0 = \beta \times \kappa \times D$ [2] donde β es la tasa de transmisión por contacto, κ es el número promedio de contactos por unidad de tiempo y D es la duración del periodo infeccioso.
- Modelo de generación de casos: $R_0 = C/S$ [3] donde C es el número de casos y S es el número de individuos susceptibles en la población.
- Modelo de crecimiento exponencial: $R_0 = 1 + r \times D$ [4] donde r es la tasa de crecimiento logarítmica y D es la duración del periodo infeccioso.
- Modelos matemáticos complejos. En algunos casos se utilizan modelos computacionales complejos que incorporan factores como la estructura poblacional, la posible variabilidad en la susceptibilidad al contagio o la heterogeneidad de los contactos.

Siendo así, el valor del umbral para alcanzar la inmunidad de grupo varía al estar influido por diferentes factores. En la Hepatitis A, se estima un umbral del 55 % al 80 % (118), mientras que otras enfermedades como el sarampión o la difteria, por sus altas tasas de contagio, requieren de un umbral de más del 90 % para alcanzar la inmunidad de grupo. En el caso del SARS-CoV-2, se estimó inicialmente un umbral del 60-70 % a nivel global, aunque este valor ha ido variando debido a las diferencias en las características de la población, factores ambientales y metodologías utilizadas para calcular R_0 (119,120).

La principal estrategia para lograr la inmunidad de grupo contra la COVID-19 ha sido la vacunación masiva. Las vacunas desarrolladas han demostrado ser altamente efectivas para prevenir la enfermedad grave y la muerte, y su implementación ha sido crucial para reducir la carga de la pandemia. La inmunidad adquirida por aquellos que han superado la infección de forma natural también contribuye a alcanzar la inmunidad colectiva. Sin embargo, depender únicamente de la inmunidad natural no es una estrategia viable debido al riesgo de enfermedad grave y muerte, así como a la presión sobre los sistemas de salud. De hecho, estudios que comparan diferentes estrategias de confinamiento y restricción social muestran que, aunque alcanzar la inmunidad de grupo de forma natural con menos restricciones (el caso de Suecia, Brasil o Estados Unidos) puede llevar a acortar la duración de la pandemia, también resulta en una tasa de mortalidad más alta en comparación con las zonas que implementaron medidas de distanciamiento social (121,122). Estas diferencias están esquematizadas en la Figura 7.

Las estrategias de vacunación han presentado desafíos únicos debido a la naturaleza del virus SARS-CoV-2 y su capacidad para mutar, lo que puede afectar a la eficacia de las vacunas y la duración de la inmunidad adquirida. Otras dificultades significativas han sido la variabilidad en la respuesta inmunitaria entre individuos, la reticencia a recibir la vacuna y las barreras socioeconómicas y logísticas en ciertas zonas a la hora de llevar a cabo campañas de vacunación

eficientes. Por ejemplo, en zonas como Afganistán, la ineficiencia de las cadenas de frío para mantener los fármacos a la temperatura adecuada durante su distribución, la economía o las restricciones que enfrentan las mujeres a la hora de acceder al sistema de salud, multiplica las complicaciones para alcanzar la inmunidad de grupo (123).

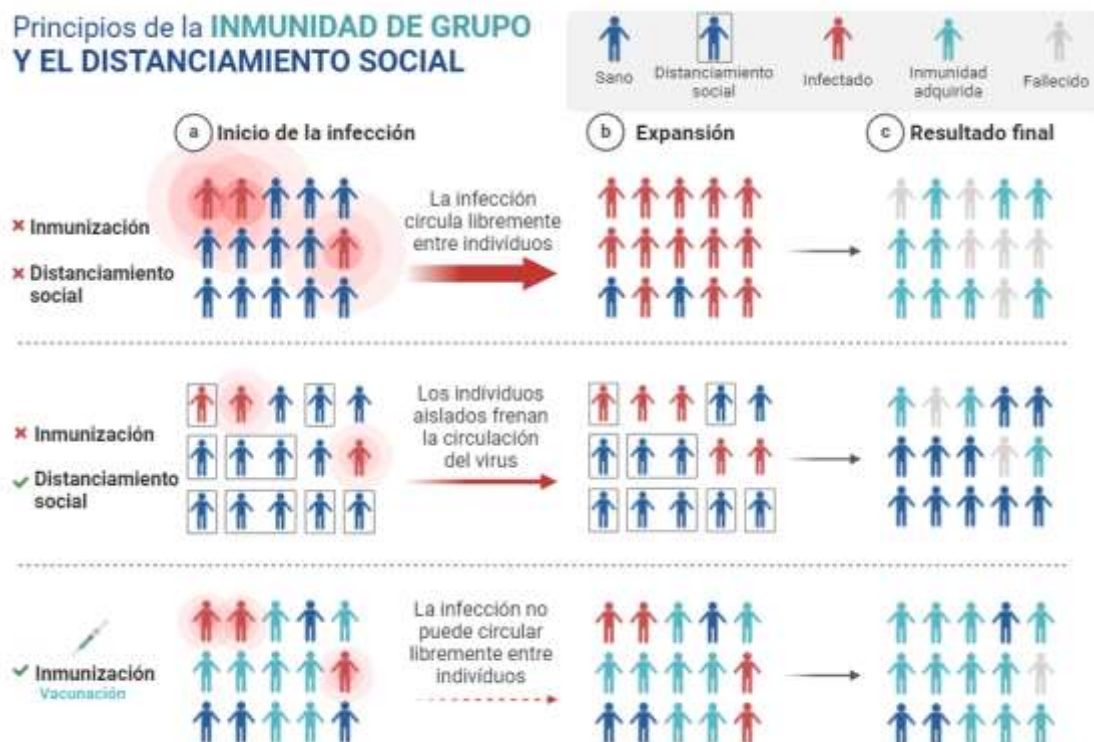


Figura 7. Ilustración del funcionamiento de la inmunidad de grupo. La infografía muestra los efectos del distanciamiento físico y social y de la inmunidad de grupo, obtenida a través de la vacunación, en comparación con los resultados de una situación sin estas medidas.

Alcanzar la inmunidad colectiva tiene como objetivo la relajación de medidas de control como el distanciamiento social y el uso de mascarillas, facilitando el retorno a la normalidad y la recuperación económica. En las zonas con sistemas de salud robustos, con alta cobertura vacunal y con capacidad para realizar una vigilancia epidemiológica, se propuso una estrategia de levantamiento gradual de restricciones y medidas de control, con el objetivo de contribuir a desarrollar una inmunidad de grupo de forma controlada, protegiendo especialmente a las personas vulnerables. Fue el caso de países como Noruega, Reino Unido, Italia, España (124), Australia o China (125). Este enfoque requiere una consideración cuidadosa de los sistemas de salud, la ética y la economía. No obstante, es importante mantener una vigilancia continua y estar preparados para ajustar las estrategias en respuesta a nuevas amenazas.

2.3. Innovación tecnológica y avances en investigación

La pandemia de COVID-19 ha sido un impulso sin precedentes para la innovación tecnológica y los avances en la investigación científica. Desde el inicio de la crisis sanitaria, la

comunidad global ha presenciado una aceleración en el desarrollo y la implementación de tecnologías avanzadas, así como en la colaboración científica a nivel mundial.

Como se ha mencionado anteriormente, el desarrollo de vacunas eficaces en tiempo récord, especialmente las de ARNm, ha marcado un hito en la historia de la medicina. Además, la inteligencia artificial (IA) y el *big data* han jugado un papel crucial en el seguimiento y la predicción de la propagación del virus, permitiendo a los gobiernos y organizaciones de salud pública tomar decisiones informadas y oportunas. Sin embargo, aún se investigan muchas otras soluciones prometedoras que continúan siendo de gran importancia en la actualidad para mejorar nuestra preparación y respuesta ante futuras amenazas sanitarias.

2.3.1. Técnicas de diagnóstico

Durante la pandemia, la atención sanitaria se vio desbordada por la necesidad de aumentar la tasa de diagnósticos con el objetivo de frenar la expansión del virus. Una tecnología que se ha propuesto para esto ha sido el diagnóstico basado en CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*). Este método, que puede detectar el virus en unos 40 minutos, se conoce como CRISPR-COVID y no requiere equipos sofisticados. Una de las estrategias basadas en esta tecnología es la técnica SHERLOCK (*Specific High-sensitivity Enzymatic Reporter unLOCKing*), en la que la proteína Cas13a se utiliza para escindir el ARN viral, liberando una señal fluorescente que indica la presencia del mismo (126). Similar a esta técnica, se ha desarrollado el sistema DETECTR (*DNA Endonuclease-Targeted CRISPR Trans Reporter*), que utiliza el mismo mecanismo, pero con la enzima Cas12 en lugar de la Cas13. Por último, también se pueden utilizar dos enzimas CRISPR en tándem, Cas13 y Csm6, para detectar la presencia de SARS-CoV-2 en una muestra. Es el caso de la técnica CARMEN (*Combinatorial Arrayed Reactions for Multiplexed Evaluation of Nucleic acids*), en la cual la escisión del ARN viral por parte de la enzima Cas13 libera una molécula activadora que activa la enzima Csm6, la cual escinde moléculas fluorescentes que se añaden como sustrato al sistema. De esta manera se amplifica dicha señal fluorescente permitiendo una detección sensible y rápida (normalmente no más de 20 minutos) (127).

Estas tecnologías han sido desarrolladas y validadas en estudios de laboratorio, pero su uso práctico a gran escala está aún en fase de desarrollo. Se han llegado a utilizar en algunos países como herramienta de emergencia, sin llegar a reemplazar a las pruebas tradicionales como la PCR (128,129). Aunque han mostrado resultados prometedores en términos de rapidez y precisión, deben realizarse estudios adicionales para confirmar su sensibilidad y fiabilidad en el ámbito clínico, además de requerir la validación y adaptación a los sistemas de salud existentes.

Otra técnica de diagnóstico que está en fase de investigación es el uso de nanomateriales bidimensionales (2DNMs, por sus siglas en inglés). Se trata de materiales extremadamente delgados con características eléctricas, ópticas y bioactivas únicas, lo que los hace muy interesantes para aplicaciones médicas. Estas propiedades permiten el desarrollo de sistemas de diagnóstico en tiempo real utilizando biosensores de alta sensibilidad en la detección de SARS-CoV-2. Integrar estos sistemas con plataformas online permitiría realizar valoraciones rápidas, precisas y rentables en formato portátil, sencillo y accesible para cualquier persona (130), lo que sería ideal para entornos de recursos limitados. Este tipo de técnica puede llegar a ser muy útil en el rastreo y predicción de la propagación de futuras pandemias o situaciones similares.

2.3.2. Vacunas de nueva generación

Se están tratando de estudiar vacunas de nueva generación contra la COVID-19 explorando diferentes mecanismos para potenciar la protección inmunológica. Una de las estrategias es desarrollar vacunas que actúen contra múltiples antígenos, como la proteína S y las proteínas de la nucleocápside, para generar una respuesta más robusta y aumentar la capacidad de protección a largo plazo frente a las diferentes variantes (131,132). En este contexto, se están desarrollando vacunas con *virus-like particles* (VLP) y vectores del virus Orf adaptado, siendo ambas plataformas prometedoras debido a su seguridad y a su capacidad para presentar múltiples antígenos (133,134).

Algunos grupos de investigación exploran nuevos métodos de administración de estas vacunas de nueva generación, con el objetivo de mejorar su seguridad y eficacia, optimizar la respuesta inmunitaria y aumentar las facilidades para su distribución y almacenamiento. Como ejemplo de esto, se ha desarrollado una plataforma de nanopartículas en capas (LbL, por sus siglas en inglés “*layer-by-layer*”), basada en quitosano trimetilado para la administración segura de múltiples antígenos de forma escalonada, mostrando resultados prometedores en estudios preclínicos en cuanto a su eficacia para una protección rápida y amplificadora (135). También se están investigando sistemas de administración de tipo lipopoliplex (LPP) para transportar el material genético, como el ARNm, a las células. Sus propiedades y características permiten limitar la toxicidad de la vacuna restringiendo la expresión de ARNm al sitio de la inyección, debido a su tamaño y composición, de tal manera que la liberación de ARNm por estas partículas es dirigida y controlada. Este sistema ha demostrado una reducción significativa de la carga viral en estudios preclínicos en ratones (136).

Una estrategia prometedora investigada a lo largo de toda la pandemia ha sido la vacunación a través de las mucosas, por ejemplo, por vía intranasal. Este tipo de vacunas provocaría fuertes respuestas inmunitarias a nivel local gracias a la activación de la inmunoglobulina A (IgA) secretora, un anticuerpo que se encuentra en las mucosas y juega un

papel crucial en la defensa contra infecciones. Este anticuerpo puede neutralizar virus y bacterias, evitando su adherencia y penetración en el organismo(137). La activación de una respuesta inmunitaria local puede, por tanto, ayudar a eliminar el virus más rápidamente reduciendo la duración y severidad de los síntomas, o incluso evitando la infección. Otras ventajas que hacen de esta vía de administración una buena estrategia son la posibilidad de autoadministración y la facilidad de almacenamiento y distribución en comparación con vacunas intramusculares.

En España, el Centro Nacional de Biotecnología asociado al Consejo Superior de Investigaciones Científicas (CNB-CSIC), en colaboración con otros grupos, está desarrollando una vacuna intranasal contra la COVID-19 que ha logrado un 100 % de eficacia en estudios preclínicos. Actualmente, está en fase de producción a mayor escala, y se encuentra a la espera de comenzar los estudios clínicos en humanos (138,139).

Estas iniciativas fomentaron la creación del proyecto NextGen, una propuesta que surge de la colaboración de varios organismos gubernamentales, que avala la investigación de vacunas y tratamientos de nueva generación para la COVID-19. Su objetivo es generar datos clínicos para abordar los posibles desafíos que plantean las nuevas variantes emergentes del virus, mejorar las capacidades de los países en la lucha contra la COVID-19 y potenciar la preparación para futuras pandemias (89).

2.3.3. Nuevos tratamientos

A pesar de que se ha instaurado un tratamiento estándar para los pacientes con COVID-19, la investigación continúa avanzando para desarrollar nuevas y más efectivas opciones terapéuticas. Estos esfuerzos han dado lugar a una serie de posibles tratamientos que están actualmente en fases de investigación y que buscan no solo mejorar la eficacia en la lucha contra el SARS-CoV-2 y sus variantes, sino también ofrecer tratamientos más seguros y adaptados a las necesidades de los pacientes. Algunos ejemplos de estas nuevas investigaciones son fármacos como el ensitrelvir, el asunercept o el molnupiravir, y tratamientos de terapias avanzadas como las células CAR-T o las terapias génicas y tisulares.

El ensitrelvir, un inhibidor de proteasa oral dirigido a la 3CLpro del SARS-CoV-2, crucial para su replicación, ha demostrado una elevada eficacia al reducir rápidamente los niveles de carga viral en el organismo. Su uso ha sido aprobado en Japón como tratamiento de emergencia, pero está en proceso de investigación clínica para confirmar su seguridad y eficacia en la práctica habitual (140).

Para casos no severos se está estudiando el molnupiravir, un antiviral que ha manifestado una alta eficacia en la reducción de la hospitalización, la disminución de la mortalidad de pacientes no hospitalizados, la resolución rápida de síntomas y una eficaz eliminación de la carga

viral. Su mecanismo de acción se conoce como “catástrofe de error viral”, ya que su metabolito activo (β -D-N4-hidroxicitidina o NHC) se incorpora al ARN viral, provocando una acumulación de errores en su genoma que inhibe su replicación. Ha mostrado ser especialmente útil cuando se administra en las primeras fases de la infección (141,142). La combinación de molnupiravir con el compuesto denominado Mpro61 ha demostrado efectos positivos en estudios con animales, creando un efecto sinérgico que potencia sus efectos. Mpro61 es un inhibidor de la proteasa principal del SARS-CoV-2 Mpro, crucial para la replicación. Aunque realiza una función semejante a la de nirmatrelvir, tiene un metabolismo más lento y, por tanto, no requiere de un inhibidor metabólico. Por ello, esta combinación molnupiravir-Mpro61 podría ser una alternativa más segura en comparación con el Paxlovid (nirmatrelvir/ritonavir), utilizado actualmente en la práctica clínica habitual, siendo una opción más tolerable en el caso de pacientes inmunocomprometidos o que ya tomen una medicación de base (143).

Para casos severos, se está estudiando, entre otros, el asunercept, un inhibidor del ligando Fas (FasL) o CD95L. La sobreexpresión de este ligando ocurre debido a la liberación excesiva de citoquinas proinflamatorias, como ocurre durante la infección por SARS-CoV-2. El fármaco ha demostrado una recuperación rápida y una reducción de la mortalidad en un 20 % en pacientes hospitalizados con COVID-19 severo (144).

Fármacos como nafamostat o camostat, que pueden actuar como inhibidores de TMPRSS2 y se utilizan en Japón y Corea como anticoagulantes y antiinflamatorios (145), o diferentes inhibidores de furina (146), se han estudiado como posibles tratamientos contra la COVID-19. El objetivo era evaluar su eficacia inhibiendo la activación de la proteína S y la fusión de la membrana viral con la de la célula huésped. Sin embargo, además de no haber obtenido aún resultados suficientemente prometedores a nivel preclínico, se debe tener en cuenta que inhibir estas proteínas puede generar efectos nocivos dado que tienen funciones importantes en los procesos fisiológicos normales (147). Otros fármacos que no han arrojado, por el momento, resultados suficientemente concluyentes como para ser adoptados como tratamiento estándar para la COVID-19 son la ivermectina, un antiparasitario de amplio espectro utilizado en países de América Latina y Asia (148); o el favipiravir, un antiviral aprobado para uso de emergencia en Japón, Rusia e India (149).

Otra estrategia prometedora sería la de potenciar la respuesta inmune para que su acción sea más rápida y eficiente. Esto se puede conseguir mediante técnicas de estimulación celular, junto con el uso de adyuvantes y citoquinas. Se ha visto, por ejemplo, que utilizar bloqueadores de los receptores de las células NK, como el NKG2A o el KIR2DL1, parece potenciar su acción citotóxica contra células infectadas (150). Por otro lado, se han descrito técnicas para activar, directa o indirectamente, las células B y T, ya que la combinación de sus funciones características

asegura que el sistema inmune pueda responder de forma específica y consistente. Algunas opciones para ello son el uso de adyuvantes como el hidróxido de aluminio o el MF59 para activar las células presentadoras de antígenos, la exposición a antígenos específicos del SARS-CoV-2 o la infusión de citoquinas. Aunque todos estos métodos podrían utilizarse en la fabricación de vacunas y tratamientos, se necesitan más estudios clínicos para comprobar su seguridad y eficacia en pacientes con COVID-19 (151).

Los últimos avances en terapias avanzadas han centrado sus investigaciones en aprovechar las capacidades inmunomoduladoras y regenerativas de varios tipos de células inmunes. Una de estas estrategias se basa en la expansión de células CTL específicas para proteínas estructurales del SARS-CoV-2 en un medio rico en interleucinas proinflamatorias. Experimentos *in vitro* han demostrado que estas células mantienen su capacidad citotóxica durante su expansión y aislamiento. La estrategia está en fase de ensayos clínicos para evaluar su eficacia en casos de COVID-19 leve, moderado y severo. Otra innovadora técnica sería el tratamiento con células CAR-T como antiviral contra la COVID-19. Habitualmente, las células CAR-T son células T modificadas genéticamente para conferirles la habilidad específica de reconocer células tumorales, pero han surgido como tratamiento prometedor para la COVID-19 al modificarse para reconocer antígenos virales del SARS-CoV-2, concretamente la proteína N. Además de las mencionadas, se están llevando a cabo ensayos clínicos y preclínicos para determinar la seguridad y eficacia de terapias avanzadas basadas en células NK, células T_{reg}, células dendríticas, monocitos y macrófagos, siendo las MSCs las de mayor interés en este campo (152).

Por su parte, la terapia tisular también ha generado interés en el campo de la COVID-19. Se han desarrollado tejidos impresos en 3D utilizando una matriz de bio-tinta de quitosano/policaprolactona con el objetivo de regenerar tejidos pulmonares (153). Sin embargo, este enfoque aún se mantiene en fase de laboratorio y requiere más investigación.

2.3.4. Investigación con organoides

Los organoides son sistemas de cultivos celulares tridimensionales que pretenden simular estructuras orgánicas humanas y sus funciones, actuando como modelos para estudiar el comportamiento de dichos órganos ante patologías o tratamientos. Esta tecnología ha surgido como una potente herramienta durante la pandemia, permitiendo realizar estudios relacionados con la patogenia del virus y la identificación de dianas terapéuticas.

En el caso de la COVID-19, esta técnica se ha utilizado para crear modelos de diferentes órganos, como organoides pulmonares, que han permitido estudiar cómo las MSCs interactúan con los fibroblastos durante la reparación tisular (154); organoides cerebrales, que han ayudado a comprender el impacto del virus en las neuronas y células gliales para explicar los síntomas

neurrológicos asociados al *long* COVID (155); o los organoides renales, que han mostrado la habilidad del virus para infectar directamente a células del túbulo proximal del riñón siendo esencial el receptor ACE2 para ello (156).

Estos modelos apoyan la detección de fármacos y el desarrollo de vacunas al proporcionar una plataforma que refleja la fisiología humana. Por ejemplo, se han identificado 33 genes que podrían ser buenas dianas terapéuticas para el tratamiento de la COVID-19. Entre ellos, el que codifica para IFNAR2 (*interferon- α/β receptor subunit 2*), proteína que forma parte del receptor de IFN involucrado en la vía de señalización JAK-STAT y en la regulación de la respuesta inmune; o el que codifica para TYK2 (*tyrosine kinase 2*), enzima de la familia JAK involucrada en la señalización de citoquinas e interferones (154).

A pesar de sus ventajas, los organoides no tienen, actualmente, la capacidad para madurar completamente o desarrollar la complejidad de un órgano o sistema orgánico real. Esto los convierte en un buen enfoque preclínico pero que requiere de una validación adicional de los resultados mediante otros estudios clínicos y preclínicos.

2.3.5. Tecnología *organ-on-a-chip*

Los sistemas de tipo *organ-on-a-chip* consisten en plataformas microfluídicas que replican la fisiología orgánica creando un modelo adecuado para estudiar enfermedades y tratamientos, en comparación con otros métodos tradicionales de estudios *in vitro*. La pandemia ha acelerado el desarrollo y aplicación de estos sistemas, ya que permiten el estudio del virus en diferentes sistemas de órganos, lo que facilita obtener resultados sobre su patogenicidad y tropismo.

El sistema PREDICT96-ALI, una plataforma *organ-on-a-chip* de alto rendimiento, se ha utilizado, por ejemplo, para evaluar de forma rápida la eficacia de medicamentos para la COVID-19 como remdesivir, tocilizumab o MPro6 (157). También se han utilizado estos sistemas para estudiar diferentes respuestas inflamatorias en pacientes infectados que cuentan además con comorbilidades previas, como la diabetes tipo II, mostrando que estos pacientes tienen una respuesta inflamatoria más potente (158).

Este tipo de estudios presenta ciertas limitaciones, como la necesidad de vascularización constante o la dificultad para integrar varios sistemas de órganos en un solo chip. Se espera que los próximos avances en estas investigaciones permitan integrar la inteligencia artificial para potenciar la funcionalidad y eficiencia de estas plataformas. Además, se debe realizar una adaptación adecuada de los laboratorios para el uso seguro y eficaz de esta tecnología en la investigación de enfermedades infecciosas, lo cual dificulta, encarece y atrasa los avances.

2.4. Inteligencia artificial

El concepto de usar ordenadores para simular la inteligencia humana y el pensamiento crítico se describió por primera vez en 1950 por Alan Turing, aunque no fue hasta 1956 cuando John McCarthy dio una definición oficial de Inteligencia Artificial (IA), siendo esta “la ciencia e ingeniería de hacer máquinas inteligentes” (159). Actualmente definimos la IA como “la capacidad de una máquina para imitar funciones cognitivas humanas como el aprendizaje y la resolución de problemas” (160). Caracterizada por su habilidad para emular el pensamiento lógico, el aprendizaje y la resolución de problemas de los humanos, la IA ha evolucionado en las últimas décadas y ha pasado del plano teórico al práctico, aplicándose de forma habitual en campos como las finanzas o la medicina.

A medida que la cantidad de datos generados por diversas fuentes crece exponencialmente, se vuelve evidente que los métodos tradicionales de análisis de datos no son suficientes para extraer información útil de manera eficiente, especialmente teniendo en cuenta la creciente complejidad de los problemas que enfrentamos en campos como la medicina o la tecnología. Es por esto que surgen subcategorías de la inteligencia artificial con tecnologías más avanzadas y específicas para diferentes funciones, como el *Machine learning* (ML) o aprendizaje automático, que permite a los sistemas aprender de los datos que procesan, o el Deep Learning (DL) o aprendizaje profundo, que utiliza redes neuronales artificiales para modelar patrones complejos en grandes volúmenes de datos (161).

El ML se centra en el desarrollo de algoritmos que permiten a las máquinas usar diferentes métodos para extraer un aprendizaje de los datos y mejorar su rendimiento con el tiempo sin intervención humana explícita. Esto permite automatizar tareas que serían demasiado complejas o imposibles de programar manualmente, como el reconocimiento de patrones en conjuntos masivos de información. Gracias a esto los sistemas pueden hacer predicciones y/o automatizar procesos de toma de decisiones de forma precisa mejorando su rendimiento a medida que se le proporcionan más datos, adaptándose a nuevos patrones y tendencias (162).

En concreto, esta tecnología ha revolucionado el campo de la medicina gracias a su capacidad para procesar grandes volúmenes de datos en poco tiempo, ya que permite realizar diagnósticos precisos, personalizar tratamientos o predecir la evolución de enfermedades. Además, facilita el monitoreo continuo de pacientes mediante dispositivos portátiles para enfermedades crónicas o la intervención temprana o preventiva.

A pesar de sus múltiples ventajas, los algoritmos utilizados, especialmente en medicina, son supervisados por humanos para evitar predicciones o diagnósticos erróneos (163). De hecho, todavía existe cierta desconfianza en los sistemas automatizados por parte de la población general,

pero también de los clínicos y sanitarios, ya que la naturaleza de “caja negra” de estos modelos dificulta que se pueda comprender completamente cómo se toman las decisiones o se generan los resultados (164). Para abordar estos desafíos, es crucial implementar estrategias que garanticen la calidad de los datos y aseguren una formación e infraestructura adecuada para la integración con los sistemas de salud existentes, manteniendo un equilibrio entre el uso de la tecnología y el juicio clínico. Por ello, sigue siendo imprescindible la validación clínica del rendimiento de cada modelo por los profesionales sanitarios.

Esta tecnología se ha aplicado de forma regular en el campo del análisis y diagnóstico por imagen (165–167), donde el reconocimiento de patrones es fundamental. De hecho, la FDA ya en 2021 avalaba más de 400 algoritmos de IA como dispositivos médicos aprobados (168). Otros ejemplos de las aplicaciones de modelos de ML son el diagnóstico de cáncer mediante la clasificación de células obtenidas en biopsias de mama (169), el diagnóstico, pronóstico y desarrollo de nuevas terapias para enfermedades neurodegenerativas (170), o el uso de algoritmos para el fenotipado y la predicción de riesgo de enfermedades cardiovasculares (171).

Durante la pandemia de la COVID-19, el ML demostró ser una herramienta valiosa, ya que se utilizaron algoritmos para tareas como la predicción de la propagación del virus (172), la identificación de pacientes de alto riesgo o la asignación de recursos médicos (173). Además, se aplicaron estos modelos en el análisis de imágenes médicas para una detección rápida de signos de infección pulmonar en radiografías y TCs, mejorando la precisión y velocidad del diagnóstico. También se ha utilizado para analizar biomarcadores en parámetros de laboratorio en busca de signos de infección, facilitando el triaje en los servicios de urgencias mediante una detección temprana efectiva antes de los resultados de las PCR (174). Estas aplicaciones ayudaron a los profesionales de la salud a tomar decisiones informadas en tiempo real y contribuyeron a la gestión eficiente de la pandemia.

3. Hipótesis y objetivos

La pandemia de COVID-19, causada por el virus SARS-CoV-2, ha tenido un impacto sin precedentes en la salud pública global desde su aparición a finales de 2019. La rápida propagación del virus llevó a la implementación de medidas de emergencia sanitaria en todo el mundo, incluyendo confinamientos, distanciamiento social y el desarrollo acelerado de vacunas.

A lo largo de las diferentes olas de la pandemia, se han observado variaciones significativas en la presentación clínica de la enfermedad, así como en los factores que influyen en la mortalidad de los pacientes hospitalizados. Estas variaciones pueden atribuirse a múltiples factores, incluyendo la evolución del virus, la implementación de nuevas estrategias de tratamiento y la creciente inmunidad de la población, tanto natural como inducida por las vacunas.

Las características clínicas y los predictores de mortalidad en pacientes hospitalizados con COVID-19 varían significativamente entre las diferentes olas de la pandemia. La primera ola, caracterizada por una falta de inmunidad previa y la ausencia de tratamientos específicos, presentó un perfil clínico y predictores de mortalidad distintos en comparación con las olas posteriores. La evolución de los patrones de mortalidad está influenciada tanto por la inmunidad natural como por la inmunidad otorgada por las vacunas, que comenzaron a administrarse masivamente durante las olas intermedias y finales. Por otra parte, los modelos de predicción basados en técnicas de aprendizaje automático, por sus múltiples ventajas y utilidades, pueden ser validados e implementados para su uso en futuras epidemias o pandemias, mejorando así la capacidad de respuesta del sistema de salud.

Por todo ello, el objetivo principal de este trabajo es profundizar en las variables que pueden tener mayor influencia en la predicción de mortalidad de pacientes hospitalizados por COVID-19, a la vez que se implementa y valida una herramienta predictiva y diagnóstica en el seguimiento de la enfermedad. Con esta finalidad, se plantearon los siguientes objetivos secundarios:

1. Analizar las características clínicas de los pacientes hospitalizados con COVID-19 en el Hospital Universitario Río Hortega de Valladolid.
2. Determinar los predictores de mortalidad de estos pacientes en los tres grupos temporales elegidos: 1ª ola, 2ª a 4ª olas y 5ª a 6ª olas.
3. Estudiar la evolución de los patrones de mortalidad en estos pacientes relacionándola con la situación de emergencia sanitaria en cada momento de la pandemia mencionado.
4. Identificar biomarcadores de inflamación en pacientes hospitalizados con COVID-19 en la Unidad de Cuidados Intensivos del Hospital Universitario Río Hortega de Valladolid durante la primera ola.

5. Proponer modelos de predicción de patrones de mortalidad gracias a técnicas de aprendizaje automático, con el objetivo de identificar enfermos con mayor riesgo de mortalidad y validar el método de cara a su potencial uso frente a nuevas epidemias o pandemias.

4. Pacientes y métodos

4.1. Pacientes

Para la realización de este estudio se ha incluido un total de 898 pacientes con diagnóstico de COVID-19 confirmado por PCR hospitalizados en el Hospital Universitario Río Hortega de Valladolid.

De ellos, 291 pacientes se incluyen en el periodo de la primera ola (entre febrero y abril de 2020), 138 en el periodo de la segunda ola (entre julio y diciembre de 2020), 114 en la tercera ola (de enero a febrero de 2021), 111 en la cuarta ola (de marzo a mayo de 2021), 121 en la quinta ola (de junio a septiembre de 2021) y 123 en la sexta ola (de octubre de 2021 a marzo de 2022).

Los pacientes se clasificaron en tres grupos temporales: primera ola (291 pacientes), segunda a cuarta olas (363 pacientes) y quinta a sexta olas (244 pacientes). Además, se creó un grupo específico con 36 pacientes de la primera ola que requirieron ingreso en la UCI.

4.1.1. Recogida de datos

Se realizó un estudio retrospectivo observacional mediante la revisión sistemática de historias clínicas en formato papel y/o electrónico de todos los pacientes incluidos. Los datos recogidos fueron de tipo demográfico, clínico y de laboratorio. Se recogieron datos tanto del momento de admisión del paciente, como de su evolución durante la estancia hospitalaria. En los pacientes que requirieron admisión en la UCI, también se recogieron estos datos tras quince días con VMI en dicha unidad.

Los datos se introdujeron manualmente en la base de datos prediseñada RedCap®, que incluía todas las variables de estudio. Posteriormente, los datos se volcaron a una tabla de Microsoft Excel para su manejo. La inclusión de datos se realizó de forma anonimizada, dando a cada paciente un número aleatorio.

4.1.2. Variables de estudio

4.1.2.1. Variables relacionadas con el estado previo del paciente

Los datos demográficos recogidos de cada paciente fueron el sexo y el año de nacimiento.

Se recogieron como antecedentes médicos aquellas comorbilidades consideradas relevantes y que hubieran sido diagnosticadas según criterios clínicos, analíticos y de imagen habituales: hipertensión arterial (HTA), diabetes mellitus tipo II, cardiopatía isquémica, insuficiencia cardíaca, enfermedad pulmonar crónica (enfisema/bronquitis crónica o EPOC, fibrosis/enfermedad pulmonar intersticial crónica o EPIC, apnea del sueño y tromboembolismo pulmonar crónico), asma, dislipemia, coagulopatías (secundarias al uso crónico de anticoagulantes, a hepatopatías crónicas o de origen primario/hereditario), insuficiencia renal

crónica, neoplasia activa, enfermedades inmunomediadas (inmunodeficiencias primarias, enfermedades autoinflamatorias y enfermedades autoinmunes), sobrepeso, obesidad, infección por virus de la inmunodeficiencia humana (VIH) y hepatitis B o C crónicas.

También se anotaron los datos relacionados con los tratamientos crónicos que tomaban los pacientes en el momento del ingreso hospitalario: antihipertensivos, betabloqueantes, diuréticos, antidiabéticos, antiagregantes, anticoagulantes, hipolipemiantes, quimioterapia, inmunosupresores/inmunomoduladores, antirretrovirales y antivíricos.

Por su relevancia tanto en la situación de emergencia sanitaria como en este estudio, se obtuvo la información vacunal de los pacientes, registrando cuántas dosis habían recibido de la vacuna contra la COVID-19 y qué vacuna se les había administrado, en caso de haber recibido alguna.

Se anotó si el paciente había tenido algún episodio previo de COVID-19, lo que implicaba que estaba experimentando una reinfección. En el caso de pacientes inmunocompetentes, se consideró reinfección un nuevo episodio si se daba la siguiente situación: una PCR positiva obtenida al menos 15 días después de la primera PCR positiva, junto con la aparición de síntomas nuevos compatibles con la COVID-19 que ocasionaran un ingreso hospitalario(175). En pacientes inmunocomprometidos (cáncer, tumores hematológicos, enfermedades autoinmunes, etc.), se consideró reinfección si se cumplían las siguientes condiciones: se obtuvo, o bien una PCR positiva al menos tres meses después de la primera, o bien una PCR positiva al menos 15 días después de la primera con niveles altos de ARN viral, y además la aparición de síntomas nuevos consistentes con la COVID-19 que requirieran ingreso hospitalario (176).

4.1.2.2. Variables relacionadas con el ingreso

En relación al ingreso hospitalario, se tomaron datos sobre el diagnóstico de COVID-19 confirmado por PCR y la fecha de este diagnóstico, así como la fecha de ingreso y de alta hospitalaria, de ingreso y alta en UCI, si aplicaba, o de éxitus.

Desde el punto de vista clínico, para valorar la carga de la enfermedad, se registraron las manifestaciones clínicas propias de la enfermedad en el momento del ingreso: disnea, dolor o presión torácica, tos, rinorrea, anosmia, ageusia, odinofagia, mialgias, fiebre, distermia, cefalea, náuseas o vómitos, diarrea, astenia, confusión, mareo/síncope y expectoración. Además, se recogieron las constantes vitales, incluyendo temperatura, presión arterial diastólica (PAD), presión arterial sistólica (PAS), frecuencia cardíaca (FC) y frecuencia respiratoria (FR).

Con el objetivo de evaluar el estado de gravedad del paciente, se utilizaron las escalas *scales for the assessment of severity in community-acquired pneumonia* (CURB-65), *Glasgow coma scale* (GCS), y las escalas *sequential organ failure assessment* (SOFA) y *acute physiology*

and chronic health disease classification system (APACHE) II en el momento de admisión en UCI en los casos que correspondiera.

A nivel analítico, las variables recogidas incluyeron las siguientes: leucocitos, neutrófilos, linfocitos, monocitos, eosinófilos, basófilos, eritrocitos, hemoglobina, hematocrito, VCM. (Volumen Corpuscular Medio), plaquetas, dímero D, actividad de protrombina (TP), INR. (Índice Internacional Normalizado), TTPA (tiempo de tromboplastina parcial activada), ratio TTPA, fibrinógeno derivado, sodio, potasio, cloruro, glucosa, urea, creatinina, tasa de filtrado glomerular estimada (TFGe, fórmula de CKD-EPI 2009), alanina aminotransferasa (ALT/GPT), aspartato aminotransferasa (AST/GOT), gamma-glutamyl transferasa (GGT), bilirrubina total, fosfatasa alcalina, lactato deshidrogenasa (LDH), fosfato, proteína C reactiva, procalcitonina, albúmina, pH, pCO₂, pO₂, HCO₃⁻, FiO₂, pO₂/FiO₂, Gradiente alveolo-arterial (Aa) de O₂ y lactato.

Se recogieron los datos de análisis de imagen por radiografía de tórax al ingreso de los pacientes, evaluando si existían o no opacidades y consolidaciones y cuántas zonas estaban afectadas.

En el grupo de 36 pacientes que ingresaron en la UCI durante la primera ola, adicionalmente, se evaluaron los niveles de las citoquinas IFN- γ , LIF y TNF- α tras quince días con VMI durante su ingreso en UCI, dada su importancia como indicadores de la severidad de la infección y el estado inflamatorio del paciente. Los niveles de estas citoquinas fueron analizados por ELISA (*enzyme-linked immunosorbent assay*) siguiendo los siguientes protocolos de la casa comercial Invitrogen (Thermo Fisher Scientific): Human IFN gamma ELISA Kit (BMS228), Human LIF ELISA Kit (BMS242) y Human TNF alpha ELISA Kit (BMS223-4). Para ello, las muestras se recolectaron, se almacenaron a -20 °C hasta su uso y se descongelaron antes de realizar el análisis. Todos los procedimientos, así como la preparación de reactivos y estándares, se realizaron de acuerdo con las instrucciones del fabricante proporcionadas en los manuales de los kits. Para estos análisis se tomaron como valores de referencia analítica los estipulados por los manuales de la casa comercial, siendo estos los indicados en la Tabla 1.

Tabla 1. Citoquinas estudiadas en los pacientes ingresados en UCI y los rangos de referencia analíticos y la sensibilidad analítica del correspondiente kit de ELISA utilizado.

Analito	Rango de referencia analítico	Sensibilidad analítica
IFN- γ	1,6-100 pg/mL	0,99 pg/mL
LIF	3,15-200 pg/mL	0,7 pg/mL
TNF- α	7,8-500 pg/mL	2,3 pg/mL

Por último, se registró si el diagnóstico inicial que se le dio al paciente en el momento de admisión hospitalaria incluía SDRA, insuficiencia respiratoria aguda o fallo multiorgánico.

4.1.2.3. Variables relacionadas con la estancia hospitalaria

En aquellos pacientes que requirieron soporte ventilatorio, se registró si necesitaron VMI, VMNI, CNAF, ECMO y/o sesiones en posición de prono. Además, se recogió información sobre las coinfecciones e infecciones nosocomiales que se dieron durante la estancia hospitalaria, así como el tipo de muestra y tipo de infección en caso de disponer de la información.

Se anotaron también todas aquellas complicaciones que surgieron en el periodo de hospitalización: insuficiencia renal aguda, shock séptico, complicaciones cardíacas (arritmias, insuficiencia cardíaca, paro cardíaco, miocarditis y cardiomiopatía de estrés), complicaciones gastrointestinales (aumento de transaminasas, íleo, isquemia mesentérica, colecistitis, pancreatitis y suboclusión), complicaciones neurológicas (delirio, encefalopatía, encefalitis, síndrome de Guillain-Barré, neuropatía periférica y miositis), coagulopatías (trombosis venosa profunda, tromboembolismo pulmonar, infarto agudo de miocardio, ictus, isquemia y sangrado) y distrés respiratorio.

Otro conjunto de datos recogido de forma detallada por su trascendencia dado el desconocimiento y el progreso de las investigaciones al respecto, fue el tratamiento específico para COVID-19 que se administró a los pacientes. Se documentaron los fármacos y tratamientos activos que el paciente recibió durante su estancia en el hospital, siendo estos: oxígeno suplementario, corticosteroides, remdesivir, lopinavir/ritonavir, hidroxicloroquina, tocilizumab, baricitinib, anakinra, interferón β -1 β (IFN β -1 β), ceftriaxona, azitromicina y heparina. En muchos casos, no fue posible obtener información sobre la dosis y los días que los pacientes recibieron cada tratamiento. Los tratamientos utilizados en cada periodo siguieron el protocolo del centro según regulaciones locales y nacionales, de acuerdo con las indicaciones de la OMS.

4.1.3. Aspectos éticos

El protocolo de estudio fue aprobado por la Comisión de Investigación y el Comité de Ética de Investigación Clínica del Hospital Universitario Río Hortega de Valladolid.

Se han seguido las directrices nacionales e internacionales, el código deontológico médico y la Declaración de Helsinki, y los datos han sido tratados de forma confidencial y anónima, en cumplimiento de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

La información del estudio fue centralizada en una base de datos segura, controlada y custodiada por la Universidad de Castilla-La Mancha.

4.2. Metodología

4.2.1. Modelos de aprendizaje automático (*machine learning*)

En este estudio se han utilizado diferentes modelos de ML para obtener los mejores resultados en cuanto a la identificación de predictores de mortalidad en las diferentes olas. Para ello, tras estudiar los parámetros a analizar, se realizaron análisis comparativos con diferentes modelos de ML ampliamente reconocidos en la comunidad científica. Posteriormente, se seleccionaron los modelos de aprendizaje que mejor se adaptaban al contexto clínico de la COVID-19 y se valoró qué algoritmos se ajustaban mejor a las características del proyecto.

Tras la obtención y preparación del conjunto de datos, se eliminaron los datos perdidos, se categorizaron los valores de diferentes variables y se normalizaron los valores para que diferentes conjuntos fueran comparables de forma coherente.

A continuación, se describen los algoritmos de ML que se utilizaron para el análisis en este estudio.

4.2.1.1. *Decision tree (DT)*

Los árboles de decisión o *decision tree (DT)* son modelos de aprendizaje supervisado ampliamente utilizados tanto en clasificación como en regresión. Su estructura se asemeja a un árbol, donde cada nodo representa una condición sobre un atributo, las ramas corresponden a las posibles respuestas, y los nodos hoja contienen las predicciones finales. La construcción de un árbol de decisión implica seleccionar iterativamente los atributos que mejor dividen los datos, utilizando criterios como la ganancia de información basada en la entropía o el índice de Gini, una métrica utilizada para evaluar la impureza o heterogeneidad de los nodos en un árbol de decisión. Cuanto menor es su valor, más homogéneos son los datos dentro de un nodo.

La profundidad del árbol es un factor crítico que influye en la capacidad de generalización del modelo. Un árbol demasiado profundo puede memorizar los datos de entrenamiento, generando sobreajuste (*overfitting*), mientras que un árbol demasiado superficial puede no capturar patrones significativos, llevando a un subajuste (*underfitting*). Para evitar estos problemas, se emplean técnicas como la poda (*pruning*), que reduce la complejidad del árbol eliminando nodos irrelevantes, y la validación cruzada para evaluar su rendimiento.

Los árboles de decisión tienen varias ventajas, como su facilidad de interpretación, la capacidad para manejar tanto datos numéricos como categóricos y la ausencia de necesidad de normalización. Sin embargo, también presentan desventajas, como su susceptibilidad a datos ruidosos y la tendencia a crear modelos inestables cuando se entrenan en diferentes subconjuntos

de datos. Son comúnmente utilizados en aplicaciones de diagnóstico médico, predicción de fallos en sistemas industriales y sistemas de recomendación (177,178).

- **Configuración del modelo DT:**

Dado un conjunto de datos de entrenamiento con n muestras y m características, representado como:

$$D = \{(x_i, y_i)\}_i^n = 1 \quad [5]$$

donde x_i es el vector de características e y_i la variable objetivo, el modelo final del modelo DT se define como una estructura de árbol de decisión que divide iterativamente el espacio en regiones homogéneas.

Para construir el modelo, el árbol de decisión sigue un proceso de partición recursiva en el que cada nodo se divide en dos hijos basándose en una característica y un valor de umbral que minimiza una función de pérdida. La función de pérdida comúnmente utilizada es la entropía o el índice de Gini para problemas de clasificación, y el error cuadrático medio para problemas de regresión.

La función objetivo se define como:

$$L(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) \quad [6]$$

donde $l(y_i, \hat{y}_i)$ es la función de pérdida que mide la discrepancia entre la variable objetivo y_i y la predicción del modelo $\hat{y}_i$.

Para controlar la complejidad del modelo y evitar el sobreajuste, se utilizan varios hiperparámetros, tales como:

- Profundidad máxima del árbol (d): limita el número de niveles en el árbol.
- Número mínimo de muestras por nodo: define el número mínimo de muestras necesarias para realizar una división.
- Número mínimo de muestras en una hoja: define el número mínimo de muestras que debe tener una hoja, la cual representa una predicción final del modelo.

El proceso de construcción del árbol se realiza de manera iterativa, donde en cada iteración se selecciona la mejor división basada en la reducción de la función de pérdida. La selección de hiperparámetros se puede realizar mediante técnicas como *grid search* y *Bayesian optimization*, buscando mejorar la precisión y evitar el sobreajuste.

4.2.1.2. **Gaussian Naïve Bayes (GNB)**

El clasificador *Naïve Bayes* es un modelo probabilístico basado en el teorema de Bayes, que asume la independencia condicional entre las características. La variante *Gaussian Naïve Bayes* (GNB) es un caso particular en el que se asume que las características siguen una distribución gaussiana (normal) dentro de cada clase.

Este algoritmo calcula la probabilidad condicional de que una observación pertenezca a una clase dada la combinación de sus características. La clasificación se basa en la fórmula del teorema de Bayes, donde se estima la probabilidad de cada clase multiplicada por la probabilidad de que los datos observados pertenezcan a dicha clase, considerando una distribución normal de los datos.

Una de las principales ventajas de *Naïve Bayes* es su rapidez y eficiencia, ya que no requiere entrenamientos computacionalmente costosos. También es robusto con conjuntos de datos pequeños y proporciona una interpretación probabilística clara. No obstante, su rendimiento puede verse afectado si las características no son realmente independientes, lo que ocurre en muchos problemas del mundo real. Se utiliza ampliamente en la detección de *spam*, clasificación de documentos y diagnóstico médico basado en probabilidades (177,179).

Concretamente, el modelo GNB es una herramienta poderosa en el análisis de datos médicos. Este modelo tiene varias ventajas en este sentido, como su simplicidad y eficiencia, lo que es crucial cuando se trabaja con grandes conjuntos de datos. Tiene la peculiaridad de, además de manejar datos faltantes de manera efectiva, lo cual es común en estudios médicos, funciona bien incluso con conjuntos de datos relativamente pequeños, proporcionando resultados robustos en cualquiera de los casos (179).

GNB es especialmente útil cuando se necesita una rápida clasificación y cuando las características siguen una distribución normal. Otros modelos como BLDA pueden no ofrecer ventajas significativas en términos de velocidad y simplicidad comparados con GNB (180,181).

– **Configuración del modelo GNB:**

Dado un conjunto de datos de entrenamiento con n muestras y m características, representado como se definió en la Ecuación 5, el modelo final GNB se define como:

$$P(x|y) \propto P(y) \prod_{j=1}^m P(x_j|y) \quad [7]$$

Para optimizar el modelo, GNB asume que las características son independientes entre sí, lo que simplifica el cálculo de las probabilidades condicionales. La probabilidad condicional $P(x_j|y)$ se modela como una distribución gaussiana:

$$P(x_j|y) = \frac{1}{\sqrt{2\pi\sigma_{y,j}^2}} \exp\left(-\frac{(x_j - \mu_{y,j})^2}{2\sigma_{y,j}^2}\right) \quad [8]$$

donde $\mu_{y,j}$ y $\sigma_{y,j}^2$ son la media y la varianza de la característica x_j para la clase y .

Para construir el modelo, GNB sigue un proceso en el que se estiman los parámetros $\mu_{y,j}$ y $\sigma_{y,j}^2$ a partir del conjunto de datos de entrenamiento. En cada iteración, se calcula la probabilidad de cada clase y dado el vector de características x y se selecciona la clase con la mayor probabilidad como la predicción final:

$$\hat{y} = \arg \max_y P(y) \prod_{j=1}^m P(x_j|y) \quad [9]$$

Para ajustar los hiperparámetros del modelo, se pueden utilizar técnicas de *grid search* y *Bayesian optimization*, buscando mejorar la precisión y evitar el sobreajuste. Aunque los hiperparámetros ajustados en GNB son la estimación de la media ($\mu_{y,j}$) y la estimación de la varianza ($\sigma_{y,j}^2$), el principal hiperparámetro ajustado en GNB es la varianza de suavización o suavización de la varianza ($\sigma_{y,j}^2 + \text{var_smoothing}$). Este hiperparámetro se utiliza en el modelo GNB para evitar problemas con varianzas muy pequeñas en los datos. En términos simples, añade una pequeña cantidad a las varianzas calculadas para cada característica, lo que ayuda a estabilizar los cálculos y prevenir divisiones por cero o valores extremadamente pequeños que podrían causar inestabilidad en el modelo.

4.2.1.3. **Bayesian linear discriminant analysis (BLDA)**

El Análisis Discriminante Lineal (LDA, por sus siglas en inglés) es una técnica estadística utilizada para clasificación y reducción de dimensionalidad. La variante *Bayesian Linear Discriminant Analysis* (BLDA) extiende este enfoque al introducir suposiciones bayesianas, lo que permite una estimación más robusta de los parámetros en situaciones donde los datos son escasos o presentan alta variabilidad.

BLDA asume que los datos de cada clase siguen una distribución normal multivariante y que la varianza puede diferir entre clases. Emplea métodos de inferencia bayesiana para ajustar los parámetros del modelo, combinando información previa con la evidencia observada en los datos. Esto le permite manejar la incertidumbre de manera más efectiva que el LDA tradicional y evitar el sobreajuste en conjuntos de datos pequeños.

Este algoritmo es particularmente útil en problemas donde es crucial capturar la incertidumbre en la clasificación, como en el reconocimiento de patrones y el análisis de señales biométricas. Aunque su rendimiento depende de la suposición de normalidad de los datos, su capacidad para mejorar la generalización en escenarios con información limitada lo hace valioso

en diversas aplicaciones científicas y médicas. En el caso de un análisis de datos de pacientes en diferentes olas de la pandemia, donde los patrones de los datos pueden cambiar drásticamente, su capacidad para manejar la variabilidad y la incertidumbre es crucial (177,182).

– **Configuración del modelo BLDA:**

Dado un conjunto de datos de entrenamiento con n muestras y m características, representado como se definió en la Ecuación 5, el modelo final en BLDA se define como una combinación lineal de las características con un enfoque bayesiano para la clasificación.

Para construir el modelo, BLDA sigue un proceso en el que se asume que las características de cada clase siguen una distribución normal multivariante. El objetivo es encontrar la combinación lineal de características que maximiza la separación entre las clases. La función de discriminante lineal se define como:

$$\delta_k(x) = x^T \sum^{-1} \mu_k - \frac{1}{2} \mu_k^T \sum^{-1} \mu_k + \log \pi_k \quad [10]$$

donde μ_k es el vector de medias de la clase k , $\sum$ es la matriz covarianza compartida, y π_k es la probabilidad a priori de la clase k .

La función objetivo se define como se indicó en la Ecuación 6. Para controlar la complejidad del modelo y evitar el sobreajuste, se utilizan varios hiperparámetros, tales como:

- Media a priori (μ_0): define la media a priori de las distribuciones de las clases.
- Covarianza a priori ($\sum_0$): define la matriz de covarianza a priori.
- Número de iteraciones (n_{iter}): define el número de iteraciones para la convergencia del algoritmo.

El proceso de construcción del modelo se realiza de manera iterativa, donde en cada iteración se actualizan los parámetros del modelo basándose en las distribuciones a priori y los datos observados. La selección de hiperparámetros se puede realizar mediante técnicas como *grid search* y *Bayesian optimization*, buscando mejorar la precisión y evitar el sobreajuste.

4.2.1.4. *K-nearest neighbors (KNN)*

K-nearest neighbors (KNN) es un algoritmo de aprendizaje supervisado basado en instancias que se emplea en tareas de clasificación y regresión. Su principio fundamental es que las observaciones con características similares tienden a agruparse en el espacio, por lo que la clasificación de un nuevo punto se basa en la mayoría de sus k vecinos más cercanos, es decir, los puntos más cercanos con características similares.

El proceso de clasificación de KNN implica calcular la distancia entre la nueva observación y todas las muestras del conjunto de entrenamiento. Luego, se seleccionan los k vecinos más cercanos y se asigna la clase más frecuente entre ellos.

Uno de los principales desafíos de KNN es la elección del valor de k . Un valor pequeño puede hacer que el modelo sea más sensible a ruidos y anomalías en los datos, mientras que un valor grande puede hacer que el modelo sea más general, pero menos sensible a las variaciones locales. Además, su alto costo computacional en la fase de predicción limita su aplicación en conjuntos de datos de gran tamaño. Sin embargo, sigue siendo un algoritmo muy utilizado en reconocimiento facial, sistemas de recomendación y análisis de datos geoespaciales (177,183).

Su capacidad para clasificar datos basándose en la similitud de características es particularmente útil cuando se analizan patrones en diferentes conjuntos de datos, como pueden ser los datos de las diferentes olas de la pandemia. De hecho, como KNN puede manejar datos faltantes de manera efectiva y adaptarse a cambios en los patrones de datos a lo largo del tiempo, algunos estudios han demostrado que este modelo puede ser altamente efectivo en la predicción de enfermedades y en la clasificación de datos médicos, proporcionando resultados precisos y fiables incluso en escenarios complejos (183).

- **Configuración del modelo KNN:**

Dado un conjunto de datos de entrenamiento con n muestras y m características, representado como se definió en la Ecuación 5, el modelo KNN se basa en encontrar los k vecinos más cercanos a una muestra dada para hacer predicciones.

Para construir el modelo, KNN sigue un proceso en el que se calcula la distancia entre la muestra de interés y todas las muestras del conjunto de datos. Las distancias más comunes utilizadas son la distancia euclidiana, la distancia de Manhattan y la distancia de Minkowski. La predicción se realiza tomando la mayoría de las clases (para clasificación) o el promedio de los valores (para regresión) de los k vecinos más cercanos.

La función objetivo se define como se indicó en la Ecuación 6. Para controlar la complejidad del modelo y evitar el sobreajuste, se utilizan varios hiperparámetros, tales como:

- Número de vecinos (k): define cuántos puntos vecinos se consideran para hacer la predicción.
- Métrica de distancia: define cómo se mide la distancia entre las muestras.
- Peso de los vecinos: define si todos los vecinos contribuyen por igual a la predicción o si se ponderan según su distancia.

El proceso de construcción del modelo se realiza de manera iterativa, donde en cada iteración se selecciona el mejor valor de k y la métrica de distancia basándose en la validación cruzada. La selección de hiperparámetros se puede realizar mediante técnicas como *grid search* y *Bayesian optimization*, buscando mejorar la precisión y evitar el sobreajuste.

4.2.1.5. **Support vector machine (SVM)**

Support Vector Machine (SVM) es un algoritmo de clasificación supervisado que busca encontrar el hiperplano óptimo que separa distintas clases en un espacio de características. Su objetivo es maximizar el margen entre las clases, lo que reduce el riesgo de error en nuevas predicciones.

Cuando los datos no son linealmente separables, SVM emplea el truco del *kernel* para transformarlos en un espacio de mayor dimensión donde sí puedan ser separados. Los kernels más utilizados incluyen el kernel lineal, polinómico, sigmoide y de base radial (RBF, por sus siglas en inglés), que permiten modelar relaciones más complejas en los datos.

SVM es altamente efectivo en problemas de clasificación con datos de alta dimensión y es utilizado en reconocimiento de imágenes, clasificación de texto y bioinformática. Sin embargo, su entrenamiento puede ser computacionalmente costoso, especialmente cuando se trabaja con grandes volúmenes de datos (177,184). Aun así, investigaciones recientes han demostrado que SVM puede mejorar la precisión y la robustez de los modelos predictivos en el análisis de datos médicos, proporcionando una herramienta valiosa para la toma de decisiones clínicas (185,186).

- **Configuración del modelo SVM:**

Dado un conjunto de datos de entrenamiento con n muestras y m características, representado como se definió en la Ecuación 5, el modelo SVM se basa en encontrar el hiperplano que mejor separa las diferentes clases en el espacio.

Para construir el modelo, SVM sigue un proceso en el que se transforman los datos en un espacio de características de alta dimensionalidad utilizando una función kernel.

La función de discriminante se define como:

$$f(x) = w^T x + b \quad [11]$$

donde w es el vector de pesos y b es el sesgo. El hiperplano se define por los vectores de soporte, que son los puntos de datos más cercanos al hiperplano.

La función objetivo se define como:

$$L(\theta) = \frac{1}{2} \|w^2\| + C \sum_{i=1}^n l(y_i, f(x_i)) \quad [12]$$

donde $l(y_i, f(x_i))$ es la función de pérdida que mide la discrepancia entre la variable objetivo y_i y la predicción del modelo $f(x_i)$, C es un hiperparámetro que controla el *trade-off* entre maximizar la margen y minimizar el error de clasificación.

Para controlar la complejidad del modelo y evitar el sobreajuste, se utilizan varios hiperparámetros, tales como:

- Tipo de *kernel*: define la función *kernel* utilizada para transformar los datos.
- Parámetro de regularización (C): controla el equilibrio (*trade-off*) entre la complejidad del modelo y el error de clasificación.
- Coeficiente del kernel (γ): define el coeficiente para los *kernels* polinómicos y RBF.

El proceso de construcción del modelo se realiza de manera iterativa, donde en cada iteración se optimizan los parámetros del modelo basándose en la maximización del margen y la minimización del error de clasificación. La selección de hiperparámetros se puede realizar mediante técnicas como *grid search* y *Bayesian optimization*, buscando mejorar la precisión y evitar el sobreajuste.

4.2.1.6. **Extreme Gradient Boosting (XGB)**

Extreme Gradient Boosting (XGBoost o XGB) es una implementación avanzada del algoritmo de *Gradient Boosting*, optimizada para maximizar la eficiencia y el rendimiento en tareas de aprendizaje supervisado. Su éxito radica en su capacidad para construir modelos altamente precisos a partir de la combinación de múltiples árboles de decisión débiles, donde cada árbol intenta corregir los errores del anterior.

Debido a su alto rendimiento, XGB es ampliamente utilizado en ciencia de datos y en aplicaciones reales que requieren alta precisión, como detección de fraudes, predicción de series temporales y análisis de datos estructurados en sectores como la banca y la medicina (187,188).

XGB ofrece ventajas significativas como su alta velocidad de ejecución, escalabilidad y capacidad para manejar datos faltantes. Además, incorpora técnicas de regularización L1 (*lasso*) y L2 (*ridge*) para reducir el sobreajuste y mejorar la generalización del modelo, lo que lo hace ideal para una amplia gama de problemas científicos, superando en muchos aspectos a otros algoritmos (188). Esto es especialmente importante en estudios médicos, donde la generalización del modelo es fundamental para su aplicación en diferentes cohortes de pacientes. Además, su optimización del uso de memoria y paralelización permite entrenar modelos de manera rápida, incluso en grandes volúmenes de datos.

La capacidad de XGB para realizar una búsqueda exhaustiva de hiperparámetros y su robustez frente a datos ruidosos lo hacen ideal para identificar patrones complejos en los predictores de mortalidad. Otros modelos, como KNN y SVM, pueden ser menos eficientes en grandes conjuntos de datos, mientras que BLDA puede no manejar bien la no linealidad presente en los datos y DT puede ser más propenso al sobreajuste (189).

- **Configuración del modelo XGB:**

Dado un conjunto de datos de entrenamiento con n muestras y m características, representado como se definió en la Ecuación 5, el modelo final en XGB se define como:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad [13]$$

donde f_k pertenece al espacio de los árboles de regresión y K es el número total de árboles.

Para optimizar el modelo, XGB minimiza una función de pérdida con regularización, lo que ayuda a evitar el sobreajuste y mejora la generalización. La función objetivo se define como:

$$L(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad [14]$$

El término de regularización $\Omega(f_k)$ controla la complejidad del modelo y se define como:

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad [15]$$

donde T es el número de hojas en el árbol y γ y λ son hiperparámetros de regularización.

Para construir el modelo, XGB sigue un proceso iterativo en el que cada nuevo árbol intenta minimizar el error del modelo anterior mediante una expansión de segundo orden de la función de pérdida. En cada iteración t , se optimiza un nuevo árbol f_t con:

$$L^{(t)} \approx \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \quad [16]$$

donde g_i es el gradiente de la función de pérdida y h_i la hessiana o segunda derivada, que proporciona información sobre la curvatura de la función.

Para ajustar los hiperparámetros del modelo, se utilizaron técnicas de *grid search* y *Bayesian optimization*, buscando mejorar la precisión y evitar el sobreajuste. Los principales hiperparámetros ajustados fueron:

- Número de árboles (K)
- Profundidad máxima de árbol (d)

- Tasa de aprendizaje (η)
- Gamma (γ)
- Lambda (λ) y Alpha (α)

4.2.1.7. **Random Forest (RF)**

Random Forest (RF) es un algoritmo de *ensemble learning* basado en la combinación de múltiples árboles de decisión mediante la técnica de *bagging*. En este enfoque, cada árbol se entrena con una muestra aleatoria del conjunto de datos, lo que introduce diversidad y reduce la varianza del modelo, mejorando su capacidad de generalización.

Además, en cada nodo de los árboles se selecciona aleatoriamente un subconjunto de características en lugar de utilizar todas las variables, lo que minimiza la correlación entre los árboles y aumenta la robustez del modelo.

RF se utiliza en aplicaciones como la detección de fraudes, la predicción de enfermedades y el análisis financiero. Su capacidad para manejar datos ruidosos y proporcionar interpretabilidad mediante la importancia de características lo convierte en una herramienta valiosa en muchas disciplinas (190,191).

Una de las principales ventajas del modelo RF es la reducción del sobreajuste. Al utilizar múltiples árboles de decisión y promediar sus resultados, se reduce significativamente el riesgo de sobreajuste. La técnica de *bagging* ayuda a crear modelos más robustos al entrenar cada árbol con diferentes subconjuntos de datos. La robustez y estabilidad del modelo son otras ventajas clave. RF es menos susceptible a ser influenciado por valores atípicos en los datos y, al utilizar múltiples árboles, las predicciones son más estables y menos sensibles a cambios en los datos de entrenamiento. También puede manejar datos faltantes de manera efectiva, permitiendo realizar predicciones incluso cuando algunas variables no están disponibles.

Otra ventaja importante es su capacidad para manejar datos complejos. RF puede trabajar tanto con variables continuas como categóricas sin necesidad de preprocesamiento adicional. Además, es capaz de capturar interacciones complejas y no lineales entre las variables predictoras, lo cual es crucial en estudios médicos donde las relaciones entre variables pueden ser intrincadas. RF también proporciona una medida de importancia de las características, lo que permite identificar rápidamente las variables más relevantes para la predicción de mortalidad (192).

La capacidad de RF para generalizar bien a nuevos datos es crucial en estudios médicos, donde los modelos deben ser aplicables a diferentes poblaciones de pacientes. Además, RF es menos susceptible a la variabilidad en los datos y puede manejar bien los datos faltantes, ofreciendo un equilibrio entre precisión y eficiencia (190,191).

- **Configuración del modelo RF:**

En el algoritmo RF propuesto, un conjunto de árboles de decisión $\{T_1, T_2, \dots, T_m\}$ se construye con enfoque de agregación por *bagging*. Para construir cada árbol T_i , se realizan los siguientes pasos:

Dado un conjunto de datos de entrenamiento con n observaciones y p características semejante al de la Ecuación 5, se genera un subconjunto de datos D seleccionando n muestras aleatorias con reemplazo del conjunto de datos original. Esta técnica de muestreo permite que ciertos puntos de datos aparezcan múltiples veces en D , mientras que otros pueden no aparecer en absoluto.

En cada nodo de cada árbol, en lugar de evaluar todas las p características, se selecciona un subconjunto aleatorio de k características, donde $k = \sqrt{p}$. Esto reduce la correlación entre los árboles individuales, aumentando así la capacidad de generalización del modelo.

Cada nodo se divide según un criterio de reducción de impureza, que puede ser la entropía o el índice de Gini para la clasificación. En este estudio, se utiliza el índice de Gini. La impureza de Gini G de un nodo con una proporción p_k de elementos pertenecientes a la clase k se define como:

$$G = 1 - \sum_{k=1}^K p_k^2 \quad [17]$$

Una vez que los árboles están entrenados, la predicción del modelo RF se obtiene mediante agregación. Para un conjunto de árboles $\{T_1, T_2, \dots, T_m\}$, la predicción final $\hat{y}$ se calcula utilizando el voto mayoritario:

$$\hat{y} = \text{mode}\{T_1(x), T_2(x), \dots, T_m(x)\} \quad [18]$$

La importancia de cada característica se mide evaluando el cambio en el criterio de división cuando la característica se permutaba aleatoriamente en el conjunto de datos. Para la importancia basada en el índice de Gini, si permutar una característica específica aumenta la impureza de los nodos del árbol, esa característica se considera importante.

4.2.2. Configuración y entrenamiento de los modelos

Estos modelos se diseñaron utilizando los sistemas *MatLab Statistical* y *Machine learning Toolbox* (MatLab 2024a; The MathWorks, Natick, MA, USA).

Las técnicas de aprendizaje automático suelen involucrar uno o más hiperparámetros que permiten ajustar el algoritmo durante el proceso de entrenamiento. Los diversos valores asignados a estos hiperparámetros (*number of splits*, *learners*, *neighbors*, *distance metric*, *distant weight*,

kernel, *box constraint level*, *multiclass method*, etc.) resultan en algoritmos con diferentes rendimientos de predicción para lograr resultados óptimos. Para optimizar estos hiperparámetros en cada técnica de aprendizaje automático empleada en este estudio, se aplicó un enfoque de optimización bayesiana. La optimización bayesiana tiene como objetivo determinar la configuración de hiperparámetros que maximice el rendimiento del algoritmo basándose en evaluaciones previas, operando bajo la suposición de que existe una correlación entre los diversos hiperparámetros y el rendimiento logrado por el algoritmo.

Estos modelos de ML requieren de una fase de entrenamiento previa, mediante la cual se introducen diferentes conjuntos o vectores de entrada y se van ajustando los parámetros según un procedimiento determinado. El sistema modifica automáticamente los valores de tal manera que se ajustan para generar salidas correctas. Por este motivo, la base de datos se dividió en dos segmentos: los datos del 70 % de los pacientes para el entrenamiento del modelo y los datos del 30 % restante para el test, sin solapamiento de datos.

Para validar los resultados, es decir, para evaluar el sistema, limitar el posible sobreajuste del sistema, y comprobar que los resultados del análisis eran independientes de la clasificación utilizada (entrenamiento y test), se utilizó la técnica de *k-fold cross-validation* (177). Esta técnica es ampliamente utilizada en aprendizaje automático y análisis de datos, ya que permite obtener estimaciones más robustas del rendimiento del modelo al utilizar múltiples particiones del conjunto de datos.

El procedimiento consiste en dividir aleatoriamente el conjunto de datos en k subconjuntos o *folds* de tamaño similar. A lo largo de k iteraciones, cada uno de estos subconjuntos se utiliza una vez como conjunto de prueba (*test set*), mientras que los $k - 1$ subconjuntos restantes se emplean para entrenar el modelo. De este modo, cada observación del conjunto de datos tiene la oportunidad de ser utilizada tanto para entrenamiento como para validación, lo que mejora la generalización del modelo y minimiza el riesgo de sobreajuste.

Las fases descritas en este proceso se ilustran en la Figura 8. Inicialmente, el conjunto de datos se divide aleatoriamente en los dos subconjuntos mencionados. En cada iteración o ciclo, el subconjunto de entrenamiento permite estimar los parámetros del modelo, mientras que el subconjunto de prueba se emplea para evaluar su desempeño. Este procedimiento se repite hasta completar las k iteraciones, tras lo cual se calcula un promedio de las métricas de evaluación obtenidas en cada ciclo, proporcionando así una estimación más estable y fiable del rendimiento del sistema.

Este enfoque no solo ayuda a mitigar los efectos del sobreajuste, sino que también permite comparar de manera objetiva diferentes configuraciones del modelo, facilitando la selección de

hiperparámetros óptimos y garantizando que el desempeño obtenido no sea producto de una partición específica del conjunto de datos.

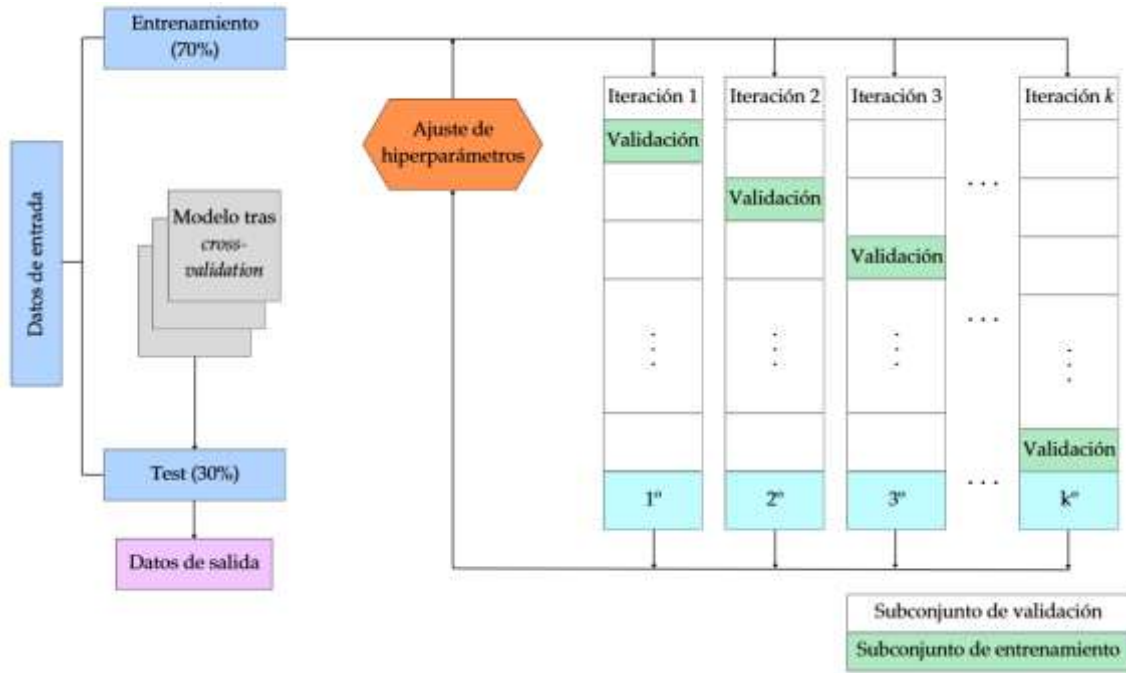


Figura 8. Esquema del mecanismo de entrenamiento y test de los modelos de aprendizaje automático.

4.2.3. Evaluación del rendimiento de los modelos

Para estudiar cuantitativamente el rendimiento de los diferentes modelos predictivos respecto al conjunto de datos se recurrió a métricas estándar conocidas de ML: coeficiente de correlación de Matthews (MCC), *F1-score*, *balanced accuracy* o exactitud equilibrada, *recall* o sensibilidad, precisión o valor predictivo positivo (VPP), especificidad, índice de Youden degenerado (DYI), índice kappa de Cohen (CKI) o simplemente kappa, curvas *receiver operating characteristic* (ROC) y el área bajo la curva (AUC, por sus siglas en inglés) (177).

El MCC obtiene una puntuación más alta cuanto mejores son los resultados obtenidos en las cuatro categorías de la matriz de confusión (verdaderos positivos, falsos negativos, verdaderos negativos y falsos positivos). Tiene un valor entre -1 y $+1$, siendo $MCC = -1$ el resultado de una contradicción total entre el resultado predicho y la verdad fundamental correspondiente; $MCC = 0$ el resultado de una incapacidad de predicción mejor del modelo que de manera aleatoria; y $MCC = +1$ el resultado de un clasificador que predice el resultado verdadero. Se define como:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FN)}} \quad [19]$$

donde FP representa el número de predicciones incorrectas, TP el número de predicciones correctas, TN los verdaderos negativos, FN los falsos negativos, P los positivos y N los negativos (177).

El *F1-score* es una escala que combina los valores de precisión y sensibilidad. Se describe de la siguiente manera:

$$F_1 \text{ score} = 2 \frac{\text{Precisión} \cdot \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}} \quad [20]$$

donde la precisión o VPP se calcula como:

$$\text{Precisión o VPP} = \frac{TP}{TP + FP} \quad [21]$$

y la sensibilidad o *recall* se calcula como:

$$\text{Sensibilidad o recall} = \frac{TP}{TP + FN} \quad [22]$$

La exactitud equilibrada o *balanced accuracy* es la media de la sensibilidad y la especificidad. Se calcula como:

$$\text{Balanced accuracy} = \frac{\text{Sensibilidad} + \text{Especificidad}}{2} \quad [23]$$

donde la sensibilidad queda definida en la Ecuación 22 y la especificidad se calcula como:

$$\text{Especificidad} = \frac{TN}{TN + FP} \quad [24]$$

El DYI es una medida estadística utilizada para evaluar la efectividad de una prueba diagnóstica. Se calcula como la máxima distancia vertical entre la curva de tasas de verdaderos positivos identificados como tales (sensibilidad) y la tasa de falsos positivos (1 - especificidad) en una curva ROC. Este parámetro permite elegir el punto de corte óptimo en la curva. Se define como:

$$DYI = \text{Sensibilidad} + \text{Especificidad} - 1 \quad [25]$$

donde la sensibilidad queda definida en la Ecuación 22 y la especificidad en la Ecuación 24.

El CKI o kappa es una medida estadística utilizada para evaluar el nivel de acuerdo entre dos evaluadores que clasifican ítems en categorías mutuamente excluyentes. Es especialmente útil en situaciones donde las decisiones son subjetivas y las categorías son nominales (es decir, no tienen un orden natural). Este índice intenta aprovechar la exactitud normalizándola a la probabilidad de que la clasificación coincida por azar. Se define como:

$$\kappa = \frac{P_0 - P_e}{1 - P_e} \quad [26]$$

donde P_0 es la proporción de acuerdo observado entre dos evaluadores, y P_e es la proporción de acuerdo esperado por azar.

El gráfico de las curvas ROC muestra el rendimiento del modelo clasificatorio en todos los umbrales de decisión, siendo una representación de la tasa de casos positivos correctamente identificados como tales (sensibilidad) frente a la tasa de falsos positivos (1 - especificidad). A partir de esta curva se obtiene la AUC calculada con el área bajo la curva ROC, pudiendo obtener un valor de 0,5 (idéntico a la clasificación aleatoria) a 1,0 (perfección en las predicciones). Esta AUC se define como:

$$AUC = \frac{1}{n_+n_-} \sum_{x^+} \sum_{x^-} I(f(x^+) > f(x^-)) \quad [27]$$

La AUC y la exactitud equilibrada sirvieron como métricas de rendimiento a maximizar.

Además, para problemas de regresión, se utilizó el error cuadrático medio (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad [28]$$

Debido a la naturaleza estocástica tanto del proceso de inicialización de la máquina como del aprendizaje automático en todas las simulaciones, se llevaron a cabo 100 repeticiones con el objetivo de garantizar la estabilidad de los resultados y proporcionar estimaciones más fiables de las métricas de rendimiento. Este enfoque permitió calcular no solo los valores medios, sino también la desviación estándar, lo que facilita una mejor interpretación de la variabilidad inherente en los resultados y la comparación objetiva entre diferentes modelos o configuraciones.

Para mitigar el impacto del ruido en los datos y reducir la influencia de fluctuaciones aleatorias, se implementó un esquema de replicación sistemática, asegurando que las variaciones observadas fueran consecuencia del comportamiento del modelo y no de factores externos. Además, este enfoque permitió obtener cálculos precisos de AUC y otras métricas clave, garantizando que los resultados fueran estadísticamente significativos y representativos del rendimiento real del sistema.

Para ello, se adoptó una estrategia de muestreo aleatorio uniforme, en la que cada repetición de la simulación seguía un procedimiento consistente, evitando sesgos en la selección de datos y asegurando una distribución equitativa de las condiciones iniciales. De esta manera, se optimizó la robustez del análisis y se mejoró la reproducibilidad de los resultados, lo que es fundamental para la validez del estudio.

5. Resultados

5.1. Perfil y evolución de los pacientes

5.1.1. Pacientes de la primera cohorte (1º ola)

5.1.1.1. Estado previo de los pacientes

En la cohorte de pacientes de la primera ola se han incluido 291 pacientes hospitalizados con diagnóstico de COVID-19 confirmado por PCR. De ellos, fallecieron 60 (20,62 %). Ninguno de estos pacientes había tenido un episodio de COVID-19 previo ni había recibido ninguna dosis de la vacuna frente a la enfermedad.

Como muestra la Tabla A1 (Anexo 8.2.1.), la media de edad de los pacientes fallecidos fue de 77,1 años, mientras que de los vivos fue de 64,5 años.

El resto de información relacionada con el estado previo de los pacientes se encuentra reflejada en la Tabla A2 (Anexo 8.2.1.). Entre estos pacientes hubo un 7,22 % más de hombres que de mujeres, falleció un 16,67 % más de hombres que de mujeres, y vivió un 3,92 % más de mujeres que de hombres.

Las comorbilidades previas más frecuentes entre los pacientes en general fueron la hipertensión (51,7 % de los fallecidos, 42 % de los supervivientes), la diabetes mellitus (18,3 % de los fallecidos, 19 % de los supervivientes), la cardiopatía isquémica (40 % de los fallecidos, 23 % de los supervivientes) y la dislipemia (26,7 % de los fallecidos, 88 % de los supervivientes). Entre los pacientes fallecidos, destacan, a mayores, la EPOC (11,7 % de los pacientes fallecidos), la insuficiencia renal crónica (11,7 %), las enfermedades inmunomediadas (10 %) y la obesidad (10 %). En el caso de los pacientes que superaron la enfermedad, destaca el asma con un 10 % del total de supervivientes.

El porcentaje de pacientes que tomaban tratamientos crónicos en el momento del ingreso fue de un 73,3 % entre los fallecidos, y de un 71 % entre los supervivientes.

5.1.1.2. Clínica de los pacientes al ingreso

Los resultados relacionados con la clínica de los pacientes al ingreso se reflejan en las Tablas A3 y A4 (Anexo 8.2.2.).

Los síntomas más frecuentes en los pacientes al momento del ingreso hospitalario fueron la disnea (13,3 % de los fallecidos, 12 % de los vivos), la tos (66,7 % de los fallecidos, 61 % de los vivos), las mialgias (16,7 % de los fallecidos, 12 % de los vivos), la fiebre (28,3 % de los fallecidos, 26 % de los vivos), las náuseas o vómitos (15 % de los fallecidos, 14 % de los vivos), la diarrea (16,7 % de los fallecidos, 18 % de los vivos), la astenia (28,3 % de los fallecidos, 38 % de los vivos) y la expectoración (13,3 % de los fallecidos, 12 % de los vivos). En el caso de los

pacientes fallecidos, destaca adicionalmente la confusión en el momento de la exploración física (10 % de los pacientes fallecidos). Tanto en pacientes supervivientes como fallecidos se evidencia la presencia de consolidaciones (más del 20 % de los pacientes) y opacidades (más del 50 % de los pacientes) en el análisis de la radiografía de tórax.

En los pacientes que fallecieron se observaron valores significativamente mayores de FC y FR respecto a los vivos.

El diagnóstico más frecuente en el momento del ingreso hospitalario en este grupo de la primera ola fue la insuficiencia respiratoria aguda: 74 % de los fallecidos y 34 % de los supervivientes, siendo esta diferencia estadísticamente significativa.

Las escalas CURB-65 y APACHE-II mostraron valores significativamente diferentes entre pacientes supervivientes y fallecidos.

En cuanto a la analítica de ingreso, cabe destacar diferencias significativas en los valores de pacientes fallecidos y pacientes que superaron la enfermedad en los niveles de neutrófilos, hemoglobina, glucosa, urea, TFGe, albúmina, LDH, proteína C reactiva, FiO₂, pO₂/FiO₂, Gradiente Aa de O₂ y lactato.

5.1.1.3. Complicaciones durante la estancia hospitalaria

Los datos relacionados con las complicaciones durante la estancia hospitalaria en este grupo de pacientes se reflejan en las Tablas A7 y A8 (Anexo 8.2.3.).

Los pacientes fallecidos permanecieron hospitalizados una media de 16,7 días, mientras que los supervivientes lo hicieron durante 18,3 días, no siendo esta una diferencia estadísticamente significativa.

Solamente el 10 % de los fallecidos ingresó en UCI, frente al 26 % de los supervivientes, siendo la diferencia estadísticamente significativa.

Los pacientes que superaron la enfermedad requirieron un uso notablemente menor de soporte ventilatorio frente a los que fallecieron, siendo las técnicas más utilizadas la VMI (43 % de los pacientes fallecidos y 17 % de los supervivientes) y las sesiones de pronó (10 % de los pacientes fallecidos y 2,2 % de los supervivientes). La elevada proporción de pacientes fallecidos que requirieron este soporte respiratorio adicional es significativamente mayor que la de los pacientes supervivientes que lo necesitaron.

Tanto los pacientes fallecidos como los supervivientes sufrieron infecciones nosocomiales, mayormente bacterianas, durante la estancia hospitalaria.

Las complicaciones durante el ingreso fueron más frecuentes en los pacientes fallecidos, destacando el fallo renal agudo (20 % de los pacientes fallecidos), las complicaciones gastrointestinales (10 % de los pacientes fallecidos), el shock séptico (10 % de los pacientes fallecidos) y el SDRA (30 % de los pacientes fallecidos), siendo este último también frecuente en los pacientes supervivientes afectando a un 15 % de los mismos.

El tratamiento utilizado para la COVID-19 en estos pacientes fue variado según la severidad de los mismos. Los tratamientos más utilizados fueron el oxígeno suplementario (87 % de los pacientes fallecidos, 55 % de los supervivientes), la hidroxicloroquina (43 % de los fallecidos, 76 % de los supervivientes), y la combinación de antibióticos de ceftriaxona (43 % de los fallecidos, 65 % de los supervivientes) y azitromicina (47 % de los fallecidos, 72 % de los supervivientes). Además, el 29 % de los pacientes que superaron la enfermedad recibieron tratamiento con lopinavir/ritonavir. Este fármaco no se empleó en ningún paciente fallecido.

Las causas de éxitus más frecuentes en este grupo fueron el fallo multiorgánico (48 %) y la insuficiencia respiratoria (28 %).

5.1.2. Pacientes de la segunda cohorte (2ª a 4ª olas)

5.1.2.1. Estado previo de los pacientes

En la cohorte de pacientes de las olas 2ª a 4ª, se han incluido 363 pacientes hospitalizados con diagnóstico de COVID-19 confirmado por PCR. De ellos, fallecieron 40 (11,0 %).

El 2,5 % de los pacientes fallecidos había presentado al menos un episodio anterior de COVID-19, frente a solo un 0,6 % de los pacientes que sobrevivieron. Además, el 4,7 % del total de pacientes había recibido al menos una dosis de la vacuna contra la enfermedad, siendo este porcentaje muy similar entre pacientes fallecidos y supervivientes. Por otra parte, ninguno de los pacientes fallecidos había recibido la pauta completa vacunación, mientras que el 2 % de los supervivientes sí la había recibido.

Como muestra la Tabla A1 (Anexo 8.2.1.), la media de edad de los pacientes fallecidos fue de 79,5 años, mientras que de los vivos fue de 64,4 años.

El resto de información relacionada con el estado previo de los pacientes se encuentra reflejada en la Tabla A2 (Anexo 8.2.1.). Entre estos pacientes hubo un 27,3 % más de hombres que de mujeres, falleció un 65 % más de hombres que de mujeres, y sobrevivió un 9 % más de mujeres que de hombres.

Las comorbilidades previas más frecuentes entre los pacientes en general fueron la hipertensión (60 % de los fallecidos, 45 % de los supervivientes), la diabetes mellitus (25 % de los fallecidos, 21 % de los supervivientes), la dislipemia (40 % de los fallecidos, 31 % de los

supervivientes) y las coagulopatías (17,5 % de los fallecidos, 41 % de los supervivientes). Entre los pacientes fallecidos, destacan, además de las anteriores, la insuficiencia renal crónica (12,5 % de los pacientes fallecidos) y la cardiopatía isquémica (17,5 %). En el caso de los pacientes que superaron la enfermedad, además de las comunes a ambos grupos, fueron frecuentes las enfermedades inmunomediadas (12 % de los pacientes supervivientes).

El porcentaje de pacientes que tomaban tratamientos crónicos en el momento de la hospitalización fue de un 80 % entre los fallecidos y de un 57 % entre los supervivientes, siendo esta diferencia estadísticamente significativa.

5.1.2.2. Clínica de los pacientes al ingreso

Los resultados relacionados con la clínica de los pacientes al ingreso se reflejan en las Tablas A3 y A5 (Anexo 8.2.2.).

Los síntomas más frecuentes en los pacientes al momento del ingreso hospitalario fueron la disnea (47,5 % de los fallecidos, 53 % de los vivos), la tos (52,5 % de los fallecidos, 63 % de los vivos), las mialgias (10 % de los fallecidos, 18 % de los vivos), la fiebre (65 % de los fallecidos, 70 % de los vivos), las náuseas o vómitos (12,5 % de los fallecidos, 15 % de los vivos), la astenia (17,5 % de los fallecidos, 31 % de los vivos) y la expectoración (15 % de los fallecidos, 14 % de los vivos). En el caso de los pacientes fallecidos, destaca adicionalmente la confusión en el momento de la exploración física (22,5 % de los pacientes fallecidos). En los pacientes que superaron la enfermedad destacan, además de los comunes, la presión torácica (14 % de los pacientes supervivientes), la pérdida de gusto y de olfato (15 %) y las cefaleas (10 % de los pacientes supervivientes). Tanto en pacientes supervivientes como en fallecidos se evidencia la presencia de consolidaciones (más del 16 % de los pacientes) y opacidades (el 67,5 % de los pacientes fallecidos y el 83 % de los vivos) en el análisis de la radiografía de tórax.

El diagnóstico más frecuente en el momento del ingreso hospitalario fue la insuficiencia respiratoria aguda: 50 % de los fallecidos y 39 % de los supervivientes, siendo esta diferencia estadísticamente significativa. Por otro lado, el 15 % de los fallecidos fue diagnosticado con SDRA al ingreso.

La escala SOFA muestra diferencias significativas entre los valores obtenidos para pacientes que fallecieron en comparación con los que superaron la enfermedad.

En cuanto a la analítica de ingreso, cabe destacar diferencias significativas en los valores de pacientes fallecidos y pacientes que superaron la enfermedad en los niveles de hemáties, VCM, plaquetas, urea, TFG_e, creatinina, albúmina, LDH y proteína C reactiva.

5.1.2.3. Complicaciones durante la estancia hospitalaria

Los datos relacionados con las complicaciones durante la estancia hospitalaria en este grupo de pacientes se reflejan en las Tablas A7 y A8 (Anexo 8.2.3.).

Los pacientes fallecidos permanecieron una media de 20,6 días hospitalizados, mientras que los que superaron la enfermedad estuvieron ingresados una media de 14,9 días, siendo esta una diferencia estadísticamente significativa.

El 53 % de los pacientes fallecidos ingresó en UCI, frente al 19 % de los supervivientes. Esta diferencia fue estadísticamente significativa.

La proporción de pacientes que superaron la enfermedad que necesitaron soporte ventilatorio fue menor que la de los que fallecieron, siendo la técnica más utilizada la VMI (48 % de los pacientes fallecidos frente al 17 % de los pacientes que sobrevivieron). En los pacientes que fallecieron, destaca el uso de otras técnicas de soporte ventilatorio que se utilizaron en menor medida en los pacientes supervivientes, como la CNAF (10 % de los pacientes fallecidos), la VMNI (30 % de los fallecidos) o las sesiones de prono (30 % de los fallecidos).

Tanto los pacientes fallecidos como los supervivientes sufrieron infecciones nosocomiales, mayormente bacterianas, durante la estancia hospitalaria. Sin embargo, la proporción de pacientes fallecidos que las padecieron (28 % del total) es considerablemente mayor que en el caso de pacientes supervivientes (9 % de los pacientes supervivientes).

Las complicaciones durante el ingreso fueron más frecuentes en los pacientes fallecidos, destacando el fallo renal agudo (33 % de los pacientes fallecidos), las complicaciones cardíacas (23 % de los pacientes fallecidos) y gastrointestinales (13 % de los pacientes fallecidos), las coagulopatías (23 % de los pacientes fallecidos), el shock séptico (13 % de los pacientes fallecidos) y el SDRA (38 % de los pacientes fallecidos), siendo este último frecuente también entre los supervivientes, aunque en menor medida (11 %).

El tratamiento utilizado para la COVID-19 en estos pacientes varió según la gravedad. Los tratamientos más utilizados fueron el oxígeno suplementario (80 % de los pacientes fallecidos, 54 % de los supervivientes), los corticoides (58 % de los pacientes fallecidos, 40 % de los supervivientes), la heparina como anticoagulante (20 % de los pacientes fallecidos, 35 % de los supervivientes) y los antibióticos ceftriaxona (53 % de los fallecidos, 44 % de los supervivientes) y azitromicina (38 % de los fallecidos, 34 % de los supervivientes). El tratamiento con lopinavir/ritonavir y/o con hidroxiclороquina dejó de utilizarse por completo tras la primera ola.

Las causas más frecuentes de éxitus en este grupo fueron la insuficiencia respiratoria (28 %) y el fallo multiorgánico (43 %).

5.1.3. Pacientes de la tercera cohorte (5ª y 6ª olas)

5.1.3.1. Estado previo de los pacientes

En la cohorte de pacientes de las olas 5ª y 6ª se han incluido 244 pacientes hospitalizados con diagnóstico de COVID-19 confirmado por PCR. De ellos, fallecieron 27 (11,1 %).

El 3,7 % de los pacientes fallecidos había tenido al menos un episodio previo de COVID-19, frente a solo un 0,9 % de los pacientes que sobrevivieron. Además, el 44,4 % de los pacientes fallecidos había recibido la pauta completa de vacunación, frente al 36 % de los supervivientes.

Como muestra la Tabla A1 (Anexo 8.2.1.), la media de edad de los pacientes fallecidos fue de 82 años, mientras que de los vivos fue de 61 años.

El resto de información relacionada con el estado previo de los pacientes se encuentra reflejada en la Tabla A2 (Anexo 8.2.1.). Entre estos pacientes hubo un 18,8 % más de hombres que de mujeres. El porcentaje de hombres fallecidos es un 33,3 % mayor que el de mujeres, mientras que superaron la enfermedad un 3,3 % más de mujeres que de hombres.

Las comorbilidades previas más frecuentes entre los pacientes en general fueron la hipertensión (48,1 % de los fallecidos, 43 % de los supervivientes), la diabetes mellitus (11,1 % de los fallecidos, 21 % de los supervivientes), la dislipemia (40,7 % de los fallecidos, 24 % de los supervivientes), las coagulopatías (37 % de los fallecidos, 12 % de los supervivientes) y las neoplasias activas (25,9 % de los fallecidos, 13 % de los supervivientes). Entre los pacientes fallecidos, destacan, a mayores, la cardiopatía isquémica (18,5 % de los fallecidos), la EPOC (26,9 % de los pacientes fallecidos), la insuficiencia renal crónica (22,2 % de los fallecidos), y las enfermedades inmunomediadas (18,5 % de los fallecidos).

El porcentaje de pacientes que tomaban tratamientos crónicos en el momento del ingreso fue de un 88,9 % entre los fallecidos, y de un 54 % entre los supervivientes. Esta diferencia fue estadísticamente significativa.

5.1.3.2. Clínica de los pacientes al ingreso

Los resultados relacionados con la clínica de los pacientes al ingreso se reflejan en las Tablas A3 y A6 (Anexo 8.2.2.).

Los síntomas más frecuentes en los pacientes al momento del ingreso hospitalario fueron la disnea (74,1 % de los fallecidos, 49 % de los supervivientes), la tos (44,4 % de los fallecidos, 59 % de los supervivientes), la fiebre (48,1 % de los fallecidos, 64 % de los supervivientes), la astenia (11,1 % de los fallecidos, 29 % de los supervivientes) y la expectoración (18,5 % de los fallecidos, 15 % de los supervivientes). En el caso de los pacientes fallecidos, destaca

adicionalmente la confusión en el momento de la exploración física (18,5 % de los pacientes fallecidos). Por otro lado, entre los pacientes que superaron la enfermedad fueron frecuentes la presión torácica (18 % de los supervivientes), la pérdida de gusto (13 %) y olfato (15 %), las cefaleas (10 %), las náuseas o vómitos (12 %) y la diarrea (18 %). Tanto en pacientes supervivientes como fallecidos se evidencia la presencia de consolidaciones (más del 20 % de los pacientes) y opacidades (más del 60 % de los pacientes) en el análisis de la radiografía de tórax.

El diagnóstico más frecuente en el momento del ingreso hospitalario en este grupo de las olas 5ª y 6ª fue la insuficiencia respiratoria aguda: 35 % de los fallecidos y 85,2 % de supervivientes, siendo esta diferencia estadísticamente significativa.

Las escalas GCS, CURB-65, así como la FC y la FR, muestran diferencias significativas entre los valores obtenidos para pacientes que fallecieron en comparación con los que superaron la enfermedad.

En cuanto a la analítica de ingreso, cabe destacar que se encontraron diferencias significativas entre los valores de pacientes fallecidos y los de pacientes que superaron la enfermedad en los niveles de eosinófilos, VCM, urea, creatinina, TFG_e, albúmina y proteína C reactiva.

5.1.3.3. Complicaciones durante la estancia hospitalaria

Los datos relacionados con las complicaciones durante la estancia hospitalaria en este grupo de pacientes se reflejan en las Tablas A7 y A8 (Anexo 8.2.3.).

Los pacientes fallecidos permanecieron una media de 10,3 días hospitalizados, mientras que los que superaron la enfermedad estuvieron ingresados una media de 23,2 días, siendo esta una diferencia estadísticamente significativa.

El 11 % de los fallecidos ingresó en UCI, frente al 9 % de los supervivientes.

El 10 % de los supervivientes requirió soporte ventilatorio por VMI, mientras que, entre los fallecidos, este porcentaje fue de un 7 %. Por otra parte, la técnica de soporte ventilatorio por CNAF fue más común entre los pacientes fallecidos (un 22 % de los mismos) que en los supervivientes (un 6 % de los mismos).

Tanto los pacientes fallecidos como los supervivientes sufrieron infecciones nosocomiales, mayormente bacterianas, durante la estancia hospitalaria. Sin embargo, en esta cohorte el porcentaje de pacientes que mostraron este tipo de infecciones fue bajo en general, no superando el 10 % en ninguno de los dos grupos.

Las complicaciones durante el ingreso fueron más frecuentes en los pacientes fallecidos, destacando el fallo renal agudo (22 %) y las coagulopatías (11 %). La proporción de pacientes

supervivientes que padecieron complicaciones durante la estancia hospitalaria fue menor que en las olas anteriores, con una prevalencia inferior al 7 % en este grupo.

Las estrategias terapéuticas utilizadas para tratar la COVID-19 en estos pacientes varió según la gravedad del cuadro clínico. Los tratamientos más utilizados fueron el oxígeno suplementario (89 % de los pacientes fallecidos, 54 % de los supervivientes), la dexametasona como corticoide (48 % de los pacientes en general), la heparina como anticoagulante (22 % de los pacientes fallecidos, 47 % de los supervivientes) y los antibióticos ceftriaxona (41 % de los fallecidos, 28 % de los supervivientes) y azitromicina (11 % de los fallecidos, 15 % de los supervivientes).

La causa más frecuente de éxitos en este grupo fue la insuficiencia respiratoria (85 %).

5.2. Identificación de predictores de mortalidad

5.2.1. Comparación y selección de modelos de ML

Se han utilizado diferentes métodos para identificar los predictores de mortalidad en estos pacientes. El objetivo de comparar varios algoritmos es utilizar de forma definitiva para cada cohorte aquel que proporcione los mejores resultados en términos de capacidad predictiva.

5.2.1.1. Primera cohorte (1º ola)

Los resultados obtenidos para las métricas de rendimiento en cada modelo (*balanced accuracy* (%), precisión, MCC, F1 Score, AUC, sensibilidad, especificidad, *kappa* y DYI) para la primera cohorte de pacientes se muestran en las Tablas 2 y 3.

Tabla 2. Resultados finales de *balanced accuracy* (%), precisión, MCC, F1 Score y AUC para cada modelo de ML evaluado en el análisis del grupo de pacientes de la primera ola.

Modelo	<i>Balanced Accuracy</i> (%)	Precisión	MCC	F1 Score	AUC
SVM	87,48	86,85	77,62	87,22	0,87
DT	86,02	85,40	76,32	85,76	0,86
BLDA	81,91	81,33	72,68	81,67	0,81
GNB	78,75	78,19	69,88	78,52	0,78
KNN	89,44	88,80	79,36	89,17	0,89
XGB	96,02	95,33	85,20	95,73	0,96

Resultados

Tabla 3. Resultados finales de sensibilidad, especificidad, kappa y DYI para cada modelo de ML evaluado en el análisis del grupo de pacientes de la primera ola.

Modelo	Sensibilidad (Recall)	Especificidad	Kappa	DYI
SVM	87,58	87,38	77,88	87,48
DT	86,12	85,91	76,58	86,02
BLDA	82,01	81,82	72,92	81,91
GNB	78,84	78,66	70,11	78,75
KNN	89,55	89,34	79,63	89,44
XGB	96,13	95,91	85,48	96,02

Tal como muestran los resultados, el modelo XGB proporcionó los valores más altos de adquisición y sensibilidad. El algoritmo XGB se ejecutó de manera consistente y uniforme con un valor de precisión superior al 95 %. Además, la curva ROC, representando las medidas de sensibilidad y especificidad, permitió comparar la capacidad de clasificación de los diferentes algoritmos de ML. La Figura 9 muestra los resultados obtenidos. Nuevamente, el sistema basado en XGB obtiene un AUC mayor, lo que significa que permite una mejor precisión en la predicción para este grupo de pacientes.

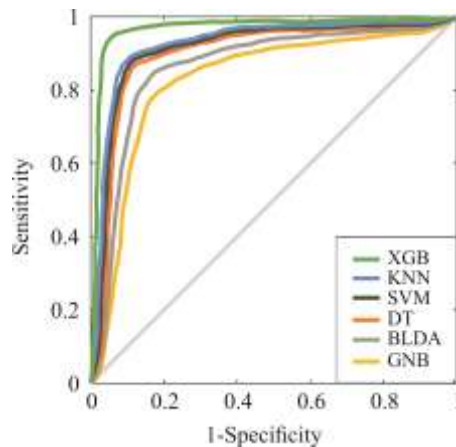


Figura 9. Curvas ROC para los seis modelos de ML evaluados en el grupo de pacientes de la 1ª ola.

En el presente estudio, los subconjuntos de entrenamiento del modelo presentan altas puntuaciones para todas las métricas y, en general, muestran puntuaciones ligeramente más bajas para el subconjunto de prueba. Estas similitudes se deben a que el algoritmo logra un nivel óptimo de entrenamiento sin sobreajuste ni subajuste. Como se muestra en los gráficos de radar en la Figura 10, el modelo XGB obtiene un área mayor que los otros métodos probados, y es un ejemplo adecuado de un modelo bien equilibrado con una alta capacidad de generalización, lo que significa que el algoritmo proporciona una salida correcta para cada nueva entrada.

Resultados

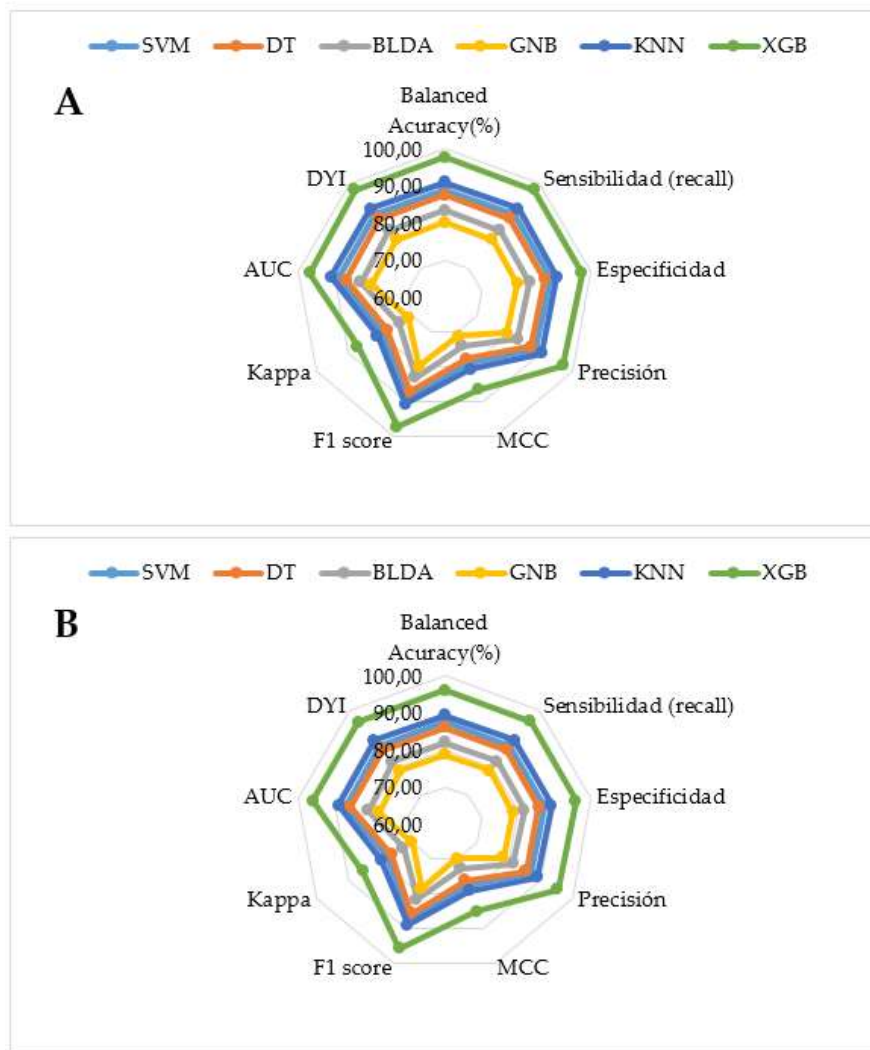


Figura 10. Gráficos de radar para las variables y modelos de ML utilizados en el grupo de pacientes de la 1ª ola. A: fase de entrenamiento. B: fase de test.

Como se puede ver en la Figura 11, según el modelo XGB elegido, los parámetros clínicamente más relevantes que se asocian a la mortalidad de los pacientes hospitalizados con COVID-19 en la primera ola, listados en orden descendente de relevancia, son linfocitos, urea, FiO₂, potasio, pH sérico, basófilos, tumores activos, bilirrubina total, temperatura, TFGe, alanina aminotransferasa (ALT/GPT), disnea y edad. Estas son variables que se pueden obtener fácilmente con un simple análisis de sangre y una exploración física en el momento de la admisión hospitalaria.

Resultados

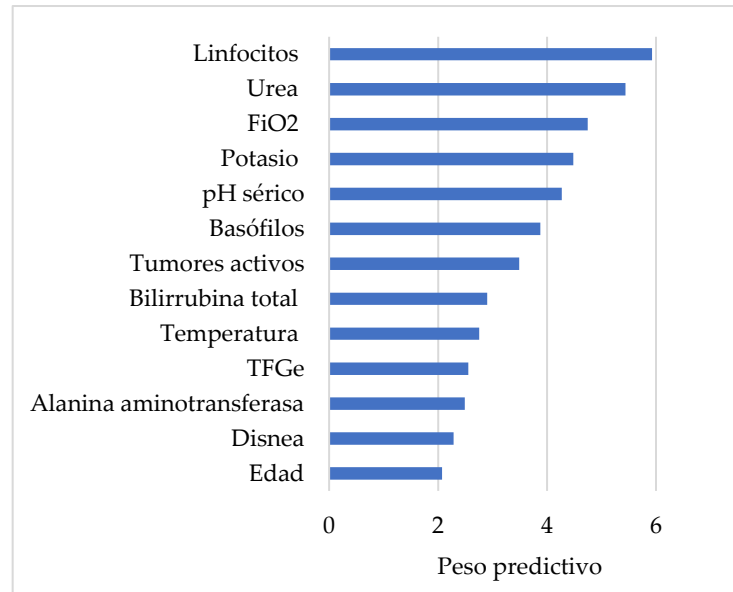


Figura 11. **Parámetros más relevantes asociados a la mortalidad de los pacientes hospitalizados por COVID-19 durante la primera ola.** Las unidades de medida de cada parámetro se muestran en las Tablas A1-A8 (Anexo 8.2.).

5.2.1.2. Segunda cohorte (olas 2ª a 4ª)

Los resultados obtenidos para las métricas de rendimiento en cada modelo (*balanced accuracy* (%), precisión, MCC, F1 Score, AUC, sensibilidad, especificidad, *kappa* y DYI) para la segunda cohorte de pacientes se muestran en las Tablas 4 y 5.

Tabla 4. Resultados finales de *balanced accuracy* (%), precisión, MCC, F1 Score y AUC para cada modelo de ML evaluado en el análisis del grupo de pacientes de las olas 2ª a 4ª.

Modelo	Balanced accuracy (%)	Precisión	MCC	F1 score	AUC
SVM	83,32	82,73	73,93	83,07	0,82
BLDA	80,37	79,78	71,31	80,12	0,79
DT	85,52	84,98	75,94	85,29	0,84
GNB	74,82	74,35	65,83	74,63	0,74
KNN	89,24	89,14	79,48	89,23	0,89
RF	95,83	95,15	86,83	95,53	0,95

Tabla 5. Resultados finales de sensibilidad, especificidad, *kappa* y DYI para cada modelo de ML evaluado en el análisis del grupo de pacientes de las olas 2ª a 4ª.

Modelo	Sensibilidad (Recall)	Especificidad	Kappa	DYI
SVM	83,42	83,22	73,39	83,32
BLDA	80,47	80,27	71,41	80,37
DT	85,62	85,42	75,46	85,52
GNB	74,91	74,73	66,67	74,82
KNN	89,31	89,18	79,01	89,24
RF	95,92	95,73	86,32	95,83

Resultados

En este caso, el modelo RF proporcionó los valores más altos de adquisición y sensibilidad, ejecutándose con un valor de precisión superior al 95 %. Esto se refleja también en la curva ROC representada en la Figura 12. Al comparar la capacidad de clasificación de los diferentes algoritmos de ML, se observó que el modelo RF obtiene un área mayor, lo que significa que permite una mejor precisión en la predicción para este grupo de pacientes.

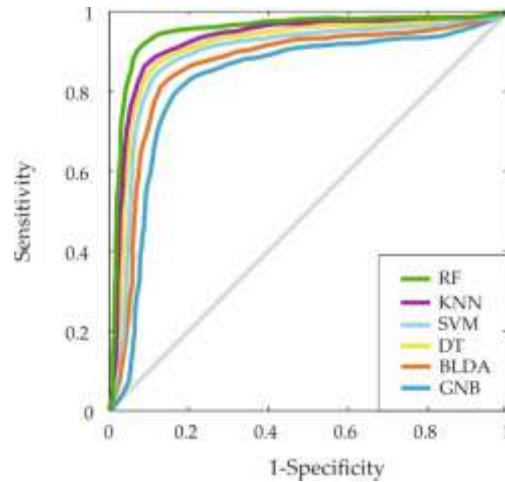


Figura 12. Curvas ROC para los seis modelos de ML evaluados en el grupo de pacientes de las olas 2ª a 4ª.

De nuevo, los subconjuntos de entrenamiento presentan altas puntuaciones para todas las métricas y, en general, muestran puntuaciones ligeramente más bajas para el subconjunto de prueba, evidenciando que este algoritmo también logra un nivel óptimo de entrenamiento sin sobreajuste ni subajuste. Como se muestra en los gráficos de radar en la Figura 13, el modelo RF obtiene un área mayor que los otros métodos probados, es un ejemplo adecuado de un modelo bien equilibrado con una alta capacidad de generación y, por tanto, también proporciona una salida correcta para cada nueva entrada.

Resultados

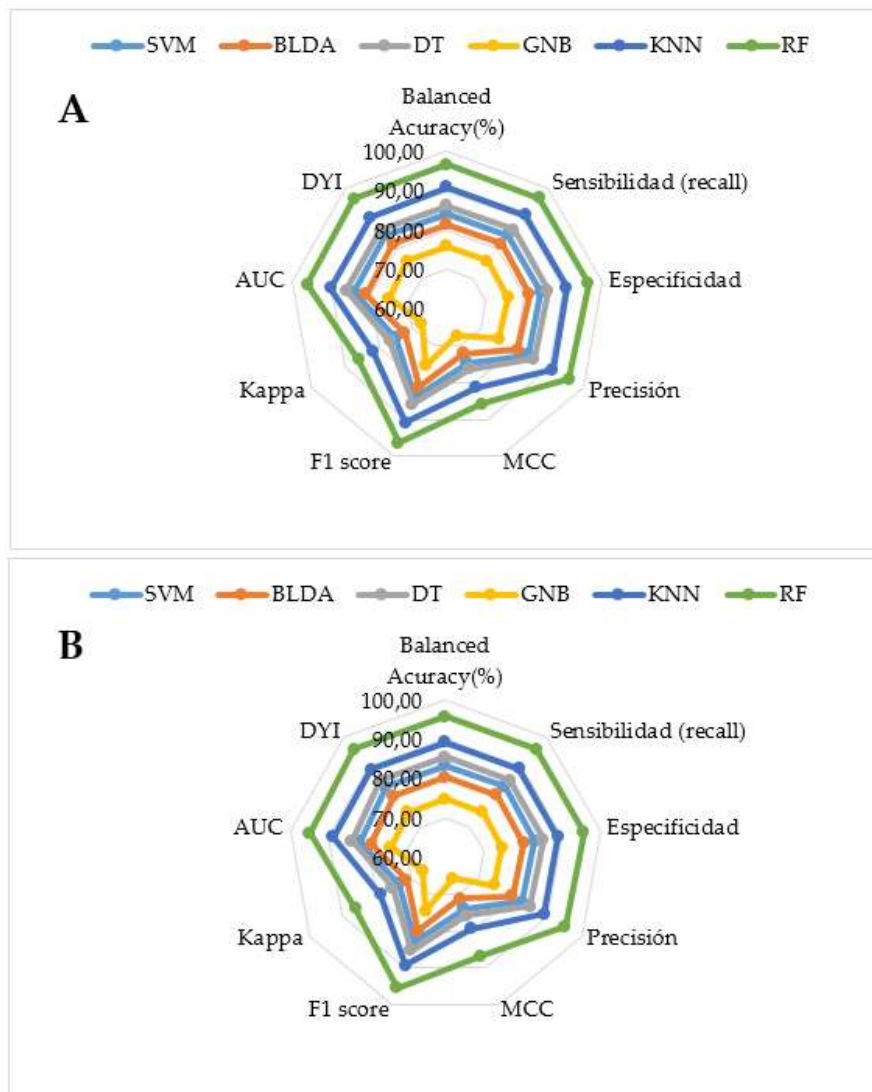


Figura 13. Gráficos de radar para las variables y modelos de ML utilizados en el grupo de pacientes de las olas 2ª a 4ª. A: fase de entrenamiento. B: fase de test.

Una vez elegido el modelo RF para el análisis de este grupo de pacientes, la Figura 14 muestra el histograma con los parámetros clínicamente más relevantes que se asocian a la mortalidad de los pacientes hospitalizados con COVID-19 en el periodo de la 2ª a la 4ª ola. Listados en orden descendente de relevancia, son: edad, escala CURB-65, urea, procalcitonina, TFGe, linfocitos, proteína C reactiva, fallo renal agudo, albúmina, sesiones de pronóstico, oxígeno suplementario y la escala GCS. Estas, de nuevo, son variables que se pueden obtener fácilmente con un simple análisis de sangre y una exploración física en el momento de la admisión hospitalaria.

Resultados

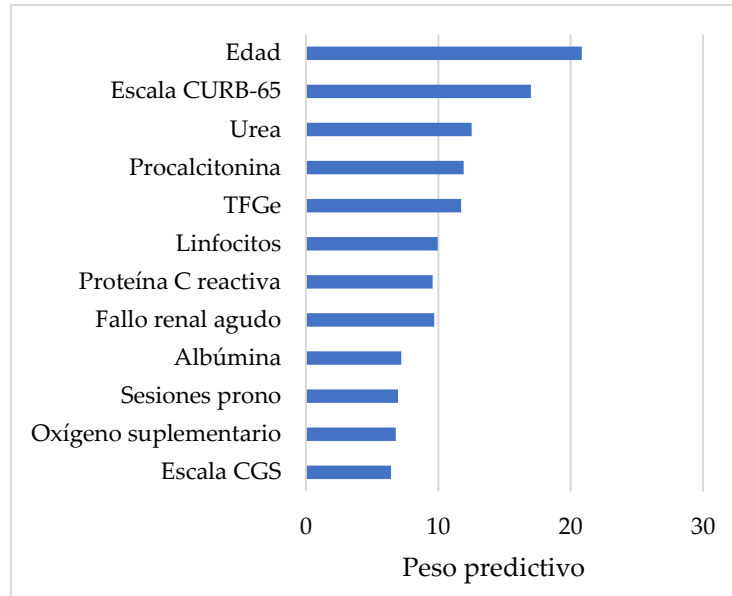


Figura 14. **Parámetros más relevantes asociados a la mortalidad de los pacientes hospitalizados por COVID-19 durante las olas 2ª a 4ª.** Las unidades de medida de cada parámetro se muestran en las Tablas A1-A8 (Anexo 8.2.).

5.2.1.3. Tercera cohorte (olas 5ª y 6ª)

Los resultados obtenidos para las métricas de rendimiento en cada modelo (*balanced accuracy* (%), precisión, MCC, F1 Score, AUC, sensibilidad, especificidad, *kappa* y DYI) para la tercera cohorte de pacientes se muestran en las Tablas 6 y 7.

Tabla 6. Resultados finales de *balanced accuracy* (%), precisión, MCC, F1 Score y AUC para cada modelo de ML evaluado en el análisis del grupo de pacientes de las olas 5ª y 6ª.

Modelo	Balanced Accuracy (%)	Precisión	F1 score	MCC	AUC
SVM	81,91	81,33	81,66	72,68	0,82
BLDA	78,19	77,64	77,96	69,38	0,78
DT	84,04	83,44	83,79	74,57	0,84
GNB	73,78	73,25	73,56	65,46	0,73
KNN	88,05	87,42	87,79	78,13	0,88
XGB	94,73	94,06	94,45	84,06	0,94

Tabla 7. Resultados finales de sensibilidad, especificidad, *kappa* y DYI para cada modelo de ML evaluado en el análisis del grupo de pacientes de las olas 5ª y 6ª.

Modelo	Sensibilidad (Recall)	Especificidad	DYI	Kappa
SVM	82,01	81,81	81,91	72,92
BLDA	78,29	78,10	78,19	69,61
DT	84,14	83,94	84,04	74,82
GNB	73,86	73,69	73,78	65,68
KNN	88,15	87,95	88,05	78,39
XGB	94,84	94,62	94,73	84,34

Resultados

Al estudiar esta cohorte, destacó el modelo XGB nuevamente, proporcionando los valores más altos de adquisición y sensibilidad, con un valor de precisión superior al 94%. Al comparar la capacidad de clasificación de los diferentes algoritmos de ML en la curva ROC, representada en la Figura 15, se observó que el modelo XGB obtiene un área mayor, lo que significa que permite una mejor precisión en la predicción para este grupo de pacientes.

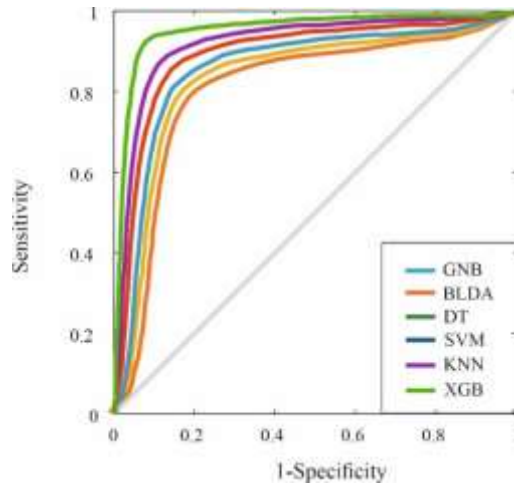


Figura 15. Curvas ROC para los seis modelos de ML evaluados en el grupo de pacientes de las olas 5ª y 6ª

De nuevo, los subconjuntos de entrenamiento presentan altas puntuaciones para todas las métricas y, en general, muestran puntuaciones ligeramente más bajas para el subconjunto de prueba, evidenciando que el XGB, en este caso, también logra un nivel óptimo de entrenamiento sin sobreajuste ni subajuste. Como se muestra en los gráficos de radar en la Figura 16, el modelo obtiene un área mayor que los otros métodos probados. Vistos los resultados, el XGB vuelve a ser un buen ejemplo de modelo bien equilibrado con una alta capacidad de generación otorgando una salida correcta para cada nueva entrada.

Resultados

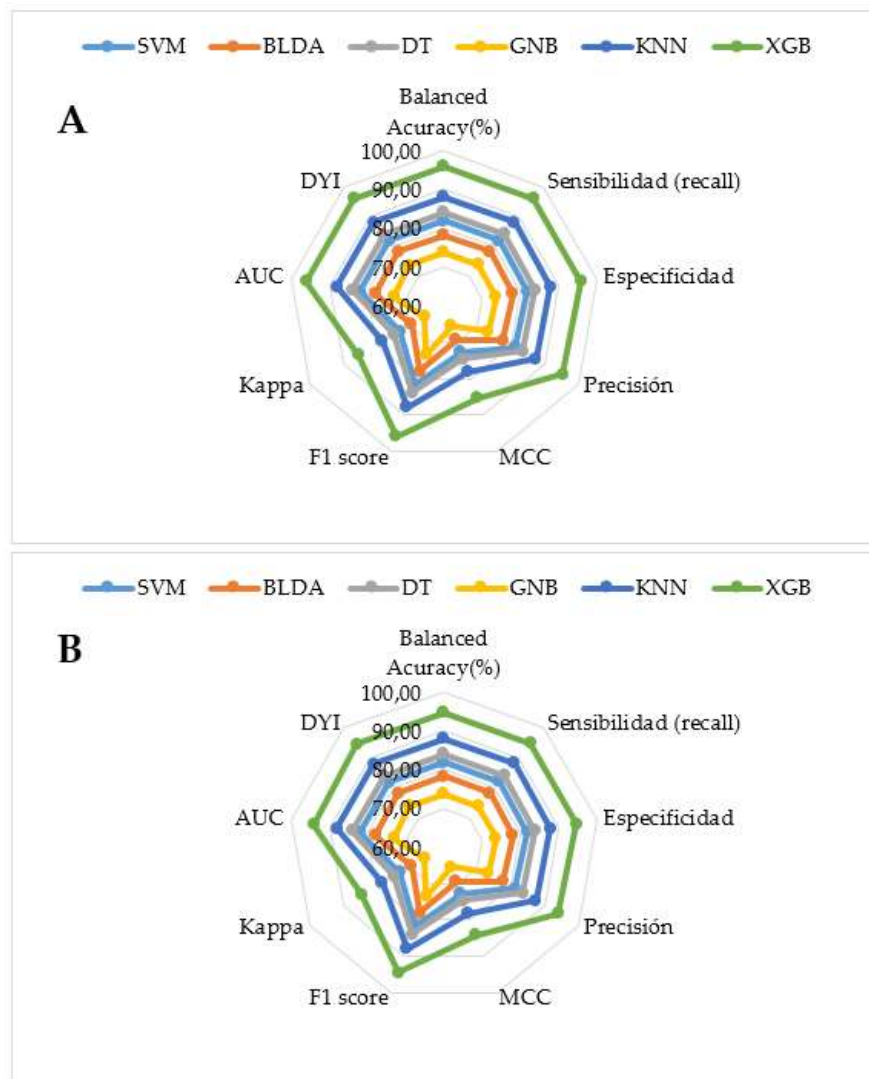


Figura 16. Gráficos de radar para las variables y modelos de ML utilizados en el grupo de pacientes de las olas 5ª y 6ª. A: fase de entrenamiento. B: fase de test.

La Figura 17 muestra el histograma con los parámetros clínicamente más relevantes que se asocian a la mortalidad de los pacientes hospitalizados con COVID-19 en el periodo de la 5ª a la 6ª ola según el modelo elegido XGB. Listados en orden descendente de relevancia, son: urea, escala CURB-65, TFGe, edad, albúmina, frecuencia respiratoria, diagnóstico de distrés respiratorio al ingreso, procalcitonina, volumen corpuscular medio, potasio, escala GCS y dímero D. De la misma manera que en las cohortes anteriores, son variables que se pueden obtener fácilmente con un simple análisis de sangre y una exploración física en el momento de la admisión hospitalaria.

Resultados

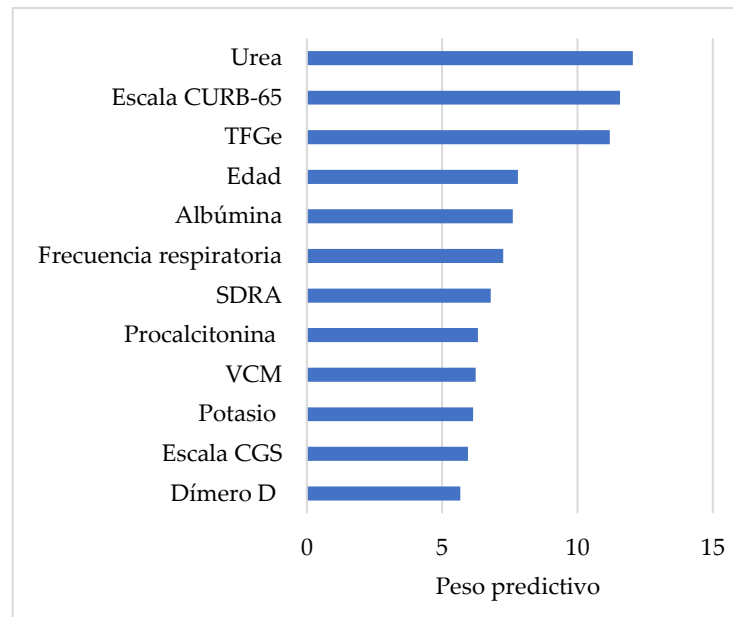


Figura 17. *Parámetros más relevantes asociados a la mortalidad de los pacientes hospitalizados por COVID-19 durante las olas 5ª y 6ª. Las unidades de medida de cada parámetro se muestran en las Tablas A1-A8 (Anexo 8.2.).*

5.2.2. Estudio de pacientes en UCI: biomarcadores

Se estudiaron 36 pacientes diagnosticados con COVID-19, hospitalizados en la UCI durante la primera ola, con el objetivo de identificar biomarcadores relacionados con una mayor tasa de mortalidad.

La Tabla 8 presenta los resultados obtenidos para las concentraciones de las citoquinas IFN- γ , LIF, TNF- α en muestras de pacientes tras 15 días de intubación en la UCI. En la Figura 18 se representan los histogramas que muestran la concentración de las mencionadas citoquinas en cada uno de los 36 pacientes estudiados.

Tabla 8. *Concentración de citoquinas en pacientes de la primera ola tras 15 días con VMI en la UCI.*

	1ª OLA (36 pacientes de UCI)		
	Fallecidos (12 pacientes)	Vivos (24 pacientes)	Global (36 pacientes)
	M $\pm$ DE	M $\pm$ DE	M $\pm$ DE
Concentración IFN γ (pg/ml)	1,64 $\pm$ 0,53	2,14 $\pm$ 1	1,98 $\pm$ 0,89
Concentración LIF (pg/ml)	36,48 $\pm$ 17,46	47,32 $\pm$ 28,86	43,71 $\pm$ 25,88
Concentración TNF α (pg/ml)	23,38 $\pm$ 46,82	12,79 $\pm$ 18,2	16,32 $\pm$ 30,53

De estos resultados se concluye que no se observaron diferencias estadísticamente significativas en las concentraciones de estas citoquinas entre los pacientes fallecidos y los supervivientes.

Resultados

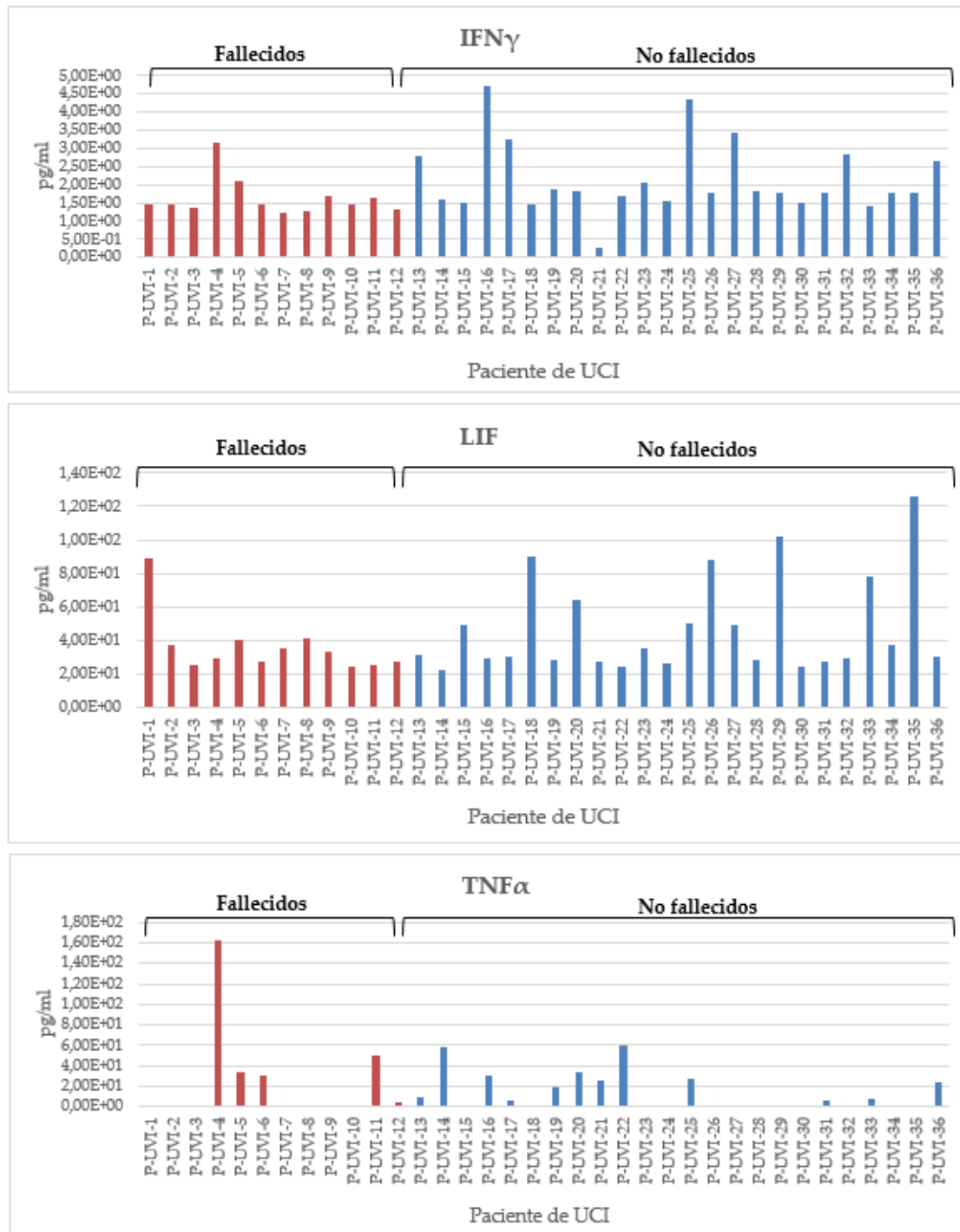


Figura 18. Concentración de las citoquinas IFN- γ , LIF y TNF- α en cada uno de los 36 pacientes ingresados en la UCI durante la primera ola.

6. Discusión

La pandemia de COVID-19 ha tenido un impacto devastador en la salud global, con millones de personas afectadas y una alta tasa de mortalidad entre pacientes hospitalizados. La aparición de nuevas variantes y el aumento de casos en determinados periodos ponen de manifiesto la necesidad de seguir investigando y mejorando las estrategias de respuesta. Además, la pandemia ha evidenciado los puntos débiles de los sistemas sanitarios y la importancia de investigar intervenciones y protocolos más eficaces para ofrecer una atención de mayor calidad.

En este contexto, identificar predictores de mortalidad en pacientes hospitalizados es crucial para mejorar la atención médica y optimizar los recursos sanitarios, ya que permite a los profesionales de la salud identificar rápidamente a los pacientes con mayor riesgo y tomar decisiones informadas sobre su tratamiento y manejo. Además, evaluar cómo estos predictores evolucionan con el avance de la pandemia permite observar el comportamiento de la enfermedad ante la aparición de nuevas variantes y/o el aumento progresivo de la inmunidad poblacional: desde una situación sin inmunidad (1ª ola), pasando por una inmunización natural tras el contacto inicial con el virus (olas 2ª a 4ª), hasta una combinación de inmunidad natural y vacunación masiva (olas 5ª y 6ª).

Las técnicas de ML ofrecen una herramienta poderosa para analizar grandes volúmenes de datos médicos y descubrir patrones que pueden no ser evidentes mediante métodos tradicionales. Al aplicar ML, es posible desarrollar modelos predictivos que integren múltiples variables clínicas y de laboratorio, proporcionando una evaluación precisa del riesgo de mortalidad. Estos modelos pueden utilizarse en tiempo real para apoyar la toma de decisiones, mejorando así los resultados de los pacientes y reduciendo la carga sobre los sistemas de salud.

La elección de modelos se fundamenta en sus características distintivas y en su capacidad para abordar diferentes tipos de datos y patrones. Cada grupo de pacientes presenta características específicas que requieren enfoques diferenciados para obtener los mejores resultados. Además, la diversidad metodológica permite una mejor interpretación y comprensión de los factores que influyen en la mortalidad en distintas etapas de la pandemia.

Por todo ello, en este estudio se utilizaron estos algoritmos de ML en la identificación de los parámetros más relevantes para predecir el riesgo de mortalidad en pacientes hospitalizados por COVID-19 en tres periodos distintos de la pandemia. El objetivo fue analizar la evolución de estos predictores y de la enfermedad en sí misma a medida que aumentaba la inmunidad global.

De todos los clasificadores de ML aplicados para la predicción, se proponen los métodos XGB y RF como modelos con mejor rendimiento de clasificación. Ambos modelos obtuvieron gráficos de radar equilibrados, lo que implica que no hubo grandes diferencias entre la fase de entrenamiento y la de prueba. No solo superaron al resto de algoritmos, sino que lo hicieron manteniendo altos niveles de precisión clínica en todas las cohortes. En sus respectivos escenarios

de entrenamiento, ambos alcanzaron valores de precisión superiores al 94 % y valores muy próximos a 1 en las curvas ROC, lo que respalda su utilidad práctica como herramientas de apoyo a la toma de decisiones clínicas.

Durante la primera ola, XGB fue el modelo con mejor rendimiento: *balanced accuracy* 96,0 % y precisión 95,3 %. Su estrategia de *boosting* secuencial permite que cada árbol corrija el error del anterior y, gracias a la regularización L1/L2 y al manejo nativo de valores faltantes, el algoritmo es capaz de modelar interacciones no lineales sin incurrir en sobreajuste. Esta arquitectura con bajo sesgo y control de varianza es ideal cuando se busca captar señales débiles y lograr alta exactitud de forma simultánea.

En la segunda cohorte de pacientes, entre las olas 2ª y 4ª, el mejor desempeño lo obtuvo RF, que registró *balanced accuracy* 95,8 % y precisión 95,1 %. El *bagging* en paralelo crea multitud de árboles independientes que se entrenan sobre muestras *bootstrap* y subconjuntos aleatorios de predictores, logrando baja varianza y gran estabilidad con apenas dos hiperparámetros críticos. Además, la estimación directa de importancia de variables facilita la interpretación clínica, algo esencial cuando los predictores están fuertemente correlacionados.

En la tercera cohorte, correspondiente a las olas 5ª y 6ª, XGB volvió a imponerse con *balanced accuracy* 94,7 % y precisión 94,1 %. En esta fase, la aparición de patrones sutiles (vacunación, nuevas variantes) volvió a favorecer un enfoque que combina optimización fina de la función de pérdida y ajustes automáticos de hiperparámetros para detectar gradientes muy pequeños sin perder generalización.

Ambos algoritmos, por tanto, exhibieron un comportamiento sobresaliente y consistente. XGB destaca por su capacidad para modelar interacciones complejas y mitigar el ruido mediante *boosting* y regularización; resulta óptimo cuando se requiere capturar detalles finos sin sacrificar precisión. RF, por su parte, sobresale cuando se necesita robustez y baja varianza con mínima necesidad de ajuste, aprovechando el *bagging* para estabilizar predicciones y ofrecer interpretabilidad directa.

Así, la evidencia demuestra que ambos métodos son fiables y precisos, pero cada uno maximiza su potencial en entornos de datos distintos gracias a su diferente equilibrio interno de sesgo-varianza y a la forma en que integra la información de los atributos.

Gracias a sus características (escalabilidad, alta velocidad de ejecución, entre otras), estos dos modelos han permitido lograr un rendimiento óptimo detectando las variables predictivas con más peso. Ambos métodos han sido utilizados en otros campos de la medicina (193–195), pero también en el estudio de la pandemia de la COVID-19, especialmente en la identificación de biomarcadores de severidad, ingreso en UCI y mortalidad (173,196,197).

De los resultados obtenidos en este trabajo, se concluye que tres de las variables analizadas aparecen como parámetros con elevado peso predictivo en las tres cohortes: la edad, los niveles de urea y la TFG_e. La combinación de estos tres parámetros proporciona una visión integral del estado de salud del paciente y su capacidad para resistir la infección por COVID-19.

La edad es un factor crítico debido a su asociación con la inmunosenescencia, que es el deterioro gradual del sistema inmunológico asociado con el envejecimiento. Este fenómeno aumenta la susceptibilidad de los pacientes mayores a infecciones graves y complicaciones. Por su parte, la urea y la TFG_e podrían ser indicadores de deficiencias en la función renal, que puede verse comprometida en pacientes con COVID-19 debido a la inflamación sistémica y el daño directo al tejido renal causado por el virus.

Según los resultados, la edad media de los pacientes fallecidos fue aumentando al avanzar en las diferentes olas a la vez que la de los pacientes que superaron la enfermedad fue disminuyendo, de manera que, en la primera ola, la diferencia en la media de edad entre fallecidos y supervivientes fue de tres años, mientras que en las olas 5ª y 6ª hubo una diferencia de once años. Este aumento en la diferencia de edad entre fallecidos y supervivientes sugiere que los pacientes más jóvenes desarrollaron una mayor resistencia al virus con el tiempo, posiblemente debido a la inmunidad adquirida por la exposición previa o la vacunación. Precisamente por esto tiene sentido que esta variable tenga un mayor peso predictivo a medida que avanzan las olas de la pandemia. De hecho, el estudio con aprendizaje automático coloca a la edad en una posición importante en la predicción de mortalidad de estos pacientes en cualquiera de las tres cohortes estudiadas, siendo las últimas olas en las que adquirió mayor relevancia.

Coincidiendo con nuestros resultados, en el estudio observacional de San Martín-López et al. (114), llevado a cabo en Madrid con más de 5.000 pacientes, la media de edad de los pacientes fallecidos por COVID-19 fue aumentando a medida que avanzó la pandemia mostrando una evolución en la mortalidad de la enfermedad y, por tanto, de la enfermedad en sí misma. Nuestro estudio también destaca la importancia de la edad como un factor de riesgo independiente, lo que refuerza la necesidad de estrategias de protección específicas para los grupos de mayor edad.

Algunos autores defienden que la aparición frecuente de la edad como posible factor crítico para la severidad y mortalidad de la COVID-19 se debe a su asociación directa con las comorbilidades de los pacientes de edad más avanzada. El estudio de Valero-Bover et al. (198) sostiene que la presencia simultánea de múltiples condiciones médicas en un paciente puede explicar el 98,2 % del efecto de la edad en la severidad de la enfermedad. También argumenta que comorbilidades como la hipertensión, la diabetes y las enfermedades cardiovasculares son comunes en los pacientes mayores y, por ello, pueden intensificar los efectos de la COVID-19, aumentando el riesgo de complicaciones graves y mortalidad.

En nuestro estudio, sin embargo, la edad aparece con un valor predictivo mucho más significativo que las comorbilidades. De hecho, aunque la mayoría de los pacientes incluidos en este estudio eran adultos mayores, esto no disminuye la importancia de la edad como factor predictivo. Más que un mero reflejo de la mayor mortalidad comúnmente asociada a las poblaciones de mayor edad, la edad en sí misma desempeña un papel crucial en la progresión y el desenlace de la COVID-19. Estudios como los de Rodríguez et al. (199), Garrafa et al. (200) y Tezza et al. (201) han analizado datos de pacientes hospitalizados por COVID-19 utilizando modelos de ML obteniendo como resultado que la edad es un buen predictor independiente del riesgo de mortalidad. Esto está respaldado por una amplia evidencia en la literatura médica que destaca que los cambios inmunológicos y fisiológicos asociados con el envejecimiento aumentan la susceptibilidad a la enfermedad grave (202).

Por su parte, la urea, un producto de desecho del metabolismo de las proteínas, puede reflejar, a partir de sus niveles en sangre, la función renal y el estado metabólico del paciente. Los niveles elevados de urea pueden indicar una disfunción renal, que es común en pacientes con COVID-19 debido a la inflamación y el daño renal directo causado por el virus. Según un estudio realizado por Murri et al. (203), los niveles de urea, junto con otros parámetros de laboratorio, están estrechamente relacionados con la mortalidad de los pacientes, lo que sugiere una actividad renal anormal en los pacientes que fallecieron. Además, en nuestro estudio, la TFG_e también se encuentra entre los parámetros más relevantes en las tres cohortes, coincidiendo con los resultados del estudio de Subudhi et al. (204), en el que se demuestra que la TFG_e es un predictor independiente de mortalidad en pacientes con COVID-19 utilizando más de 10 modelos de ML.

Dado que la urea y la TFG_e aparecen junto con la edad como variables de elevado peso predictivo, debe tenerse en cuenta que estos parámetros podrían estar reflejando aspectos similares del estado de salud del paciente, lo que podría limitar su utilidad como predictores independientes. El estudio llevado a cabo por Sánchez-Montalvá et al. (205) muestra que ciertos parámetros de laboratorio tienen un valor predictivo limitado debido a su colinearidad o redundancia, es decir, que sus valores aumentan y disminuyen siempre a la par, lo cual reduce el poder predictivo al combinar varios parámetros entre sí. Este fenómeno de colinearidad es común en estudios clínicos y evidencia la necesidad de utilizar técnicas estadísticas avanzadas para identificar los predictores más relevantes y evitar la redundancia en los modelos predictivos.

Los niveles de linfocitos aparecen como buenos predictores de mortalidad en las dos primeras cohortes. Los linfocitos son un componente crucial del sistema inmunológico y su recuento puede reflejar la capacidad del organismo para combatir infecciones. La investigación durante la pandemia ha demostrado que los pacientes que fallecen a causa de la COVID-19 tienden a tener un recuento de linfocitos significativamente bajo frente a los que sobreviven

(206,207). Por tanto, este parámetro se ha clasificado como un buen predictor de mortalidad (201,204), tal como muestra nuestro estudio en las dos primeras cohortes.

En la tercera cohorte de pacientes, es decir, en el periodo correspondiente a las olas 5ª y 6ª, el recuento de linfocitos deja de aparecer entre los predictores de mayor peso. Esto puede deberse a varios factores. Las distintas variantes del virus pueden afectar el sistema inmunológico de manera diferente, influyendo en la relevancia de los linfocitos como predictores de mortalidad. Sin embargo, el factor más importante que diferencia a esta cohorte de pacientes de las anteriores es la introducción y amplia distribución de las vacunas contra la COVID-19.

Las vacunas han demostrado ser altamente efectivas en la reducción de la gravedad de la enfermedad y la mortalidad. Esto podría implicar que, en estas últimas olas, el nivel de inmunidad alcanzado gracias a la combinación de la población inmunizada de forma natural y la inmunizada a través de las vacunas era ya lo suficientemente elevado como para modificar el perfil inmunológico de los pacientes. De esta forma, los pacientes tendrían una respuesta inmunitaria diferente a la infección, lo que podría afectar la relevancia de los linfocitos como predictores de mortalidad. También pueden influir las características demográficas y epidemiológicas de las poblaciones hospitalizadas y los tratamientos cada vez más efectivos, como el uso de antivirales, corticosteroides y terapias con anticuerpos monoclonales, que han mejorado significativamente los resultados clínicos.

Cabe destacar que las primeras y más frecuentes vacunas utilizadas fueron las de ARNm. Estas vacunas son de desarrollo rápido y tienen la capacidad de ser modificadas rápidamente para adaptarse a las nuevas variantes del virus. Sin embargo, aunque este tipo de vacunas había sido investigado previamente, nunca se había usado en humanos. Vacunas tradicionales como las de virus inactivados o las de vectores virales habrían requerido menos ensayos clínicos previos a su comercialización, debido a la experiencia previa con tecnologías similares. Además, las vacunas tradicionales implican una mayor accesibilidad en países de recursos limitados, ya que su almacenamiento no requiere infraestructuras complejas. Esto ha suscitado un debate científico sobre la posibilidad de que las vacunas tradicionales pudieran haberse lanzado antes al mercado para su uso inicial de emergencia, en caso de no haberse priorizado el desarrollo de las vacunas de ARNm.

Sin embargo, las vacunas tradicionales han mostrado una eficacia menor frente a la enfermedad: 60-65 % frente a un 90-95 % en el caso de las vacunas de ARNm. Los efectos adversos observados en cualquiera de ellas han sido similares y esperables (85). Algunos argumentan que las farmacéuticas y los gobiernos se centraron en las vacunas de ARNm debido a intereses económicos (208,209). Sin embargo, no hay evidencias concluyentes de que las

vacunas tradicionales hubieran podido lanzarse antes que las de ARNm o de que su lanzamiento o investigación se haya retrasado deliberadamente.

La eficacia de las vacunas de ARNm, su flexibilidad y su rapidez en la producción a gran escala han hecho de ellas la opción principal en la que se basa la vacuna de la COVID-19, y se ha demostrado que estas ventajas priman frente a los elevados costes de producción y las dificultades de envío y almacenamiento que conllevan (210–212).

Continuando con el análisis de predictores de mortalidad, nuestro estudio muestra una alta relevancia de la FiO_2 durante la primera ola. Este parámetro está estrechamente vinculado a la insuficiencia respiratoria hipoxémica aguda, así como lo están el potasio y el pH, que también aparecen como variables de elevado peso predictivo. La insuficiencia respiratoria hipoxémica aguda es una complicación grave de la COVID-19 que requiere intervenciones intensivas, como la ventilación mecánica, y está asociada con una alta mortalidad. Estos hallazgos son consistentes con los resultados obtenidos por Bolourani et al. (213), Khuluq et al. (214), y Bendavid et al. (215).

Se ha demostrado que la hipopotasemia, que puede afectar la función cardíaca y muscular aumentando el riesgo de complicaciones graves, es un indicador de un peor pronóstico. Se cree que esto puede deberse a la interferencia del virus con ACE2, un factor clave en las cascadas enzimáticas que mantienen niveles adecuados de potasio en las células, causando la desregulación de la actividad fisiológica celular (216). Esta interrupción de los sistemas fisiológicos celulares también parece causar la desregulación de la homeostasis de electrolitos y el desequilibrio del pH (217). Un desajuste en el pH puede afectar múltiples sistemas del organismo, incluyendo el sistema respiratorio y el cardiovascular, aumentando el riesgo de complicaciones graves.

El recuento de basófilos fue otra variable con un peso significativo en la predicción de la mortalidad en nuestro estudio. Se ha demostrado que los basófilos juegan un papel importante en enfermedades inflamatorias crónicas a través de la expresión de Th17 y Th17/Th1. Estas células inmunitarias están involucradas en la respuesta inflamatoria y pueden contribuir a la patogénesis de la COVID-19. En el estudio de Murdaca et al. (218), se observó una reducción significativa de basófilos durante los primeros tres días de hospitalización, volviendo a niveles normales poco después. Esta reducción inicial podría reflejar una respuesta inflamatoria aguda al virus, seguida de una normalización a medida que la respuesta inmunitaria se estabiliza.

También en la primera ola, los tumores activos fueron un factor relevante en relación con la mortalidad de los pacientes. Los pacientes con cáncer tienen un sistema inmunológico comprometido debido tanto a la enfermedad como a los tratamientos agresivos (quimioterapia, radioterapia). Se ha sugerido que el riesgo de mortalidad podría deberse a un estado subóptimo del sistema inmunológico de estos pacientes debido a dichos tratamientos contra el cáncer, lo que

debilitaría la respuesta del organismo al SARS-CoV-2 (22). Sin embargo, los hallazgos de Várnai et al. (219) y Desai et al. (220) coinciden en que esta tendencia no se debe a los tratamientos relacionados con el cáncer, sino a la edad, el sexo y las múltiples comorbilidades asociadas que pueden agravar la infección por COVID-19. Esto sugiere que los factores demográficos y las condiciones subyacentes juegan un papel más significativo en la mortalidad de estos pacientes que los tratamientos oncológicos en sí.

Es interesante observar que estos últimos parámetros mencionados (FiO_2 , pH, potasio, basófilos y tumores activos) aparecieron como predictores de mortalidad principalmente durante la primera ola de COVID-19 pero no tanto en las posteriores. Esto puede deberse a la falta de tratamientos específicos y protocolos estandarizados al inicio de la pandemia, lo que llevó a una mayor dependencia de estos parámetros para predecir la mortalidad. Al igual que con el recuento linfocitario, los tratamientos y protocolos médicos evolucionaron rápidamente a medida que se comprendía la enfermedad, mejorando la atención y gestión clínica, lo que pudo haber mitigado la influencia de estos parámetros en las olas posteriores. También es posible que los cambios en la composición demográfica y las características de los pacientes entre las diferentes olas hayan afectado a los factores predictivos de mortalidad, de la misma manera que pueden haber tenido influencia las diferentes variantes del virus con distinta tasa de virulencia y transmisibilidad(114).

Comunes a la segunda y la tercera cohorte de pacientes, y con una potencia predictiva significativa en ambas, aparecen la escala CURB-65, la procalcitonina, la albúmina y la Escala Coma de Glasgow (GCS). Estos parámetros han demostrado ser útiles en la evaluación de la gravedad de la enfermedad y en la predicción de la mortalidad en pacientes hospitalizados con COVID-19.

La escala CURB-65 ha sido evaluada en otros estudios como un predictor adecuado de mortalidad antes de la hospitalización por neumonía por COVID-19 (221,222). También se usa comúnmente para determinar la gravedad clínica en pacientes con cualquier tipo de neumonía (223). Esta escala es fácil de usar y proporciona una evaluación rápida de la gravedad de la enfermedad, lo que puede ser crucial en situaciones de emergencia, ayudando a priorizar la atención de los pacientes más graves. Incluye la tensión arterial, el bajo nivel de conciencia, el nivel de urea y la edad. Como se ha mencionado, estas dos últimas variables aparecen en nuestro estudio como potentes predictores de mortalidad. Por ello, en este estudio, la escala CURB-65 es un buen parámetro predictivo y, dado que puede determinarse en el momento de la admisión hospitalaria, esta información puede usarse para realizar un triaje rápido y eficiente.

Respecto a las razones por las cuales esta escala CURB-65 no aparece como predictor en la primera ola, se pueden barajar varias teorías. La primera ola se caracterizó por una falta de datos suficientes y de alta calidad sobre los pacientes con COVID-19 debido a la saturación del

sistema sanitario y a la necesidad de una asistencia rápida. La urgencia de tratar a un gran número de pacientes en un corto período de tiempo puede haber limitado la capacidad para recopilar datos detallados y precisos, haciendo más complicada la obtención de los datos necesarios para la evaluación de esta escala. Además, la CURB-65 es una herramienta desarrollada originalmente para la neumonía adquirida en la comunidad, así que su aplicación para la COVID-19 se exploró más a fondo en olas posteriores, cuando se disponía de más datos y se podía evaluar su eficacia en este contexto específico. De hecho, con el tiempo, se desarrollaron y validaron otras herramientas de predicción específicas para COVID-19, como las escalas ISARIC-4C, PLABAC o PRABLE, que han obtenido resultados comparables e incluso mejores en algunos casos (224,225). Estas nuevas herramientas han sido diseñadas específicamente para abordar las características únicas de la COVID-19 y han demostrado ser valiosas en la predicción de la mortalidad y la gestión clínica de los pacientes.

En relación a los valores de procalcitonina y albúmina, son marcadores conocidos de inflamación y sepsis, y su elevación refleja la gravedad de la respuesta inflamatoria que puede estar asociada a un peor pronóstico. Estudios como el de Jamshidi et al. (226) o el ya mencionado estudio de Subhudi et al. (204), han concluido también que los niveles elevados de procalcitonina, niveles bajos de albúmina y recuento bajo de linfocitos son buenos predictores de mortalidad. Concretamente, la albúmina, una proteína producida por el hígado, normalmente disminuye en estados de inflamación crónica y estrés fisiológico debido a su papel en la respuesta de fase aguda. Niveles bajos pueden indicar un estado de desnutrición o inflamación crónica, ambos asociados con un peor pronóstico en pacientes con COVID-19. Todos estos datos son consistentes con un estado de inflamación avanzada y daño orgánico, explicando la asociación con un mayor riesgo de mortalidad.

Por su parte, la escala GCS es una herramienta ampliamente utilizada en la evaluación neurológica y proporciona una medida objetiva del nivel de conciencia del paciente. Valores bajos se han relacionado con un mayor riesgo de mortalidad en casos de pacientes hospitalizados por COVID-19 (227). La disminución del nivel de conciencia puede ser un indicador de hipoxia cerebral, sepsis o daño neurológico, todos asociados con un peor pronóstico. Cabe destacar que la posible confusión del paciente también está incluida en la mencionada escala CURB-65, lo que podría explicar su relación con el elevado peso predictivo de la GCS. La implementación de protocolos de monitorización continua de los pacientes, así como la capacitación del personal médico en la interpretación de estas escalas es crucial para una correcta evaluación del paciente.

Además de las mencionadas, este estudio muestra que ciertas variables de monitorización hospitalaria, incluyendo la insuficiencia renal aguda, la necesidad de oxigenación suplementaria, las sesiones de prono y/o el desarrollo de un SDRA, tienen potencia predictiva en algunos grupos.

Estas variables no se pueden extraer en el momento de admisión hospitalaria, sino que rastrean la progresión de la enfermedad durante la hospitalización y proporcionan información crítica sobre el riesgo de mortalidad. La combinación de estas herramientas puede mejorar la precisión de la evaluación clínica y ayudar a identificar a los pacientes que requieren intervenciones más intensivas. Por tanto, es necesario dar mucha importancia al monitoreo continuo y a la intervención temprana en pacientes con COVID-19 con estas alteraciones.

El desarrollo de insuficiencia renal aguda se ha asociado con mayor riesgo de mortalidad en pacientes con COVID-19. Concretamente, en el estudio multicéntrico de Ponce et al. (228), se desarrolló una escala a modo de modelo predictivo con modelos de ML relacionando la insuficiencia renal aguda con la mortalidad de pacientes infectados. El elevado peso predictivo de este parámetro puede deberse a una combinación de factores que promueven el daño renal como la hipoxia, la sepsis y la toxicidad de los medicamentos utilizados en el tratamiento de la COVID-19. Por otro lado, Almenara et al. (229) observaron que la incidencia de la insuficiencia renal aguda varía entre un 5 % y un 80 % dependiendo de la población y el tipo de estudio realizado. Esta variabilidad destaca la heterogeneidad de la enfermedad según la población analizada y la necesidad de estudios adicionales para comprender mejor los factores de riesgo y los mecanismos subyacentes de la insuficiencia renal aguda en pacientes con COVID-19.

La oxigenoterapia suplementaria en pacientes con COVID-19 es un factor crítico para evaluar el perfil del paciente. Aunque no se ha identificado como predictor de mortalidad en otros estudios, la posible progresión a mayores necesidades de oxígeno o soporte respiratorio adicional puede indicar el desarrollo de un SDRA. Este empeoramiento sí se ha asociado con tasas de mortalidad mayores, tal como muestran los estudios de Aktar et al. (230) y Xu et al. (231).

A pesar de que el SDRA causado por la COVID-19 y un SDRA tradicional o por otras causas presentan manifestaciones clínicas similares, como la hipoxemia o la distensibilidad pulmonar, el SDRA relacionado con la COVID-19 se ha asociado con una respuesta inflamatoria mayor. La inflamación sistémica severa en el SDRA por COVID-19 puede llevar a un daño pulmonar extenso y a la disfunción de otros órganos. Algunos estudios moleculares, como el de Batra et al. (232), han mostrado diferencias en la abundancia de ciertos metabolitos, lípidos y proteínas al realizar comparaciones. Se han identificado vías de señalización específicas en el SDRA ligado a COVID-19, como la activación del metabolismo de la arginina y vías de señalización inflamatorias como la de IL-17, las MAPK, los inhibidores JAK y el TNF (233,234). Esto evidencia que existen algunos procesos biológicos específicos del SDRA relacionado con la COVID-19 y, por tanto, abre puertas a posibles dianas terapéuticas.

Aunque la necesidad de sesiones en prono aparezca como variable con peso predictivo, se puede asociar a la presencia de SDRA en los pacientes. Esta técnica se ha utilizado para mejorar

la oxigenación en pacientes con SDRA al redistribuir el flujo sanguíneo y mejorar la ventilación de las áreas pulmonares dependientes. Se ha demostrado que, aunque en los pacientes no críticos esta práctica no mejoró los resultados significativamente, sí fue muy útil en los pacientes que requirieron VMI (235). En este estudio, sin embargo, parece existir una colinearidad con la presencia de SDRA y, por tanto, no se considera un predictor independiente de mortalidad.

Entre la segunda y tercera cohortes se encuentran mayores similitudes que entre cualquiera de ellas y la primera cohorte. En la primera ola, los pacientes no habían tenido ningún contacto previo con el SARS-CoV-2. Además, en esta primera ola se produjo la mayor saturación de los sistemas sanitarios y tanto los protocolos como los tratamientos se realizaban contando con escasez de recursos y de conocimientos. Sin embargo, a partir de la segunda ola, los sistemas sanitarios y los protocolos estaban más estructurados, y la inmunización natural de la población había comenzado. Como se ha mencionado, la principal diferencia entre la segunda (2ª a 4ª ola) y la tercera (5ª y 6ª ola) cohorte de pacientes fue la vacunación masiva. El hecho de que, a pesar de ello, el perfil de predictores de mortalidad sea similar entre estas dos cohortes, podría sugerir que la inmunización natural y la inducida por las vacunas son igualmente potentes a la hora de inducir inmunidad frente a la enfermedad, sin olvidar que la vacunación masiva fue crucial para la aceleración de este proceso reduciendo notablemente la mortalidad.

Estudios como los de Qassim et al. (236) o Diani et al. (237) muestran que la efectividad de las inmunidad inducida naturalmente es más duradera que la inducida por vacunación, aunque tiene la desventaja de perder potencia frente a nuevas variantes del virus. Por su parte, las vacunas, especialmente las de ARNm, cuentan con la posibilidad de realizar modificaciones fácilmente para hacer frente a las nuevas variantes. Según Spinardi y Srivastava (238), la combinación de ambas, la llamada “inmunidad híbrida”, proporciona la protección más robusta. Da lugar a una respuesta inmune más fuerte produciendo células B de memoria de muy alta calidad y genera una protección extensa frente a la enfermedad sintomática.

En el análisis estadístico de la concentración de las citoquinas TNF- α , LIF e IFN- γ en pacientes de la primera ola tras quince días de intubación en la UCI, no se encontraron diferencias estadísticamente significativas entre los pacientes supervivientes y fallecidos. Cabe destacar que estos hallazgos difieren de algunos estudios previos que han identificado estas citoquinas como biomarcadores significativos en la progresión y severidad de la COVID-19. En estudios multicéntricos como el de Soy et al. (239) o el de Rabaan et al. (240), se observó que las citoquinas TNF- α , TGF- β e IFN- γ eran biomarcadores significativos para la evaluación de la gravedad en pacientes hospitalizados, algunos críticos, durante la fase aguda de la enfermedad. Estos estudios sugieren que las citoquinas pueden desempeñar un papel en la patogénesis de la COVID-19 y podrían ser útiles como biomarcadores en ciertos contextos clínicos.

Sin embargo, cabe destacar que el momento de la toma de muestras es crucial para la interpretación de los resultados. Las concentraciones de citoquinas pueden variar significativamente a lo largo del curso de la enfermedad, y las muestras tomadas en diferentes fases pueden arrojar resultados distintos. En nuestro estudio, solamente se tomó una muestra para este análisis en cada paciente en un momento concreto de su evolución, tras 15 días de intubación, por lo que no se ha podido realizar el análisis en diferentes fases de la enfermedad. No se han encontrado estudios en los que se hayan estudiado estas tres citoquinas específicamente en muestras de pacientes intubados durante más de una semana. Según los datos de este trabajo únicamente, realizado con un número reducido de pacientes y una sola muestra por paciente, la concentración de estas citoquinas no es un buen biomarcador de mortalidad en pacientes intubados con COVID-19 severo.

Por todo lo anteriormente mencionado, este estudio muestra resultados de relevancia clínica y epidemiológica: se pueden identificar variables de elevado peso predictivo a partir de una analítica básica, una exploración física y la información demográfica del paciente en el momento de la admisión hospitalaria. Esto subraya la importancia de una evaluación inicial exhaustiva y sistemática que permita determinar rápidamente qué pacientes requieren atención urgente. Una estratificación temprana del riesgo es crucial para la toma de decisiones clínicas y la optimización de los recursos sanitarios.

Relacionando esto con la reciente pandemia, se sabe que, en las primeras olas, la incertidumbre y la falta de datos robustos dificultaron la gestión eficiente de los pacientes. La implementación de modelos predictivos basados en variables fácilmente obtenibles podría haber permitido la priorización de los pacientes con mayor riesgo, mejorando los resultados clínicos y reduciendo la mortalidad. Además, este enfoque habría permitido una mejor planificación y distribución de los recursos hospitalarios, como camas de UCI y ventiladores, en un momento en que estos eran extremadamente limitados.

Es importante reconocer que la calidad y acceso de los datos varía entre diferentes centros y regiones, lo que puede limitar la aplicabilidad general de los modelos predictivos. Se ha visto que, en estudios de predictores de mortalidad, factores como la variabilidad demográfica (241) y socioeconómica (242), los diferentes tratamientos y prácticas empleadas, la diversidad de sistemas de salud (243) o las diferencias en la ejecución y programación de los sistemas de análisis (244), afectan significativamente a los resultados. Es por esto que los resultados obtenidos en este estudio son particulares y específicos de la población analizada.

Muchos grupos mencionan esta variabilidad como una limitación, ya que los resultados no pueden extrapolarse a poblaciones más amplias, dada la heterogeneidad de las mismas. Sin embargo, precisamente es este hecho el que sitúa a este tipo de herramientas en la primera línea

de actuación frente a situaciones críticas como lo es una emergencia sanitaria. Estudios como los de Marinetti et al. (241) o Paganuzzi et al. (243) sugieren que la heterogeneidad poblacional permite entender mejor las respuestas locales a la pandemia y mejorar la preparación futura para crisis similares, especialmente en zonas con sistemas de salud descentralizados. Ambos estudios defienden que adaptar los modelos a las características de cada población puede mejorar su precisión y utilidad en la práctica clínica.

Mirando hacia el futuro, el análisis de datos con técnicas de ML demuestra ser una herramienta poderosa que puede mejorar significativamente la gestión de futuras pandemias o crisis similares gracias a su capacidad para analizar grandes volúmenes de datos de forma rápida y eficiente. Su implementación temprana utilizando una muestra relativamente pequeña de datos (150-200 pacientes) permite salvar vidas mediante la personalización de la atención clínica y la optimización de los recursos sanitarios, aumentando la eficacia de las intervenciones y reduciendo la mortalidad.

Uno de los grandes desafíos del trabajo con IA y una posible limitación del presente estudio, es la dificultad en la interpretación de los resultados. Por ello, es crucial la existencia de grupos multidisciplinares en la investigación que permitan una colaboración estrecha entre investigadores, médicos e ingenieros, como ha tenido lugar en este trabajo. Por otro lado, está la confiabilidad de los algoritmos: es necesario probar su reproducibilidad y validar externamente los modelos en otras poblaciones. El objetivo de ello es validar el uso de la herramienta, y no universalizar los resultados derivados de este estudio. Como se ha mencionado, la ventaja de utilizar estos modelos radica en su capacidad para obtener resultados adaptados a cada población específica. Estudios adicionales en esta dirección permitirán estandarizar el uso de estas herramientas a nivel hospitalario en el análisis de otras enfermedades o emergencias sanitarias.

Los resultados de este estudio no solo contribuirán al desarrollo de enfoques clínicos personalizados, mejorando la atención sanitaria de los pacientes, sino que también muestran la importancia de adaptar las estrategias de intervención a las características específicas de cada población. Es fundamental recordar que la inteligencia artificial es una herramienta poderosa que, utilizada de manera responsable, puede transformar la medicina, haciéndola más eficiente y de mejor calidad. La investigación continua y la validación de estos modelos son esenciales para asegurar su eficacia y aplicabilidad en diferentes contextos, preparándonos mejor para futuras emergencias sanitarias.

7. Conclusiones

Primera. La revisión de las características clínicas de los pacientes hospitalizados con COVID-19 en el Hospital Universitario Río Hortega de Valladolid ha revelado que la edad avanzada y la presencia de comorbilidades, como la hipertensión, la diabetes y las enfermedades cardiovasculares, fueron factores comunes entre los pacientes. Al ingreso hospitalario, los síntomas más frecuentes fueron disnea, tos y fiebre, y el diagnóstico principal fue insuficiencia respiratoria aguda. Las complicaciones más frecuentes en los pacientes hospitalizados fueron el SDRA y la insuficiencia renal aguda. Las causas más comunes de éxitus fueron el fallo multiorgánico y la insuficiencia respiratoria.

Segunda. Los predictores de mortalidad variaron a lo largo de la pandemia. En la primera ola, los principales predictores fueron la edad, los niveles de urea, la TFG_e, la FiO₂, el pH, el potasio, los basófilos y los tumores activos. En las olas segunda a cuarta, emergieron como variables con mayor peso predictivo la escala CURB-65, la procalcitonina, la albúmina y la escala Glasgow Coma Scale (GCS). En las olas quinta y sexta, la edad, los niveles de urea y la TFG_e continuaron siendo predictores relevantes, pero la vacunación y los tratamientos modificaron algunos patrones de mortalidad.

Tercera. La evolución de los patrones de mortalidad mostró que las variantes del virus, la mejora en los tratamientos y la vacunación influyeron significativamente en los factores predictivos. La combinación de la inmunidad natural y la vacunación modificó el perfil inmunológico de los pacientes, afectando dichos predictores de mortalidad y provocando variaciones según la población analizada.

Cuarta. En la identificación de biomarcadores de inflamación en pacientes ingresados en la UCI durante la primera ola, no hubo diferencias significativas en las concentraciones de TNF- α , LIF y IFN- γ entre pacientes vivos y fallecidos. La variabilidad en la respuesta inmunológica y el momento de la toma de muestras pueden influir en la utilidad de estos biomarcadores. Se necesitan estudios adicionales con tamaños de muestra más extensos y un diseño estandarizado para validar su utilidad en la práctica clínica.

Quinta. Los modelos de predicción de patrones de mortalidad mediante técnicas de aprendizaje automático demostraron ser efectivos para identificar a los pacientes con mayor riesgo, utilizando variables fácilmente obtenibles en la admisión hospitalaria. La validación sugiere su potencial uso en futuras emergencias sanitarias, mejorando la personalización de los tratamientos y la optimización de los recursos. La variabilidad en la calidad y disponibilidad de los datos entre centros y regiones subraya la necesidad de adaptar los modelos a las características específicas de cada población. Continuar investigando y validando estos modelos es esencial para asegurar su eficacia y aplicabilidad en diferentes contextos.

8. Anexos

8.1. Valores de referencia en parámetros de laboratorio

PRUEBA	LIR-LSR*	UNIDADES
Alanina Aminotransferasa (ALT/GPT) [Femenino]	1-35	U/L
Alanina Aminotransferasa (ALT/GPT) [Masculino]	1-50	U/L
Aspartato Aminotransferasa (AST/GOT) [Femenino]	0-35	U/L
Aspartato Aminotransferasa (AST/GOT) [Masculino]	0-50	U/L
Basófilos	0-0,2	$\times 10^3/\mu\text{L}$
Bicarbonato (CO_3H^-) (arterial)	21-28	mmol/L
Bilirrubina Total Suero (5-60 años)	0,3-1,2	mg/dL
Bilirrubina Total Suero (60-90 años)	0,2-1,1	mg/dL
Bilirrubina Total Suero (90-150 años)	0,2-0,9	mg/dL
Calcio iónico	4,5-5,2	mg/dL
Cloruro suero	101-109	mmol/L
CO_2 Total (TCO_2)	20-28	mmol/L
Creatinina	0,6-1,3	mg/dL
Dímero D	0-230	ng/mL
E. de bases en fluido extracel (BEecf) (a)	-3-3	mmol/L
Eosinófilos	0,1-0,65	$\times 10^3/\mu\text{L}$
Fibrinógeno derivado	180-420	mg/dL
Gammaglutamil transferasa (GGT) [Femenino]	0-38	U/L
Gammaglutamil transferasa (GGT) [Masculino]	0-55	U/L
Glucosa (14-60 años)	74-110	mg/dL
Glucosa (60-90 años)	82-110	mg/dL
Glucosa (90-150 años)	75-110	mg/dL
Hematíes [Femenino]	3,8-5,1	$\times 10^6/\mu\text{L}$
Hematíes [Masculino]	4,2-5,6	$\times 10^6/\mu\text{L}$
Hematocrito [Femenino]	34-45	%
Hematocrito [Masculino]	39-50	%
Hemoglobina Total [Femenino]	11,4-15,1	g/dL
Hemoglobina Total [Masculino]	13,2-16,8	U/L
Lactato deshidrogenasa (LDH) Suero (12-60 años)	100-190	U/L
Lactato deshidrogenasa (LDH) Suero (60-90 años)	110-210	U/L
Lactato deshidrogenasa (LDH) Suero (90-150 años)	99-284	U/L
Lactato (arterial)	0,7-2,1	mmol/L
Leucocitos	4-10,5	$\times 10^3/\mu\text{L}$
Linfocitos	0,8-5	$\times 10^3/\mu\text{L}$

Monocitos	0,2-1,1	x10 ³ /μL
Neutrófilos	1,5-7,5	x10 ³ /μL
pCO ₂ (arterial)	35-48	mmHg
pH	7,35-7,45	-
Plaquetas	150-350	x10 ³ /μL
pO ₂	83-108	mmHg
pO ₂ /FiO ₂	333-446	-
Procalcitonina	0-0,5	ng/mL
Proteína C reactiva suero	0-10	mg/L
Sodio suero	136-146	mmol/L
Tiempo de Cefalina (TTPA;APTT)	25-38	seg
TP Ratio	0,7-1,3	-
TTPA Ratio	0,8-1,3	-
Urea (18-60 años)	12,8-42,8	mg/dL
Urea (60-90 años)	17,1-49,2	mg/dL
Urea (90-150 años)	21,4-66,3	mg/dL
VCM	80-98	fL

LIR=límite inferior de referencia; LSR=límite superior de referencia

8.2. Tablas de resultados

8.2.1. Resultados relacionados con el estado previo de los pacientes

Tabla A1. Características clínico-patológicas asociadas al estado previo de los pacientes. Variables numéricas.

	1ª OLA			OLAS 2ª a 4ª			OLAS 5ª y 6ª		
	Fallecidos	Vivos	Global	Fallecidos	Vivos	Global	Fallecidos	Vivos	Global
	60	231	291	40	323	363	27	217	244
	M ± DE	M ± DE	M ± DE	M ± DE	M ± DE	M ± DE	M ± DE	M ± DE	M ± DE
Edad	77,13 ± 1,50	64,50 ± 1,10	67, 10 ± 10	79,53 ± 10,92	64,4 ± 16,18	66,08 ± 16,37	82,40 ± 10,30	60,71 ± 20,23	63,18 ± 20,54
Número de dosis de vacunación administradas	0,00	0,00	0,00	1,00 ± 0,00	1,29 ± 0,46	1,26 ± 0,45	1,94 ± 0,65	2,44 ± 0,51	2,01 ± 0,65

M = media aritmética; DE = desviación estándar

Anexos

Tabla A2. Características clínico-patológicas asociadas al estado previo de los pacientes. Variables categóricas.

	1ª OLA						OLAS 2ª a 4ª						OLAS 5ª y 6ª					
	Fallecidos		Vivos		Global		Fallecidos		Vivos		Global		Fallecidos		Vivos		Global	
	60		231		291		40		323		363		27		217		244	
	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)
Sexo																		
Hombres	35	58,3	121	52,4	156	53,6	33	82,5	198	61,3	231	63,6	18	66,7	127	58,5	145	59,4
Mujeres	25	41,7	110	47,6	135	46,4	7	17,5	125	38,7	132	36,4	9	33,3	90	41,5	99	40,6
Comorbilidades																		
Hipertensión	31	51,7	97	42,0	128	44,0	24	60,0	144	44,6	168	46,3	13	48,1	93	42,9	106	43,4
Diabetes mellitus	11	18,3	45	19,5	56	19,2	10	25,0	68	21,1	78	21,5	3	11,1	45	20,7	48	19,7
Cardiopatía isquémica	24	40,0	53	22,9	77	26,5	7	17,5	17	5,3	24	6,6	5	18,5	11	5,1	16	6,6
Insuficiencia cardíaca	1	1,7	7	3,0	8	2,7	1	2,5	6	1,9	7	1,9	2	7,4	7	3,2	9	3,7
EPOC (enfisema crónico/bronquitis)	7	11,7	18	7,8	25	8,6	1	2,5	17	5,3	18	5,0	8	29,6	17	7,8	25	10,2
Asma	3	5,0	23	10,0	26	8,9	1	2,5	17	5,3	18	5,0	1	3,7	11	5,1	12	4,9
Dislipemia	16	26,7	88	38,1	104	35,7	16	40,0	100	31,0	116	32,0	11	40,7	53	24,4	64	26,2
Coagulopatías	2	3,3	8	3,5	10	3,4	7	17,5	34	10,5	41	11,3	10	37,0	25	11,5	35	14,3
Insuficiencia renal	7	11,7	7	3,0	14	4,8	5	12,5	19	5,9	24	6,6	6	22,2	18	8,3	24	9,8
Neoplasia activa	5	8,3	13	5,6	18	6,2	1	2,5	16	5,0	17	4,7	7	25,9	28	12,9	35	14,3
Enfermedades inmunomediadas	6	10,0	14	6,1	20	6,9	2	5,0	38	11,8	40	11,0	5	18,5	20	9,2	25	10,2
Sobrepeso	2	3,3	13	5,6	15	5,2	1	2,5	4	1,2	5	1,4	0	0,0	1	0,5	1	0,4
Obesidad	6	10,0	12	5,2	18	6,2	3	7,5	30	9,3	33	9,1	1	3,7	21	9,7	22	9,0
VIH	0	0,0	1	0,4	1	0,3	0	0,0	0	0,0	0	0,0	0	0,0	2	0,9	2	0,8
Hepatitis C crónica	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	1	3,7	2	0,9	3	1,2
Tratamientos crónicos	44	73,3	165	71,4	209	71,8	32	80,0	183	56,7	215	59,2	24	88,9	117	53,9	141	57,8
Antihipertensivos	28	46,7	74	32,0	102	35,1	21	52,5	115	35,6	136	37,5	12	44,4	73	33,6	85	34,8
Betabloqueantes	12	20,0	26	11,3	38	13,1	10	25,0	38	11,8	48	13,2	7	25,9	32	14,7	39	16,0
Diuréticos	14	23,3	33	14,3	47	16,2	8	20,0	37	11,5	45	12,4	4	14,8	19	8,8	23	9,4
Antidiabéticos	20	33,3	35	15,2	55	18,9	8	20,0	58	18,0	66	18,2	6	22,2	41	18,9	47	19,3
Antiagregantes	11	18,3	20	8,7	31	10,7	11	27,5	33	10,2	44	12,1	4	14,8	27	12,4	31	12,7
Anticoagulantes	13	21,7	14	6,1	27	9,3	6	15,0	34	10,5	40	11,0	9	33,3	22	10,1	31	12,7
Hipolipemiantes	17	28,3	72	31,2	89	30,6	15	37,5	90	27,9	105	28,9	10	37,0	46	21,2	56	23,0
Quimioterapia	0	0,0	0	0,0	0	0,0	1	2,5	2	0,6	3	0,8	0	0,0	8	3,7	8	3,3
Inmunosupresores / inmunomoduladores	2	3,3	8	3,5	10	3,4	3	7,5	21	6,5	24	6,6	6	22,2	18	8,3	24	9,8
Antirretrovirales	0	0,0	0	0,0	0	0,0	0	0,0	1	0,3	1	0,3	0	0,0	0	0,0	0	0,0
Antivirales	0	0,0	0	0,0	0	0,0	0	0,0	1	0,3	1	0,3	0	0,0	0	0,0	0	0,0
Biológicos	0	0,0	1	0,4	1	0,3	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0
Al menos un episodio previo de COVID-19	0	0,0	0	0,0	0	0,0	1	2,5	2	0,6	3	0,8	1	3,7	2	0,9	3	1,2
Vacunación																		
1 dosis	0	0,0	0	0,0	0	0,0	2	5,0	15	4,6	17	4,7	12	44,4	78	35,9	90	36,9
2 dosis	0	0,0	0	0,0	0	0,0	0	0,0	6	1,9	6	1,7	12	44,4	78	35,9	90	36,9
Tipo de vacuna																		
Pfizer (Comirnaty)	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	12	44,4	63	29,0	75	30,7
Astrazeneca (Vaxzevria)	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	7	3,2	7	2,9
Moderna	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	5	2,3	5	2,0
Janssen	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	3	1,4	3	1,2

8.2.2. Resultados relacionados con la clínica de los pacientes al ingreso

Tabla A3. Características clínico-patológicas relacionadas con la sintomatología y diagnóstico clínico y por imagen de los pacientes con COVID-19 al ingreso hospitalario.

	1ª OLA						OLAS 2ª a 4ª						OLAS 5ª y 6ª					
	Fallecidos		Vivos		Global		Fallecidos		Vivos		Global		Fallecidos		Vivos		Global	
	60		231		291		40		323		363		27		217		244	
	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)
Síntomas al ingreso																		
Disnea	8	13,3	27	11,7	35	12,0	19	47,5	172	53,3	191	52,6	20	74,1	107	49,3	127	52,0
Presión torácica	1	1,7	21	9,1	22	7,6	1	2,5	44	13,6	45	12,4	1	3,7	40	18,4	41	16,8
Tos	40	66,7	141	61,0	181	62,2	21	52,5	202	62,5	223	61,4	12	44,4	127	58,5	139	57,0
Rinorrea	0	0,0	5	2,2	5	1,7	0	0,0	10	3,1	10	2,8	0	0,0	13	6,0	13	5,3
Anosmia	3	5,0	11	4,8	14	4,8	0	0,0	49	15,2	49	13,5	0	0,0	29	13,4	29	11,9
Ageusia	3	5,0	17	7,4	20	6,9	2	5,0	48	14,9	50	13,8	0	0,0	33	15,2	33	13,5
Odinofagia	0	0,0	0	0,0	0	0,0	0	0,0	12	3,7	12	3,3	0	0,0	14	6,5	14	5,7
Mialgia	10	16,7	29	12,6	39	13,4	4	10,0	57	17,6	61	16,8	0	0,0	20	9,2	20	8,2
Fiebre	17	28,3	61	26,4	78	26,8	26	65,0	225	69,7	251	69,1	13	48,1	139	64,1	152	62,3
Distermia	4	6,7	7	3,0	11	3,8	3	7,5	27	8,4	30	8,3	2	7,4	10	4,6	12	4,9
Cefalea	0	0,0	7	3,0	7	2,4	2	5,0	33	10,2	35	9,6	0	0,0	22	10,1	22	9,0
Náuseas/Vómitos	9	15,0	32	13,9	41	14,1	5	12,5	47	14,6	52	14,3	2	7,4	27	12,4	29	11,9
Diarrea	10	16,7	41	17,7	51	17,5	3	7,5	65	20,1	68	18,7	2	7,4	39	18,0	41	16,8
Astenia	17	28,3	87	37,7	104	35,7	7	17,5	101	31,3	108	29,8	3	11,1	64	29,5	67	27,5
Confusión	6	10,0	9	3,9	15	5,2	9	22,5	8	2,5	17	4,7	5	18,5	13	6,0	18	7,4
Mareo	2	3,3	7	3,0	9	3,1	2	5,0	19	5,9	21	5,8	1	3,7	15	6,9	16	6,6
Expectoración	8	13,3	27	11,7	35	12,0	6	15,0	45	13,9	51	14,0	5	18,5	33	15,2	38	15,6
Diagnóstico al ingreso																		
Distrés respiratorio	0	0,0	0	0,0	0	0,0	6	15,0	21	6,5	27	7,4	0	0,0	4	1,8	4	1,6
Insuficiencia respiratoria aguda	45	75,0	78	33,8	123	42,3	20	50,0	127	39,3	147	40,5	23	85,2	76	35,0	99	40,6
Fallo multiorgánico	0	0,0	0	0,0	0	0,0	2	5,0	0	0,0	2	0,6	0	0,0	1	0,5	1	0,4
Radiografía de tórax																		
Consolidaciones	21	35,0	53	22,9	74	25,4	8	20,0	55	17,0	63	17,4	8	29,6	52	24,0	60	24,6
Opacidades	31	51,7	141	61,0	172	59,1	27	67,5	267	82,7	294	81,0	17	63,0	160	73,7	177	72,5

Anexos

Tabla A4. Características clínico-patológicas de los pacientes de la 1ª ola al ingreso hospitalario. Variables numéricas.

	1ª OLA		
	Fallecidos	Vivos	Global
	60	231	291
	M ± DE	M ± DE	M ± DE
Radiografía tórax			
Nº zonas afectadas (consolidaciones)	2,85±1,78	2,48±1,64	3,68±1,81
Nº zonas afectadas (opacidades)	2,71±1,56	2,59±1,49	3,26±1,77
Escalas al ingreso			
Escala coma de Glasgow	14,28±2,44	14,84±0,63	14,73±1,26
Escala CURB-65	1,39±0,64	0,8±0,73	0,92±0,75
Escala APACHE-II	13,64±5,33	8,46±2,82	9,92±4,32
Escala SOFA	6,23±2,31	5,11±2,01	5,48±2,15
Signos vitales			
Temperatura (°C)	36,18 ± 0,7	36,9 ± 0,1	36,9 ± 0,1
Presión arterial sistólica / PAS	132,5±24,7	133,14±19,56	133±20,73
Presión arterial diastólica / PAD	71,87±16,93	75,04±12,36	74,35±13,5
Frecuencia cardíaca / FC	93,21±18,1	87,85±16,85	88,96±17,22
Frecuencia respiratoria / FR	25,38±7,29	22,48±5,71	23,16±6,21
Parámetros de laboratorio			
Lecucocitos (×103 µL)	9,5 ± 2	6,7 ± 0,2	6,2 ± 0,1
Linfocitos (×103 µL)	0,73 ± 0,05	1,10 ± 0,11	1,02 ± 0,09
Neutrófilos (×103 µL)	6,5 ± 0,6	5 ± 0,2	4,6 ± 0,1
Monocitos (×103 µL)	0,45 ± 0,04	0,53 ± 0,02	0,51 ± 0,02
Eosinófilos (×103 µL)	0,007 ± 0,004	0,009 ± 0,002	0,009 ± 0,002
Basófilos (×103 µL)	0,016 ± 0,005	0,013 ± 0,002	0,014 ± 0,002
Eritrocitos (×106 µL)	4,6 ± 0,1	5 ± 0,2	4,7 ± 0
Hemoglobina (g/dL)	13,3 ± 0,2	13,8 ± 0,1	13,8 ± 0,1
Hematocrito (%)	40,5 ± 0,7	41,8 ± 0,3	41,5 ± 0,3
V.C.M. (fL)	88,8 ± 0,9	88,1 ± 0,3	88,3 ± 0,3
Plaquetas (×103 µL)	181,9 ± 10	191,8 ± 5,5	183 ± 4,1
Glucosa (mg/dL)	144,9 ± 6,5	126,8 ± 3,6	117 ± 1,4
Urea (mg/dL)	57,9 ± 3,9	44,3 ± 2,4	40,5 ± 1,1
Creatinina (mg/dL)	1,2 ± 0,1	1,7 ± 0,5	0,9 ± 0
Filtrado glomerular estimado (CKD-EPI 2009) (mL/min/1.73 m2)	57,7 ± 2,8	69,4 ± 1,5	66,6 ± 1,3
Sodio (mmol/L)	134,5 ± 0,6	134,6 ± 0,7	135 ± 0,2
Potasio (mmol/L)	4 ± 0,1	5,7 ± 1,8	3,9 ± 0
Cloruro (mmol/L)	100,4 ± 1,1	108,5 ± 6,5	99,9 ± 0,3
Bilirrubina total (mg/dL)	0,7 ± 0	0,9 ± 0,3	0,6 ± 0
Aspartato aminotransferasa (AST/GOT) (U/L)	34,9 ± 3,8	39,9 ± 2,2	29,7 ± 1
Alanina aminotransferasa (ALT/GPT) (U/L)	60,3 ± 7,2	48,4 ± 2,1	40,3 ± 1,1
Lactato deshidrogenasa (LDH) (U/L)	88 ± 42	107,5 ± 66,3	101 ± 43,5
Albumina (g/dL)	3,54±0,4	3,79±0,44	3,74±0,44
Proteína C reactiva (mg/dL)	136,7 ± 14,5	95,9 ± 5,7	93 ± 4,5
Procalcitonina (ng/mL)	15,3 ± 13,1	0,4 ± 0,1	0,1 ± 0
Dímero D (ng/mL)	1403 ± 610,6	647,4 ± 69,1	459 ± 16,9
Fibrinógeno (derivado)(mg/dL)	703,7 ± 22,1	698,1 ± 11,5	699 ± 10,2
Ratio (TP)	1,6 ± 0,2	1,2 ± 0	1,1 ± 0
Ratio (TTPA)	1,09 ± 0,01	1,04 ± 0,01	1,05 ± 0,01
pH	7,426 ± 0,008	7,446 ± 0,004	7,442 ± 0,003
pCO2 (mmHg)	35,26±8,06	34,38±7,63	34,56±7,71
pO2 (mmHg)	67,76±33,88	62,02±22,54	63,25±25,43
Bicarbonato (CO3H-) (mmol/L)	22,77±4,13	24,19±6	23,88±5,67
FIO2 (%)	33,8 ± 4,1	24 ± 1	21 ± 0
pO2/FIO2	227,4 ± 19,1	270,8 ± 7,9	262 ± 7,5
Gradiente O2 arterial/alveolar (mmHg)	150,28±185,13	71,97±83,61	89,37±118,09
Lactato (mmol/L)	2 ± 0,1	1,4 ± 0	1,5 ± 0

M = media aritmética; DE = desviación estándar

Anexos

Tabla A5. Características clínico-patológicas de los pacientes de las olas 2ª a 4ª al ingreso hospitalario. Variables numéricas.

	OLAS 2ª a 4ª		
	Fallecidos	Vivos	Global
	40	323	363
	M ± DE	M ± DE	M ± DE
Radiografía tórax			
Nº zonas afectadas (consolidaciones)	2,57 ± 1,27	2,47 ± 1,55	2,48 ± 1,51
Nº zonas afectadas (opacidades)	3,41 ± 1,76	3,18 ± 1,69	3,2 ± 1,69
Escalas al ingreso			
Escala coma de Glasgow	14,43 ± 1,32	14,86 ± 0,78	14,81 ± 0,87
Escala CURB-65	1,63 ± 0,81	0,74 ± 0,81	0,84 ± 0,86
Escala APACHE-II	12,8 ± 4,69	8,31 ± 3,23	9,53 ± 4,16
Escala SOFA	4,45 ± 1,23	4,07 ± 1,63	4,17 ± 1,54
Signos vitales			
Temperatura (°C)	36,87 ± 1,08	36,7 ± 0,94	36,72 ± 0,96
Presión arterial sistólica / PAS	135,41 ± 23,07	131,94 ± 22,62	132,32 ± 22,66
Presión arterial diastólica / PAD	63,31 ± 16,1	71,92 ± 16,58	70,98 ± 16,72
Frecuencia cardíaca / FC	87,56 ± 19,81	90,9 ± 17,51	90,54 ± 17,78
Frecuencia respiratoria / FR	25,32 ± 7,67	23,01 ± 6,95	23,34 ± 7,07
Parámetros de laboratorio			
Lecucocitos (×103 µL)	7,99 ± 5,12	6,96 ± 4,86	7,07 ± 4,9
Linfocitos (×103 µL)	0,71 ± 0,36	1,53 ± 5,6	1,43 ± 5,28
Neutrófilos (×103 µL)	6,42 ± 5,23	5,05 ± 3,8	5,2 ± 4
Monocitos (×103 µL)	0,5 ± 0,29	0,61 ± 0,81	0,6 ± 0,77
Eosinófilos (×103 µL)	0,03 ± 0,09	0,02 ± 0,07	0,02 ± 0,07
Basófilos (×103 µL)	0,14 ± 0,74	0,02 ± 0,04	0,03 ± 0,25
Eritrocitos (×106 µL)	4,4 ± 0,68	4,72 ± 0,7	4,69 ± 0,7
Hemoglobina (g/dL)	13,41 ± 1,87	13,79 ± 1,83	13,75 ± 1,83
Hematocrito (%)	39,62 ± 7	41,52 ± 5,36	41,3 ± 5,59
V.C.M. (fL)	90,78 ± 9,99	87,81 ± 7,25	88,15 ± 7,64
Plaquetas (×103 µL)	163,52 ± 80,76	186,62 ± 79	184,02 ± 79,42
Glucosa (mg/dL)	144,83 ± 62,46	138,84 ± 62,98	139,5 ± 62,86
Urea (mg/dL)	71,72 ± 40,2	43,57 ± 26,37	46,77 ± 29,6
Creatinina (mg/dL)	1,37 ± 0,76	1,37 ± 5,33	1,37 ± 5,02
Filtrado glomerular estimado (CKD-EPI 2009) (mL/min/1.73 m2)	56,88 ± 23,86	73,79 ± 20,22	71,84 ± 21,33
Sodio (mmol/L)	135,85 ± 4,04	134,91 ± 3,39	135,02 ± 3,48
Potasio (mmol/L)	4,14 ± 0,58	4,2 ± 2,82	4,19 ± 2,66
Cloruro (mmol/L)	101,85 ± 3,77	101,13 ± 5,7	101,21 ± 5,52
Bilirrubina total (mg/dL)	0,76 ± 0,42	1,64 ± 9,6	1,54 ± 9,05
Aspartato aminotransferasa (AST/GOT) (U/L)	72,77 ± 100,68	48,48 ± 36,15	51,2 ± 48,22
Alanina aminotransferasa (ALT/GPT) (U/L)	51,83 ± 75,99	42,17 ± 40,45	43,21 ± 45,59
Lactato deshidrogenasa (LDH) (U/L)	411,46 ± 201,74	325,42 ± 128,12	335,4 ± 140,97
Albúmina (g/dL)	3,53 ± 0,4	3,82 ± 0,38	3,79 ± 0,39
Proteína C reactiva (mg/dL)	115,47 ± 70,71	78,23 ± 69,85	82,22 ± 70,79
Procalcitonina (ng/mL)	0,81 ± 2,93	0,25 ± 0,92	0,31 ± 1,3
Dímero D (ng/mL)	1500,51 ± 2568,3	1155 ± 2318,82	1192,1 ± 2345,08
Fibrinógeno (derivado)(mg/dL)	688,53 ± 171,63	667,34 ± 161,7	669,71 ± 162,71
Ratio (TP)	2,03 ± 4,62	1,9 ± 7,06	1,91 ± 6,83
Ratio (TTPA)	29,83 ± 3,87	30,78 ± 5,24	30,67 ± 5,11
pH	7,41 ± 0,07	7,43 ± 0,07	7,42 ± 0,07
pCO2 (mmHg)	36,91 ± 9,16	34,75 ± 7,05	35,05 ± 7,39
pO2 (mmHg)	64,2 ± 24,83	60,8 ± 23,28	61,27 ± 23,48
Bicarbonato (CO3H-) (mmol/L)	23,01 ± 4,61	23,07 ± 3,1	23,06 ± 3,34
FIO2 (%)	26,86 ± 10,08	24,18 ± 11,32	24,57 ± 11,16
pO2/FIO2	256,1 ± 86,47	279,83 ± 109,36	276,21 ± 106,32
Gradiente O2 arterial/alveolar (mmHg)	72,19 ± 55,01	65,13 ± 76,85	66,11 ± 74,11
Lactato (mmol/L)	1,89 ± 1,03	1,72 ± 1,16	1,74 ± 1,15

M = media aritmética; DE = desviación estándar

Anexos

Tabla A6. Características clínico-patológicas de los pacientes de las olas 5ª y 6ª al ingreso hospitalario. Variables numéricas.

	OLAS 5ª y 6ª		
	Fallecidos	Vivos	Global
	27	217	244
	M ± DE	M ± DE	M ± DE
Radiografía tórax			
Nº zonas afectadas (consolidaciones)	2,38 ± 1,77	2,12 ± 1,54	2,15 ± 1,56
Nº zonas afectadas (opacidades)	3,53 ± 1,7	3,13 ± 1,72	3,17 ± 1,71
Escalas al ingreso			
Escala coma de Glasgow	14,19 ± 1,83	14,92 ± 0,36	14,84 ± 0,72
Escala CURB-65	1,77 ± 0,91	0,81 ± 1,6	0,92 ± 1,57
Escala APACHE-II	11 ± 7,07	7,31 ± 2,98	7,8 ± 3,59
Escala SOFA	6,67 ± 1,15	3,85 ± 1,82	4,38 ± 2,03
Signos vitales			
Temperatura (°C)	36,5 ± 0,84	36,55 ± 2,56	36,54 ± 2,42
Presión arterial sistólica / PAS	124,92 ± 22,8	131,74 ± 33,21	131,01 ± 32,28
Presión arterial diastólica / PAD	66,44 ± 14,56	67,6 ± 23,88	67,48 ± 23,06
Frecuencia cardíaca / FC	83,56 ± 13,7	89,44 ± 18,37	88,78 ± 17,98
Frecuencia respiratoria / FR	30,17 ± 12,5	23,08 ± 6,25	24,5 ± 8,31
Parámetros de laboratorio			
Lecucocitos (×103 µL)	8,99 ± 5,57	7,18 ± 3,72	7,37 ± 3,99
Linfocitos (×103 µL)	1,22 ± 1,48	1,23 ± 1,29	1,23 ± 1,31
Neutrófilos (×103 µL)	7,18 ± 5,22	5,31 ± 3,06	5,51 ± 3,38
Monocitos (×103 µL)	0,52 ± 0,33	0,59 ± 0,43	0,59 ± 0,42
Eosinófilos (×103 µL)	0 ± 0,02	0,04 ± 0,08	0,03 ± 0,08
Basófilos (×103 µL)	0,03 ± 0,05	0,02 ± 0,05	0,02 ± 0,05
Eritrocitos (×106 µL)	4,48 ± 0,87	4,59 ± 0,72	4,57 ± 0,74
Hemoglobina (g/dL)	13,27 ± 2,03	13,48 ± 1,86	13,46 ± 1,87
Hematocrito (%)	40,16 ± 6,21	40,01 ± 5,48	40,02 ± 5,55
V.C.M. (fL)	90,63 ± 9,05	87,54 ± 6,62	87,87 ± 6,97
Plaquetas (×103 µL)	189,69 ± 75,7	182,81 ± 68,75	183,55 ± 69,4
Glucosa (mg/dL)	144,81 ± 43,03	134,93 ± 60,07	136,04 ± 58,41
Urea (mg/dL)	81,69 ± 53,77	42,07 ± 28,1	46,32 ± 33,96
Creatinina (mg/dL)	1,33 ± 0,74	1,01 ± 0,51	1,04 ± 0,55
Filtrado glomerular estimado (CKD-EPI 2009) (mL/min/1.73 m2)	54,74 ± 25,93	74,37 ± 21,26	72,14 ± 22,65
Sodio (mmol/L)	136,96 ± 8,04	135,51 ± 4,23	135,68 ± 4,81
Potasio (mmol/L)	4,18 ± 0,48	4,02 ± 0,61	4,04 ± 0,6
Cloruro (mmol/L)	97,33 ± 22,95	101,3 ± 5,21	100,89 ± 8,88
Bilirrubina total (mg/dL)	0,6 ± 0,26	0,91 ± 3,93	0,87 ± 3,71
Aspartato aminotransferasa (AST/GOT) (U/L)	51,27 ± 32,62	46 ± 33,94	46,51 ± 33,78
Alanina aminotransferasa (ALT/GPT) (U/L)	27,44 ± 22,51	42,22 ± 43,29	40,6 ± 41,75
Lactato deshidrogenasa (LDH) (U/L)	362,36 ± 137,61	315,26 ± 119,25	320,15 ± 121,78
Albúmina (g/dL)	3,43 ± 0,45	3,83 ± 0,38	3,79 ± 0,41
Proteína C reactiva (mg/dL)	132,36 ± 104,82	84,65 ± 77,36	89,79 ± 81,84
Procalcitonina (ng/mL)	1,79 ± 6,27	0,93 ± 6,8	1,02 ± 6,73
Dímero D (ng/mL)	2275,38 ± 3361,11	1333,48 ± 3535,04	1434,39 ± 3521,65
Fibrinógeno (derivado)(mg/dL)	637,36 ± 255,76	685,83 ± 180,39	680,65 ± 189,77
Ratio (TP)	1,59 ± 1,5	1,23 ± 0,57	1,27 ± 0,73
Ratio (TTPA)	32,24 ± 7,18	32,25 ± 4,92	32,24 ± 5,19
pH	7,42 ± 0,06	7,44 ± 0,06	7,44 ± 0,06
pCO2 (mmHg)	37,41 ± 10,08	35,67 ± 6,9	35,92 ± 7,41
pO2 (mmHg)	74,55 ± 34,24	72,61 ± 30,04	72,88 ± 30,55
Bicarbonato (CO3H-) (mmol/L)	25 ± 5,64	24,02 ± 3,34	24,16 ± 3,76
FIO2 (%)	33,53 ± 28,19	22,83 ± 6,69	24,33 ± 12,61
pO2/FIO2	302,8 ± 152,61	325,14 ± 98,96	321,95 ± 107,8
Gradiente O2 arterial/alveolar (mmHg)	87,27 ± 159,33	43,26 ± 48,12	48,63 ± 71,78
Lactato (mmol/L)	1,78 ± 1,02	1,46 ± 0,66	1,5 ± 0,73

M = media aritmética; DE = desviación estándar

8.2.3. Resultados relacionados con la estancia hospitalaria.

Tabla A7. Características clínico-patológicas relacionadas con la estancia hospitalaria de los pacientes

	1ª OLA						OLAS 2ª a 4ª						OLAS 5ª y 6ª					
	Fallecidos		Vivos		Global		Fallecidos		Vivos		Global		Fallecidos		Vivos		Global	
	60		231		291		40		323		363		27		217		244	
	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)
Ingreso en UVI	6	10,0	59	25,5	65	22,3	21	52,5	61	19,0	82	23,0	3	11,1	20	9,2	23	9,8
VMI	26	43,3	39	16,9	65	22,3	19	47,5	56	17,0	75	21,0	2	7,4	22	10,1	24	9,8
CNAF	2	3,3	11	4,8	13	4,5	4	10,0	29	9,0	33	9,0	6	22,2	14	6,5	20	8,2
VMNI	3	5,0	4	1,7	7	2,4	12	30,0	30	9,0	42	12,0	2	7,4	6	2,8	8	3,3
Sesiones de prono	6	10,0	5	2,2	11	3,8	12	30,0	28	9,0	40	11,0	1	3,7	9	4,1	10	4,1
ECMO	2	3,3	3	1,3	5	1,7	1	3,0	4	1,0	5	1,0	0	0,0	2	0,9	2	0,8
Infecciones nosocomiales													0,0					
Viral	0	0,0	0	0,0	0	0,0	0	0,0	2	1,0	2	1,0	0	0,0	0	0,0	0	0,0
Bacteriana	11	18,3	30	13,0	41	14,1	11	28,0	30	9,0	41	11,0	1	3,7	16	7,4	17	7,0
Fúngica	4	6,7	9	3,9	13	4,5	3	8,0	3	1,0	6	2,0	1	3,7	0	0,0	1	0,4
Complicaciones durante la estancia hospitalaria																		
Fallo renal agudo	12	20,0	11	4,8	23	7,9	13	33,0	16	5,0	29	8,0	6	22,2	13	6,0	19	7,8
Cardíacas	5	8,3	11	4,8	16	5,5	9	23,0	15	5,0	24	7,0	0	0,0	4	1,8	4	1,6
Gastrointestinales	6	10,0	11	4,8	17	5,8	5	13,0	22	7,0	27	7,0	2	7,4	11	5,1	13	5,3
Neurológicas	2	3,3	17	7,4	19	6,5	3	8,0	19	6,0	22	6,0	1	3,7	8	3,7	9	3,7
Coagulopatías	3	5,0	13	5,6	16	5,5	9	23,0	14	4,0	23	6,0	3	11,1	10	4,6	13	5,3
SDRA	18	30,0	35	15,2	53	18,2	15	38,0	36	11,0	51	14,0	2	7,4	14	6,5	16	6,6
Shock	6	10,0	1	0,4	7	2,4	5	13,0	1	0,3	6	2,0	2	7,4	1	0,5	3	1,2
Tratamiento para COVID durante la estancia hospitalaria																		
Oxígeno	52	86,7	127	55,0	179	61,5	32	80,0	173	53,6	205	56,5	24	88,9	117	53,9	141	57,8
Dexametasona	0	0,0	5	2,2	5	1,7	23	58,0	130	40,0	153	42,0	13	48,1	105	48,4	118	48,4
Prednisona	0	0,0	0	0,0	0	0,0	23	58,0	130	40,0	153	42,0	1	3,7	5	2,3	6	2,5
Rendesivir	0	0,0	0	0,0	0	0,0	2	5,0	26	8,0	28	8,0	0	0,0	0	0,0	0	0,0
Lopinavir/ritonavir	0	0,0	67	29,0	67	23,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0
Hidroxiclороquina	26	43,3	176	76,2	202	69,4	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0
Tocilizumab	1	1,7	13	5,6	14	4,8	1	3,0	6	2,0	7	2,0	1	3,7	2	0,9	3	1,2
Baricitinib	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0
Anakinra	2	3,3	1	0,4	3	1,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0
IFN β-1b	4	6,7	7	3,0	11	3,8	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0	0	0,0
Ceftriaxona	26	43,3	151	65,4	177	60,8	21	53,0	142	44,0	163	45,0	11	40,7	61	28,1	72	29,5
Azitromicina	28	46,7	167	72,3	195	67,0	15	38,0	111	34,0	126	35,0	3	11,1	33	15,2	36	14,8
Heparina	0	0,0	1	0,4	1	0,3	8	20,0	112	35,0	120	33,0	6	22,2	103	47,5	109	44,7
Causa del éxito (si aplica)																		
Fallo multiorgánico	29	48,3	0	0,0	29	10,0	17	43,0	0	0,0	17	5,0	0	0,0	0	0,0	0	0,0
Distrés respiratorio	4	6,7	0	0,0	4	1,4	1	3,0	0	0,0	1	0,3	0	0,0	0	0,0	0	0,0
Insuficiencia respiratoria	17	28,3	0	0,0	17	5,8	11	28,0	0	0,0	11	3,0	23	85,2	0	0,0	23	9,4

Anexos

Tabla A8. Días de estancia hospitalaria en cada grupo de pacientes analizado.

Días de estancia hospitalaria		
M ± DE		
1ª OLA	Fallecidos	16,71±23,77
	Vivos	18,3±21,02
	Global	18,01±21,59
OLAS 2ª a 4ª	Fallecidos	20,59 ± 17,44
	Vivos	14,92 ± 18,34
	Global	15,54 ± 18,3
OLAS 5ª y 6ª	Fallecidos	10,32 ± 15,22
	Vivos	23,22 ± 53,76
	Global	11,76 ± 23,11

M = media aritmética; DE = desviación estándar

9. Rendimiento científico

Article

Approaching Personalized Medicine: The Use of Machine Learning to Determine Predictors of Mortality in a Population with SARS-CoV-2 Infection

Mónica Queipo ^{1,2}, Julia Barbado ^{1,2,3}, Ana María Torres ^{4,5} and Jorge Mateo ^{4,5,*}

¹ Autoimmunity and Inflammation Research Group, Río Hortega University Hospital, 47012 Valladolid, Spain

² Cooperative Research Network Focused on Health Results—Advanced Therapies (RICORS TERAV), 28220 Madrid, Spain

³ Internal Medicine, Río Hortega University Hospital, 47012 Valladolid, Spain

⁴ Medical Analysis Expert Group, Institute of Technology, University of Castilla-La Mancha, 16071 Cuenca, Spain

⁵ Medical Analysis Expert Group, Instituto de Investigación Sanitaria de Castilla-La Mancha (IDISCAM), 45071 Toledo, Spain

* Correspondence: jorge.mateo@uclm.es

Abstract: The COVID-19 pandemic demonstrated the need to develop strategies to control a new viral infection. However, the different characteristics of the health system and population of each country and hospital would require the implementation of self-systems adapted to their characteristics. The objective of this work was to determine predictors that should identify the most severe patients with COVID-19 infection. Given the poor situation of the hospitals in the first wave, the analysis of the data from that period with an accurate and fast technique can be an important contribution. In this regard, machine learning is able to objectively analyze data in hourly sets and is used in many fields. This study included 291 patients admitted to a hospital in Spain during the first three months of the pandemic. After screening seventy-one features with machine learning methods, the variables with the greatest influence on predicting mortality in this population were lymphocyte count, urea, FiO₂, potassium, and serum pH. The XGB method achieved the highest accuracy, with a precision of >95%. Our study shows that the machine learning-based system can identify patterns and, thus, create a tool to help hospitals classify patients according to their severity of illness in order to optimize admission.

Keywords: COVID-19; mortality; predictors; risk factors; machine learning

Citation: Queipo, M.; Barbado, J.; Torres, A.M.; Mateo, J. Approaching Personalized Medicine: The Use of Machine Learning to Determine Predictors of Mortality in a Population with SARS-CoV-2 Infection. *Biomedicines* **2024**, *12*, 409. <https://doi.org/10.3390/biomedicines12020409>

Academic Editors: Monika Głodowska-Sawczuk and Monika Zajkowska

Received: 28 December 2023

Revised: 5 February 2024

Accepted: 7 February 2024

Published: 9 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Since the first cases were reported in 31 December 2019, the COVID-19 pandemic has accumulated a total of 770,875,433 confirmed cases and 6,959,316 deaths [1]. The uncontrolled spread of the virus caused by overpopulation, globalization, hyperconnectivity, and the centralization of supply chains [2] triggered a collapse of health services and resources that forced countries to take severe social actions such as isolation or lockdowns, causing serious social and economic consequences [3,4]. In Spain, health centers in the most affected areas faced problems such as inadequate intensive care capacity, insufficient equipment (both for patients and health workers), lack of medical staff, or the delay or collapse of COVID-19 helplines, which led to the cancellation of non-urgent surgeries and the need to use private health services and military facilities for public purposes [5]. In the region of Castilla y León (Spain), 585 deaths due to COVID-19 were recorded in March 2020, the first month of the pandemic in Spain, although this figure probably underestimates the actual number of deaths due to the disease during this period. In addition, the

Article

The Effect of Naturally Acquired Immunity on Mortality Predictors: A Focus on Individuals with New Coronavirus

Mónica Queipo ^{1,2}, Jorge Mateo ^{3,4,*}, Ana María Torres ^{3,4} and Julia Barbado ^{1,2,5,*} 

¹ Autoimmunity and Inflammation Research Group, Río Hortega University Hospital, 47012 Valladolid, Spain; mqueipo@saludcastillayleon.es

² Cooperative Research Network Focused on Health Results—Advanced Therapies (RICORS TERAV), 28220 Madrid, Spain

³ Medical Analysis Expert Group, Institute of Technology, University of Castilla-La Mancha, 13001 Cuenca, Spain

⁴ Medical Analysis Expert Group, Instituto de Investigación Sanitaria de Castilla-La Mancha (IDISCAM), 45071 Toledo, Spain

⁵ Internal Medicine, Río Hortega University Hospital, 47012 Valladolid, Spain

* Correspondence: jorge.mateo@uclm.es (J.M.); jbarbado@saludcastillayleon.es (J.B.)

Abstract Background/Objectives: The spread of the COVID-19 pandemic has spurred the development of advanced healthcare tools to effectively manage patient outcomes. This study aims to identify key predictors of mortality in hospitalized patients with some level of natural immunity, but not yet vaccinated, using machine learning techniques. **Methods:** A total of 363 patients with COVID-19 admitted to Río Hortega University Hospital in Spain between the second and fourth waves of the pandemic were included in this study. Key characteristics related to both the patient's previous status and hospital stay were screened using the Random Forest (RF) machine learning technique. **Results:** Of the 19 variables identified as having the greatest influence on predicting mortality, the most powerful ones could be identified at the time of hospital admission. These included the assessment of severity in community-acquired pneumonia (CURB-65) scale, age, the Glasgow Coma Scale (GCS), and comorbidities, as well as laboratory results. Some variables associated with hospitalization and intensive care unit (ICU) admission (acute renal failure, shock, PRONO sessions and the Acute Physiology and Chronic Health Evaluation [APACHE-II] scale) showed a certain degree of significance. The Random Forest (RF) method showed high accuracy, with a precision of >95%. **Conclusions:** This study shows that natural immunity generates significant changes in the evolution of the disease. As has been shown, machine learning models are an effective tool to improve personalized patient care in different periods.

Keywords: immunity; COVID-19; predictors; mortality; machine learning; Random Forest



Academic Editors: Monika Gudowska-Sawczuk and Monika Zafkowska

Received: 18 February 2025

Revised: 19 March 2025

Accepted: 24 March 2025

Published: 27 March 2025

Citation: Queipo, M.; Mateo, J.; Torres, A.M.; Barbado, J. The Effect of Naturally Acquired Immunity on Mortality Predictors: A Focus on Individuals with New Coronavirus. *Biomedicines* **2025**, *13*, 803. <https://doi.org/10.3390/biomedicines13040803>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

When a sufficient number of individuals are immunized against a pathogen, the probability of transmission between the infected and susceptible population decreases due to the interruption of the chain of transmission [1]. In the case of a contagious, direct contact-transferable pathogen that induces long-term immunity, this is maximized if the population has a random pattern of interaction [2]. Diseases such as rubella, measles or pertussis are under control, and others such as smallpox are even eradicated today thanks to herd immunity, either by natural means or through vaccination strategies [3].

10. Referencias

Referencias

1. Santos-López G, Cortés-Hernández P, Vallejo-Ruiz V, Reyes-Leyva J. SARS-CoV-2: generalidades, origen y avances en el tratamiento. *GMM*. 2021 Feb 25;157(1):4792.
2. Tyrrell DAJ, Bynoe ML. Cultivation of a Novel Type of Common-cold Virus in Organ Cultures. *BMJ*. 1965 Jun 5;1(5448):1467–70.
3. Gaunt ER, Hardie A, Claas ECJ, Simmonds P, Templeton KE. Epidemiology and Clinical Presentations of the Four Human Coronaviruses 229E, HKU1, NL63, and OC43 Detected over 3 Years Using a Novel Multiplex Real-Time PCR Method. *J Clin Microbiol*. 2010 Aug;48(8):2940–7.
4. Wilder-Smith A, Chiew CJ, Lee VJ. Can we contain the COVID-19 outbreak with the same measures as for SARS? *The Lancet Infectious Diseases*. 2020 May;20(5):e102–7.
5. MERS Situation Update [Internet]. 2020 [cited 2025 Feb 3]. Available from: <https://applications.emro.who.int/docs/EMCSR254E.pdf?ua=1&ua=1>
6. WHO-convened global study of origins of SARS-CoV-2: China Part. Joint WHO-China study: 14 January - 10 February 2021 [Internet]. World Health Organization (WHO); 2021 Mar [cited 2025 Feb 5] p. 120. Available from: <https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part>
7. Plowright RK, Parrish CR, McCallum H, Hudson PJ, Ko AI, Graham AL, et al. Pathways to zoonotic spillover. *Nat Rev Microbiol*. 2017 Aug;15(8):502–10.
8. Zoonoses [Internet]. World Health Organization (WHO); 2020 Jul [cited 2025 Feb 5]. Available from: <https://www.who.int/news-room/fact-sheets/detail/zoonoses>
9. Fisher CR, Streicker DG, Schnell MJ. The spread and evolution of rabies virus: conquering new frontiers. *Nat Rev Microbiol*. 2018 Apr;16(4):241–55.
10. Elsohaby I, Villa L. Zoonotic diseases: understanding the risks and mitigating the threats. *BMC Vet Res*. 2023 Oct 3;19(1):186, s12917-023-03736–8.
11. Hu D, Zhu C, Ai L, He T, Wang Y, Ye F, et al. Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. *Emerging Microbes & Infections*. 2018 Dec;7(1):1–10.
12. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020 Mar 12;579(7798):270–3.
13. Xiao K, Zhai J, Feng Y, Zhou N, Zhang X, Zou JJ, et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature*. 2020 Jul 9;583(7815):286–9.
14. Hamdy ME, El Deeb AH, Hagag NM, Shahein MA, Alaidi O, Hussein HA. Interspecies transmission of SARS CoV-2 with special emphasis on viral mutations and ACE-2 receptor homology roles. *International Journal of Veterinary Science and Medicine*. 2023 Dec 31;11(1):55–86.
15. Dyer O. Covid-19: WHO says laboratory escape theory is “extremely unlikely” after mission to China. *BMJ*. 2021 Feb 11;n428.
16. Holmes EC, Goldstein SA, Rasmussen AL, Robertson DL, Crits-Christoph A, Wertheim JO, et al. The origins of SARS-CoV-2: A critical review. *Cell*. 2021 Sep;184(19):4848–56.

17. Cucinotta D, Vanelli M. WHO Declares COVID-19 a Pandemic. *Acta Bio Medica Atenei Parmensis*. 2020 Mar 19;91(1):157–60.
18. WHO COVID-19 dashboard [Internet]. World Health Organization (WHO); 2020 Mar [cited 2025 May 5]. Available from: <https://data.who.int/dashboards/covid19>
19. Unruh L, Allin S, Marchildon G, Burke S, Barry S, Siersbaek R, et al. A comparison of 2020 health policy responses to the COVID-19 pandemic in Canada, Ireland, the United Kingdom and the United States of America. *Health Policy*. 2022 May;126(5):427–37.
20. Rocha R, Atun R, Massuda A, Rache B, Spinola P, Nunes L, et al. Effect of socioeconomic inequalities and vulnerabilities on health-system preparedness and response to COVID-19 in Brazil: a comprehensive analysis. *The Lancet Global Health*. 2021 Jun;9(6):e782–92.
21. Bhadra A, Mukherjee A, Sarkar K. Impact of population density on Covid-19 infected and mortality rate in India. *Model Earth Syst Environ*. 2021 Mar;7(1):623–9.
22. Jassat W, Cohen C, Tempia S, Masha M, Goldstein S, Kufa T, et al. Risk factors for COVID-19-related in-hospital mortality in a high HIV and tuberculosis prevalence setting in South Africa: a cohort study. *The Lancet HIV*. 2021 Sep;8(9):e554–67.
23. Alene KA, Wangdi K, Clements ACA. Impact of the COVID-19 Pandemic on Tuberculosis Control: An Overview. *TropicalMed*. 2020 Jul 24;5(3):123.
24. Peng F, Tu L, Yang Y, Hu P, Wang R, Hu Q, et al. Management and Treatment of COVID-19: The Chinese Experience. *Canadian Journal of Cardiology*. 2020 Jun;36(6):915–30.
25. Maier BF, Brockmann D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*. 2020 May 15;368(6492):742–6.
26. Thomson S, Ip EC. COVID-19 emergency measures and the impending authoritarian pandemic. *Journal of Law and the Biosciences*. 2020 Jul 25;7(1):lsaa064.
27. Boccia S, Ricciardi W, Ioannidis JPA. What Other Countries Can Learn From Italy During the COVID-19 Pandemic. *JAMA Intern Med*. 2020 Jul 1;180(7):927.
28. Coccia M. The impact of first and second wave of the COVID-19 pandemic in society: comparative analysis to support control measures to cope with negative effects of future infectious diseases. *Environmental Research*. 2021 Jun;197:111099.
29. Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations [Internet]. World Health Organization (WHO); 2020 Mar. Available from: <https://www.who.int/news-room/commentaries/detail/modes-of-transmission-of-virus-causing-Covid-19-implications-for-ipc-precaution-recommendations>
30. Pitol AK, Julian TR. Community Transmission of SARS-CoV-2 by Surfaces: Risks and Risk Reduction Strategies. *Environ Sci Technol Lett*. 2021 Mar 9;8(3):263–9.
31. Murphy C, Lim WW, Mills C, Wong JY, Chen D, Xie Y, et al. Effectiveness of social distancing measures and lockdowns for reducing transmission of COVID-19 in non-healthcare, community-based settings. *Phil Trans R Soc A*. 2023 Oct 9;381(2257):20230132.

32. Grépin KA, Aston J, Burns J. Effectiveness of international border control measures during the COVID-19 pandemic: a narrative synthesis of published systematic reviews. *Phil Trans R Soc A*. 2023 Oct 9;381(2257):20230134.
33. Kamyshnyi A, Krynytska I, Matskevych V, Marushchak M, Lushchak O. Arterial Hypertension as a Risk Comorbidity Associated with COVID-19 Pathology. Soma M, editor. *International Journal of Hypertension*. 2020 Dec 3;2020:1–7.
34. Karandere F. Study Hypothesis: Age, Gender, Presence of Diabetes Mellitus or Hypertension, and Anti-Hypertensive Drugs Are Independent Risk Factors for COVID-19 Mortality. ACC [Internet]. 2023 [cited 2025 Feb 9]; Available from: <https://hrcak.srce.hr/clanak/455020>
35. Liu JW, Huang X, Wang MK, Yang JS. Diabetes and susceptibility to COVID-19: Risk factors and preventive and therapeutic strategies. *World J Diabetes*. 2024 Aug 15;15(8):1663–71.
36. Qadar SMZ, Naz H, Shamim S, Hashim F, Ahmed S, Kumar Mehraj S. Prevalence of Obesity and its Effects in Patients With COVID-19: A Systematic Review and Meta-analysis. *Hosp Pharm*. 2024 Jun;59(3):341–8.
37. Widhani A, Koesnoe S, Maria S, Widjanarko AL, Karjadi TH, Hasibuan AS, et al. Factors Related to Severity, Hospitalization, and Mortality of COVID-19 Infection among Patients with Autoimmune Diseases. *TropicalMed*. 2023 Apr 18;8(4):227.
38. Pijls BG, Jolani S, Atherley A, Dijkstra JIR, Franssen GHL, Hendriks S, et al. Temporal trends of sex differences for COVID-19 infection, hospitalisation, severe disease, intensive care unit (ICU) admission and death: a meta-analysis of 229 studies covering over 10M patients. *F1000Res*. 2022 Jan 5;11:5.
39. Ho JQ, Sepand MR, Bigdelou B, Shekarian T, Esfandyarpour R, Chauhan P, et al. The immune response to COVID -19: Does sex matter? *Immunology*. 2022 Aug;166(4):429–43.
40. Mantovani A, Rinaldi E, Zusi C, Beatrice G, Saccomani MD, Dalbeni A. Coronavirus disease 2019 (COVID-19) in children and/or adolescents: a meta-analysis. *Pediatr Res*. 2021 Mar;89(4):733–7.
41. González Rodríguez P, Pérez-Moneo Agapito B, Albi Rodríguez MS, Aizpurua Galdeano P, Aparicio Rodrigo M, Fernández Rodríguez MM, et al. COVID-19 en pediatría: valoración crítica de la evidencia. *Anales de Pediatría*. 2021 Sep;95(3):207.e1-207.e13.
42. Sarker R, Roknuzzaman ASM, Nazmunnahar, Shahriar M, Hossain MdJ, Islam MdR. The WHO has declared the end of pandemic phase of COVID-19: Way to come back in the normal life. *Health Science Reports*. 2023 Sep;6(9):e1544.
43. Du J, Lang H mei, Ma Y, Chen A wen, Qin Y yi, Zhang X ping, et al. Global trends in COVID-19 incidence and case fatality rates (2019–2023): a retrospective analysis. *Front Public Health*. 2024 Jul 29;12:1355097.
44. Chan JFW, Kok KH, Zhu Z, Chu H, To KKW, Yuan S, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging Microbes & Infections*. 2020 Jan;9(1):221–36.
45. Naqvi AAT, Fatima K, Mohammad T, Fatima U, Singh IK, Singh A, et al. Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics

- approach. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 2020 Oct;1866(10):165878.
46. Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H. The Architecture of SARS-CoV-2 Transcriptome. *Cell*. 2020 May;181(4):914-921.e10.
47. Yao H, Song Y, Chen Y, Wu N, Xu J, Sun C, et al. Molecular Architecture of the SARS-CoV-2 Virus. *Cell*. 2020 Oct;183(3):730-738.e13.
48. Schmidt N, Lareau CA, Keshishian H, Ganskih S, Schneider C, Hennig T, et al. The SARS-CoV-2 RNA–protein interactome in infected human cells. *Nat Microbiol*. 2020 Dec 21;6(3):339–53.
49. Wang X, Wang W, Wang T, Wang J, Jiang Y, Wang X, et al. SARS-CoV-2 ORF8 Protein Induces Endoplasmic Reticulum Stress-like Responses and Facilitates Virus Replication by Triggering Calnexin: an Unbiased Study. Schultz-Cherry S, editor. *J Virol*. 2023 Mar 30;97(3):e00011-23.
50. Li Z, Zhang F, Jin X, Zhuang J. Elevated LIF and JAK-STAT activation drive severe COVID-19 in myeloma patients receiving the BCMA-CD3 bispecific antibody Elranatamab. *J Transl Med*. 2025 Jan 26;23(1):117.
51. Hawerkamp HC, Dyer AH, Patil ND, McElheron M, O’Dowd N, O’Doherty L, et al. Characterisation of the pro-inflammatory cytokine signature in severe COVID-19. *Front Immunol*. 2023 Mar 30;14:1170012.
52. Podzolkov VI, Pokrovskaya AE, Vanina DD, Shvedov II. sVCAM-1 as a Marker of Endothelial Dysfunction associated with Severe Course of a Novel Coronavirus Infection (COVID-19). *Racional’naâ farmakoterapiâ v kardiologii*. 2023 Apr 28;19(2):134–42.
53. Jordan SC. Innate and adaptive immune responses to SARS-CoV-2 in humans: relevance to acquired immunity and vaccine responses. *Clinical and Experimental Immunology*. 2021 May 13;204(3):310–20.
54. Cox RJ, Brokstad KA. Not just antibodies: B cells and T cells mediate immunity to COVID-19. *Nat Rev Immunol*. 2020 Oct;20(10):581–2.
55. Casanova JL, Su HC, Abel L, Aiuti A, Almuhsen S, Arias AA, et al. A Global Effort to Define the Human Genetics of Protective Immunity to SARS-CoV-2 Infection. *Cell*. 2020 Jun;181(6):1194–9.
56. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020 Mar 12;579(7798):265–9.
57. Official hCoV-19 Reference Sequence (WIV04) [Internet]. hCoV-19 Reference Sequences. 2020 [cited 2025 May 5]. Available from: <https://gisaid.org/wiv04/>
58. Okada P, Buathong R, Phuygun S, Thanadachakul T, Parnmen S, Wongboot W, et al. Early transmission patterns of coronavirus disease 2019 (COVID-19) in travellers from Wuhan to Thailand, January 2020. *Eurosurveillance* [Internet]. 2020 Feb 27 [cited 2025 May 5];25(8). Available from: <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.8.2000097>
59. WHO announces simple, easy-to-say labels for SARS-CoV-2 Variants of Interest and Concern [Internet]. World Health Organization (WHO); 2021 Dec [cited 2025 May 5]. (WHO announces simple, easy-to-say labels for SARS-CoV-2 Variants of Interest and Concern).

- Available from: <https://www.who.int/news/item/31-05-2021-who-announces-simple-easy-to-say-labels-for-sars-cov-2-variants-of-interest-and-concern>
60. Manirambona E, Okesanya OJ, Olaleke NO, Oso TA, Lucero-Prisno DE. Evolution and implications of SARS-CoV-2 variants in the post-pandemic era. *Discov Public Health*. 2024 Jun 28;21(1):16.
 61. SARS-CoV-2 variants of concern as of 25 April 2025 [Internet]. European Centre for Disease Prevention and Control; 2025 Jan [cited 2025 May 5]. Available from: <https://www.ecdc.europa.eu/en/covid-19/variants-concern>
 62. Callaway E. Heavily mutated Omicron variant puts scientists on alert. *Nature*. 2021 Dec 2;600(7887):21–21.
 63. Giurgea LT, Morens DM. Great Expectations of COVID-19 Herd Immunity. Yount J, editor. *mBio*. 2022 Feb 22;13(1):e03495-21.
 64. Genomic epidemiology of SARS-CoV-2 with subsampling focused globally since pandemic start [Internet]. 2025 [cited 2025 May 5]. Available from: <https://nextstrain.org/ncov/gisaid/global/all-time?d=tree&l=radial&m=div&p=full>
 65. Zhu J, Ji P, Pang J, Zhong Z, Li H, He C, et al. Clinical characteristics of 3062 COVID-19 patients: A meta-analysis. *Journal of Medical Virology*. 2020 Oct;92(10):1902–14.
 66. Page EM, Ariëns RAS. Mechanisms of thrombosis and cardiovascular complications in COVID-19. *Thrombosis Research*. 2021 Apr;200:1–8.
 67. Gavriatopoulou M, Korompoki E, Fotiou D, Ntanasis-Stathopoulos I, Psaltopoulou T, Kastritis E, et al. Organ-specific manifestations of COVID-19 infection. *Clin Exp Med*. 2020 Nov;20(4):493–506.
 68. Kunutsor SK, Laukkanen JA. Renal complications in COVID-19: a systematic review and meta-analysis. *Annals of Medicine*. 2020 Oct 2;52(7):345–53.
 69. Ribeiro-Junior MAF, Augusto SDS, Elias YGB, Costa CTK, Néder PR. GASTROINTESTINAL COMPLICATIONS OF CORONAVIRUS DISEASE (COVID-19). *ABCD, arq bras cir dig*. 2021;34(3):e1620.
 70. Erdinc B, Sahni S, Gotlieb V. Hematological manifestations and complications of COVID-19. *Adv Clin Exp Med*. 2021 Feb 2;30(1):101–7.
 71. Daneshgaran G, Dubin DP, Gould DJ. Cutaneous Manifestations of COVID-19: An Evidence-Based Review. *Am J Clin Dermatol*. 2020 Oct;21(5):627–39.
 72. Fred F. F. Ferri's Clinical Advisor. Elsevier; 2025. 327.e10-327.e15.
 73. Chadaga K, Prabhu S, Bhat V, Sampathila N, Umakanth S, Chadaga R. Artificial intelligence for diagnosis of mild–moderate COVID-19 using haematological markers. *Annals of Medicine*. 2023 Dec 12;55(1):2233541.
 74. Fan Y, Liu M, Sun G. An interpretable machine learning framework for diagnosis and prognosis of COVID-19. Akbar S, editor. *PLoS ONE*. 2023 Sep 21;18(9):e0291961.
 75. Lai YJ, Liu SH, Manachevakul S, Lee TA, Kuo CT, Bello D. Biomarkers in long COVID-19: A systematic review. *Front Med*. 2023 Jan 20;10:1085988.

76. Eastman A, Vervaart P, Rahman KM, Aggar C, Rosic N. Evaluating Biomarkers for Predicting Clinical Outcomes in COVID-19 Patients: A Systematic Review. *CBR*. 2024;44(2):47–90.
77. Liu A, Hammond R, Donnelly PD, Kaski JC, Coates ARM. Effective prognostic and clinical risk stratification in COVID-19 using multimodality biomarkers. *J Intern Med*. 2023 Jul;294(1):21–46.
78. O’Connell TF, Bradley CJ, Abbas AE, Williamson BD, Rusia A, Tawney AM, et al. Hydroxychloroquine/Azithromycin Therapy and QT Prolongation in Hospitalized Patients With COVID-19. *JACC: Clinical Electrophysiology*. 2021 Jan;7(1):16–25.
79. Bloch EM, Focosi D, Shoham S, Senefeld J, Tobian AAR, Baden LR, et al. Guidance on the Use of Convalescent Plasma to Treat Immunocompromised Patients With Coronavirus Disease 2019. *Clinical Infectious Diseases*. 2023 Jun 8;76(11):2018–24.
80. Pérez-Belmonte LM, López-Carmona MD, Quevedo-Marín JL, Ricci M, Martín-Carmona J, Sanz-Cánovas J, et al. Differences between Clinical Protocols for the Treatment of Coronavirus Disease 2019 (COVID-19) in Andalusia, Spain. *IJERPH*. 2020 Sep 19;17(18):6845.
81. Paxlovid [Internet]. European Medicines Agency (EMA); 2025 Feb [cited 2025 Feb 18]. Available from: <https://www.ema.europa.eu/en/medicines/human/EPAR/paxlovid#authorisation-details>
82. Tregoning JS, Brown ES, Cheeseman HM, Flight KE, Higham SL, Lemm NM, et al. Vaccines for COVID-19. *Clinical and Experimental Immunology*. 2020 Oct 30;202(2):162–92.
83. Jain S, Venkataraman A, Wechsler ME, Peppas NA. Messenger RNA-based vaccines: Past, present, and future directions in the context of the COVID-19 pandemic. *Advanced Drug Delivery Reviews*. 2021 Dec;179:114000.
84. Yadav T, Kumar S, Mishra G, Saxena SK. Tracking the COVID-19 vaccines: The global landscape. *Human Vaccines & Immunotherapeutics*. 2023 Jan 2;19(1):2191577.
85. Beladiya J, Kumar A, Vasava Y, Parmar K, Patel D, Patel S, et al. Safety and efficacy of COVID-19 vaccines: A systematic review and meta-analysis of controlled and randomized clinical trials. *Reviews in Medical Virology*. 2024 Jan;34(1):e2507.
86. Goławski M, Lewandowski P, Jabłońska I, Delijewski M. The Reassessed Potential of SARS-CoV-2 Attenuation for COVID-19 Vaccine Development—A Systematic Review. *Viruses*. 2022 May 7;14(5):991.
87. Okamura S, Ebina H. Could live attenuated vaccines better control COVID-19? *Vaccine*. 2021 Sep;39(39):5719–26.
88. Lundstrom K. COVID-19 Vaccines: Where Did We Stand at the End of 2023? *Viruses*. 2024 Jan 29;16(2):203.
89. Hofmeyer KA, Ventura CL, Armstrong KL, Houchens CR, Patel S, Disbrow GL, et al. Project NextGen: Developing the Next Generation of COVID-19 Vaccines and Therapeutics to Respond to the Present and Prepare for the Future. *Clinical Infectious Diseases*. 2024 Jul 19;79(1):115–21.

90. SeyedAlinaghi S, Pashapouryeganeh A, Dehghani S, Mirzapour P, Abbaspour F, Afroughi F, et al. Feasibility and Effectiveness of Vaccines for COVID-19: An Umbrella Review. *Arch Acad Emerg Med*. 2025;13(1):e6.
91. Soares P, Rocha JV, Moniz M, Gama A, Laires PA, Pedro AR, et al. Factors Associated with COVID-19 Vaccine Hesitancy. *Vaccines*. 2021 Mar 22;9(3):300.
92. WHO COVID-19 Dashboard: COVID-19 Vaccination, World Data [Internet]. World Health Organization (WHO); 2023 Dec [cited 2025 Apr 2]. Available from: <https://data.who.int/dashboards/covid19/vaccines?n=c>
93. Markhof Y, Wollburg P, Zezza A. Beyond the records: Data quality and COVID-19 vaccination progress in low- and middle-income countries. *Journal of Development Economics*. 2025 May;174:103449.
94. Terapias Avanzadas [Internet]. Agencia Española de Medicamentos y Productos Sanitarios (AEMPS); 2023 Nov [cited 2025 Feb 13]. Available from: <https://www.aemps.gob.es/medicamentos-de-uso-humano/medicamentos-de-fabricacion-no-industrial/terapias-avanzadas/>
95. Tang C, Dziedzic A, Khatib MN, Alhumaid S, Thangavelu L, Parameswari R, et al. Stem cell therapy for COVID-19 treatment: an umbrella review. *International Journal of Surgery*. 2024 Oct;110(10):6402–17.
96. Deng X, Terunuma H, Nieda M. Exploring the Utility of NK Cells in COVID-19. *Biomedicines*. 2022 Apr 26;10(5):1002.
97. Post COVID-19 condition (Long COVID) [Internet]. World Health Organization (WHO); 2022 Dec [cited 2025 Feb 14]. Available from: <https://www.who.int/europe/news-room/fact-sheets/item/post-COVID-19-condition>
98. Silva MMD, Benites MN, Castro YM, Moura PV, Zhang L. Prevalence of symptoms of post-COVID-19 condition (long COVID) in children hospitalized with COVID-19: A systematic review of observational studies. *Pediatric Pulmonology*. 2024 Dec;59(12):3159–69.
99. Jangnin R, Ritruangroj W, Kittisupkajorn S, Sukeiam P, Inchai J, Maneeton B, et al. Long-COVID Prevalence and Its Association with Health Outcomes in the Post-Vaccine and Antiviral-Availability Era. *JCM*. 2024 Feb 21;13(5):1208.
100. Donald J, Bilasy SE, Yang C, El-Shamy A. Exploring the Complexities of Long COVID. *Viruses*. 2024 Jun 30;16(7):1060.
101. Asadi-Pooya AA, Nemati M, Shahisavandi M, Nemati H, Karimi A, Jafari A, et al. How does COVID-19 vaccination affect long-COVID symptoms? Bastani P, editor. *PLoS ONE*. 2024 Feb 7;19(2):e0296680.
102. Yin K, Peluso MJ, Luo X, Thomas R, Shin MG, Neidleman J, et al. Long COVID manifests with T cell dysregulation, inflammation and an uncoordinated adaptive immune response to SARS-CoV-2. *Nat Immunol*. 2024 Feb;25(2):218–25.
103. Ruf W. Immune damage in Long Covid. *Science*. 2024 Jan 19;383(6680):262–3.
104. McMillan P, Turner AJ, Uhal BD. Mechanisms of Gut-Related Viral Persistence in Long COVID. *Viruses*. 2024 Aug 7;16(8):1266.

105. Kanwal A, Zhang Z. Exploring common pathogenic association between Epstein Barr virus infection and long-COVID by integrating RNA-Seq and molecular dynamics simulations. *Front Immunol.* 2024 Sep 26;15:1435170.
106. Hse University YA, Pekarski SE. Soaring public debt: Return of financial repression and high inflation? *Vopr èkon.* 2024 Jan 7;(1):33–51.
107. Chen N. Income insecurity and social protection: Examining the impact of the COVID-19 pandemic across income groups. Kálmán BG, editor. *PLoS ONE.* 2024 Sep 20;19(9):e0310680.
108. Kisa S, Kisa A. A Comprehensive Analysis of COVID-19 Misinformation, Public Health Impacts, and Communication Strategies: Scoping Review. *J Med Internet Res.* 2024 Aug 21;26:e56931.
109. Gilbert L, Parker S, Schechter L. The impact of the COVID-19 pandemic on treatment for domestic violence injuries: evidence from medical claims [Internet]. 2023 [cited 2025 Feb 14]. Available from: <https://www.researchsquare.com/article/rs-3089683/v1>
110. Dewa LH, Roberts L, Choong E, Crandell C, Demkowicz O, Ashworth E, et al. The impact of COVID-19 on young people’s mental health, wellbeing and routine from a European perspective: A co-produced qualitative systematic review. Zimmerman S, editor. *PLoS ONE.* 2024 Mar 20;19(3):e0299547.
111. Kähkönen T. Remote work during the COVID-19 pandemic: identification of working life impacts, employees’ data protection abilities and trust outcomes. *JOCM.* 2023 Jun 20;36(3):472–92.
112. Ofori EK, Zhang J, Nyantakyi G, Hayford IS, Tergu CT. Impact of Covid-19 on environmental sustainability: A bibliometric analysis. *Sustainable Development.* 2023 Aug;31(4):2176–95.
113. Vadiati M, Ghasemi L, Samani S, Islam MdA, Ahmadi A, Khaleghi S, et al. A sustainable trend in COVID-19 research: An environmental perspective. *Front Environ Sci.* 2023 Mar 29;11:1104679.
114. San Martín-López JV, Mesa N, Bernal-Bello D, Morales-Ortega A, Rivilla M, Guerrero M, et al. Seven Epidemic Waves of COVID-19 in a Hospital in Madrid: Analysis of Severity and Associated Factors. *Viruses.* 2023 Aug 30;15(9):1839.
115. Spielmanns M, Schaer CE, Pekacka-Egli AM, Spielmanns S, Ibish O, Gafina G, et al. Pulmonary Rehabilitation Outcomes of Post-Acute COVID-19 Patients during Different Waves of the Pandemic. *IJERPH.* 2023 May 22;20(10):5907.
116. Otero J. Informe científico-divulgativo sobre la labor del Instituto de Salud Carlos III (ISCIII) tras dos años de pandemia de COVID-19 [Internet]. Institutio de Salud Carlos III. Ministerio de Ciencia e Innovación; 2022 Mar [cited 2025 Feb 25]. Available from: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.isciii.es/documents/20119/537286/Informe+dos+a%C3%B1os+de+pandemia+ISCIII_final.pdf/33eef9e6-98ed-fd3c-ab04-e1cd88892147
117. Ashby B, Best A. Herd immunity. *Current Biology.* 2021 Feb;31(4):R174–7.
118. Yang J, Lo NC, Dankwa EA, Donnelly CA, Gupta R, Montgomery MP, et al. Determining Herd Immunity Thresholds for Hepatitis A Virus Transmission to Inform

- Vaccination Strategies Among People Who Inject Drugs in 16 US States. *Clinical Infectious Diseases*. 2024 Apr 10;78(4):976–82.
119. García-García D, Morales E, Fonfría ES, Vigo I, Bordehore C. Caveats on COVID-19 herd immunity threshold: the Spain case. *Sci Rep*. 2022 Jan 12;12(1):598.
120. Elsaid M, Nasef MA, Huy NT. R_0 of COVID-19 and its impact on vaccination coverage: compared with previous outbreaks. *Human Vaccines & Immunotherapeutics*. 2021 Nov 2;17(11):3850–4.
121. Kavaliunas A, Ocaya P, Mumper J, Lindfeldt I, Kyhlstedt M. Swedish policy analysis for Covid-19. *Health Policy and Technology*. 2020 Dec;9(4):598–612.
122. Silva RBTRD, Garcia S, Nääs IDA. Brazil, USA, and Germany: A comparison of legislation in the fight against COVID-19. *RSD*. 2021 Jan 24;10(1):e44810111942.
123. Wardak MF, Rahimi A, Lucero-Prisno DE, Miranda AV. Overcoming barriers to COVID-19 herd immunity in Afghanistan. *Public Health Challenges*. 2024 Mar;3(1):e161.
124. Han E, Tan MMJ, Turk E, Sridhar D, Leung GM, Shibuya K, et al. Lessons learnt from easing COVID-19 restrictions: an analysis of countries and regions in Asia Pacific and Europe. *The Lancet*. 2020 Nov;396(10261):1525–34.
125. Liu M, Shi L, Yang M, Jiao J, Yang J, Ma M, et al. Ecological comparison of six countries in two waves of COVID-19. *Front Public Health*. 2024 Feb 28;12:1277457.
126. Hou T, Zeng W, Yang M, Chen W, Ren L, Ai J, et al. Development and evaluation of a rapid CRISPR-based diagnostic for COVID-19. Krammer F, editor. *PLoS Pathog*. 2020 Aug 27;16(8):e1008705.
127. Liu TY, Knott GJ, Smock DCJ, Desmarais JJ, Son S, Bhuiya A, et al. Accelerated RNA detection using tandem CRISPR nucleases. *Nat Chem Biol*. 2021 Sep;17(9):982–8.
128. Novel SARS-CoV-2 Point-of-care Testing [Internet]. National Library of Medicine; 2023 Feb. Report No.: NCT05034978. Available from: <https://clinicaltrials.gov/study/NCT05034978?cond=covid&term=crispr&rank=1>
129. Palmgrem G. Clinical Trial Update – CRISPR-Based Diagnostics for Sars-CoV-2 and Mycobacterium. 2020 Oct 28 [cited 2025 Feb 11]; Available from: <https://crisprmedicineneeds.com/news/clinical-trial-update-crispr-based-diagnostics-for-sars-cov-2-and-mycobacterium/>
130. Sadique MohdA, Yadav S, Khan R, Srivastava AK. Engineered two-dimensional nanomaterials based diagnostics integrated with internet of medical things (IoMT) for COVID-19. *Chem Soc Rev*. 2024;53(8):3774–828.
131. Wolfe DN, Arangies E, David GL, Armstrong B, Scocca TZ, Fedler J, et al. Development of Next-Generation COVID-19 Vaccines: Biomedical Advanced Research and Development Authority (BARDA)–Supported Phase 2b Study Designs. *Clinical Infectious Diseases*. 2024 May 28;ciae286.
132. Mendoza-Ramírez NJ, García-Cordero J, Shrivastava G, Cedillo-Barrón L. The Key to Increase Immunogenicity of Next-Generation COVID-19 Vaccines Lies in the Inclusion of the SARS-CoV-2 Nucleocapsid Protein. Amdare N, editor. *Journal of Immunology Research*. 2024 Jan;2024(1):9313267.

133. Klinkhardt U, Schunk M, Ervin J, Schindler C, Sugimoto D, Rankin B, et al. A novel orf virus vector-based COVID-19 booster vaccine shows cross-neutralizing activity in the absence of anti-vector neutralizing immunity. *Human Vaccines & Immunotherapeutics*. 2024 Dec 31;20(1):2410574.
134. Hao X, Yuan F, Yao X. Advances in virus-like particle-based SARS-CoV-2 vaccines. *Front Cell Infect Microbiol*. 2024 Jun 26;14:1406091.
135. Xu Y, Masuda K, Groso C, Hassan R, Zhou Z, Broderick K, et al. Microfluidic Synthesis of Scalable Layer-by-Layer Multiple Antigen Nano-Delivery Platform for SARS-CoV-2 Vaccines. *Vaccines*. 2024 Mar 21;12(3):339.
136. Du P, Huang L, Fang Y, Zhao F, Li Q, Ma X, et al. Broad-spectrum Delta-BA.2 tandem-fused heterodimer mRNA vaccine delivered by lipopolyplex. *Best SM, editor. PLoS Pathog*. 2024 Apr 1;20(4):e1012116.
137. Zhang P, Luo S, Zou P, Deng Q, Wang C, Li J, et al. A novel simian adenovirus-vectored COVID-19 vaccine elicits effective mucosal and systemic immunity in mice by intranasal and intramuscular vaccination regimens. *Zhang B, editor. Microbiol Spectr*. 2023 Dec 12;11(6):e01794-23.
138. Wong LYR, Zheng J, Wilhelmsen K, Li K, Ortiz ME, Schnicker NJ, et al. Eicosanoid signalling blockade protects middle-aged mice from severe COVID-19. *Nature*. 2022 May 5;605(7908):146–51.
139. Enjuanes L. Luis Enjuanes, virólogo experto en covid: «Viviremos una situación parecida en un futuro» [Internet]. 2024 [cited 2025 Feb 18]. Available from: <https://theobjective.com/sanidad/2024-02-05/luis-enjuanes-virologo-covid-entrevista/#>
140. Eltayb WA, Abdalla M, Rabie AM. Novel Investigational Anti-SARS-CoV-2 Agent Ensirelvir “S-217622”: A Very Promising Potential Universal Broad-Spectrum Antiviral at the Therapeutic Frontline of Coronavirus Species. *ACS Omega*. 2023 Feb 14;8(6):5234–46.
141. Huang PY, Liu TH, Wu JY, Tsai YW, Lai CC. Clinical efficacy and safety of molnupiravir for nonhospitalized and hospitalized patients with COVID-19: A systematic review and meta-analysis of randomized control trials. *J Med Virol*. 2023 Mar;95(3):e28621.
142. Guan Y, Puenpatom A, Johnson MG, Zhang Y, Zhao Y, Surber J, et al. Impact of Molnupiravir Treatment on Patient-Reported COVID-19 Symptoms in the Phase 3 MOVE-OUT Trial: A Randomized, Placebo-Controlled Trial. *Clinical Infectious Diseases*. 2023 Nov 30;77(11):1521–30.
143. Papini C, Ullah I, Ranjan AP, Zhang S, Wu Q, Spasov KA, et al. Proof-of-concept studies with a computationally designed M^{pro} inhibitor as a synergistic combination regimen alternative to Paxlovid. *Proc Natl Acad Sci USA*. 2024 Apr 23;121(17):e2320713121.
144. Ruiz Seco MP, Paño Pardo JR, Schoergenhofer C, Dings C, Lehr T, Herth F, et al. Efficacy and safety of asunercept, a CD95L-selective inhibitor, in hospitalised patients with moderate-to-severe COVID-19: ASUNCTIS, a multicentre, randomised, open-label, controlled, phase 2 trial. *eClinicalMedicine*. 2024 Nov;77:102879.
145. UQ Centre for Clinical Research, Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia, Hernández-Mitre MP, Morpeth SC, Venkatesh B, Hills TE, Davis J, et al. TMPRSS2 inhibitors for the treatment of COVID-19 in adults: a systematic review and meta-analysis of nafamostat and camostat mesylate randomised clinical trials [Internet].

- INPLASY - International Platform of Registered Systematic Review and Meta-analysis Protocols; 2023 Feb [cited 2025 Feb 21]. Available from: <https://inplasy.com/inplasy-2023-2-0120/>
146. Thomas G, Couture F, Kwiatkowska A. The Path to Therapeutic Furin Inhibitors: From Yeast Pheromones to SARS-CoV-2. *IJMS*. 2022 Mar 22;23(7):3435.
147. Boon ACM, Bricker TL, Fritch EJ, Leist SR, Gully K, Baric RS, et al. Efficacy of host cell serine protease inhibitor MM3122 against SARS-CoV-2 for treatment and prevention of COVID-19. Liu SL, editor. *J Virol*. 2024 May 14;98(5):e01903-23.
148. Roman YM, Burela PA, Pasupuleti V, Piscocoya A, Vidal JE, Hernandez AV. Ivermectin for the Treatment of Coronavirus Disease 2019: A Systematic Review and Meta-analysis of Randomized Controlled Trials. *Clinical Infectious Diseases*. 2022 Mar 23;74(6):1022–9.
149. Manabe T, Kambayashi D, Akatsu H, Kudo K. Favipiravir for the treatment of patients with COVID-19: a systematic review and meta-analysis. *BMC Infect Dis*. 2021 Dec;21(1):489.
150. Lee G, Schauner R, Burke J, Borocz J, Vasana S, Sobieraj L, et al. NK cells from COVID-19 positive patients exhibit enhanced cytotoxic activity upon NKG2A and KIR2DL1 blockade. *Front Immunol*. 2023 Jul 7;14:1022890.
151. Pal N, Mavi AK, Kumar S, Kumar U, Joshi MD, Saluja R. Current updates on adaptive immune response by B cell and T cell stimulation and therapeutic strategies for novel coronavirus disease 2019 (COVID-19) treatment. *Heliyon*. 2021 Apr;7(4):e06894.
152. Wang Y, Liang Q, Chen F, Zheng J, Chen Y, Chen Z, et al. Immune-Cell-Based Therapy for COVID-19: Current Status. *Viruses*. 2023 Oct 25;15(11):2148.
153. Rezaei FS, Khorshidian A, Beram FM, Derakhshani A, Esmaeili J, Barati A. 3D printed chitosan/polycaprolactone scaffold for lung tissue engineering: hope to be useful for COVID-19 studies. *RSC Adv*. 2021;11(32):19508–20.
154. Ma Y, Zhou Y, Jiang D, Dai W, Li J, Deng C, et al. Integration of human organoids single-cell transcriptomic profiles and human genetics repurposes critical cell type-specific drug targets for severe COVID -19. *Cell Proliferation*. 2024 Mar;57(3):e13558.
155. García-González L, Martí-Sarrias A, Puertas MC, Bayón-Gil Á, Resa-Infante P, Martínez-Picado J, et al. Understanding the neurological implications of acute and long COVID using brain organoids. *Disease Models & Mechanisms*. 2023 Jul 1;16(7):dmm050049.
156. Vanslambrouck JM, Neil JA, Rudraraju R, Mah S, Tan KS, Groenewegen E, et al. Kidney organoids reveal redundancy in viral entry pathways during ACE2-dependent SARS-CoV-2 infection. Liu SL, editor. *J Virol*. 2024 Mar 19;98(3):e01802-23.
157. Fisher CR, Mba Medie F, Luu RJ, Gaibler RB, Mulhern TJ, Miller CR, et al. A High-Throughput, High-Containment Human Primary Epithelial Airway Organ-on-Chip Platform for SARS-CoV-2 Therapeutic Screening. *Cells*. 2023 Nov 16;12(22):2639.
158. Negi V, Gavlock D, Miedel MT, Lee JK, Shun T, Gough A, et al. Modeling mechanisms underlying differential inflammatory responses to COVID-19 in type 2 diabetes using a patient-derived microphysiological organ-on-a-chip system. *Lab Chip*. 2023;23(20):4514–27.
159. Amisha, Malik P, Pathania M, Rathaur V. Overview of artificial intelligence in medicine. *J Family Med Prim Care*. 2019;8(7):2328.

160. Haenlein M, Kaplan A. A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*. 2019 Aug;61(4):5–14.
161. Rojas A, Estid M. Una mirada a la inteligencia artificial. *Revista Ingeniería, Matemáticas y Ciencias de la Información*. 2015;2(3):27–31.
162. Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. *Electron Markets*. 2021 Sep;31(3):685–95.
163. Richens JG, Lee CM, Johri S. Improving the accuracy of medical diagnosis with causal machine learning. *Nat Commun*. 2020 Aug 11;11(1):3923.
164. Scott IA. Machine Learning and Evidence-Based Medicine. *Ann Intern Med*. 2018 Jul 3;169(1):44–6.
165. Yuan L, Yang L, Zhang S, Xu Z, Qin J, Shi Y, et al. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective multicentre clinical cohort study. *EClinicalMedicine*. 2023 Mar;57:101834.
166. Hsieh C, Laguna A, Ikeda I, Maxwell AWP, Chapiro J, Nadolski G, et al. Using Machine Learning to Predict Response to Image-guided Therapies for Hepatocellular Carcinoma. *Radiology*. 2023 Nov 1;309(2):e222891.
167. Martins TGDS, Schor P. Machine learning in image analysis in ophthalmology. *Einstein (São Paulo)*. 2021 Dec 15;19:eED6860.
168. Zhu S, Gilbert M, Chetty I, Siddiqui F. The 2021 landscape of FDA-approved artificial intelligence/machine learning-enabled medical devices: An analysis of the characteristics and intended use. *International Journal of Medical Informatics*. 2022 Sep;165:104828.
169. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol*. 2019 Dec;19(1):64.
170. Myszczyńska MA, Ojamies PN, Lacoste AMB, Neil D, Saffari A, Mead R, et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nat Rev Neurol*. 2020 Aug;16(8):440–56.
171. Shameer K, Johnson KW, Glicksberg BS, Dudley JT, Sengupta PP. Machine learning in cardiovascular medicine: are we there yet? *Heart*. 2018 Jul;104(14):1156–64.
172. Cho G, Park JR, Choi Y, Ahn H, Lee H. Detection of COVID-19 epidemic outbreak using machine learning. *Front Public Health*. 2023 Dec 18;11:1252357.
173. Lăzăroiu G, Gedeon T, Rogalska E, Andronic M, Frajtova Michalikova K, Musova Z, et al. The economics of deep and machine learning-based algorithms for COVID-19 prediction, detection, and diagnosis shaping the organizational management of hospitals. *oc*. 2024 Mar 30;15(1):27–58.
174. Dobrijević D, Vilotijević-Dautović G, Katanić J, Horvat M, Horvat Z, Pastor K. Rapid Triage of Children with Suspected COVID-19 Using Laboratory-Based Machine-Learning Algorithms. *Viruses*. 2023 Jul 8;15(7):1522.
175. Caliendo ÁM, Hanson KE. COVID-19: Diagnosis [Internet]. UpToDate. Ministerio de Sanidad; 2024 Dec [cited 2025 Feb 26]. Available from: <https://www.uptodate.com/contents/covid-19-diagnosis?search=COVID->

- 19%20reinfection%20diagnosis&source=search_result&selectedTitle=5%7E150&usage_type=default&display_rank=4#H3118018327
176. Slezak J, Bruxvoort K, Fischer H, Broder B, Ackerson B, Tartof S. Rate and severity of suspected SARS-Cov-2 reinfection in a cohort of PCR-positive COVID-19 patients. *Clinical Microbiology and Infection*. 2021 Dec;27(12):1860.e7-1860.e10.
177. Han J, Kamber M, Pei J. *Data Mining: Concepts and Techniques*. USA: Morgan Kaufmann Publishers: Burlington; 2022.
178. Charbuty B, Abdulazeez A. Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*. 2021;2(01):20–8.
179. Ontivero-Ortega M, Lage-Castellanos A, Valente G, Goebel R, Valdes-Sosa M. Fast Gaussian Naïve Bayes for searchlight classification analysis. *NeuroImage*. 2017 Dec;163:471–9.
180. As’ad I. Advancing Healthcare Diagnostics: A Study on Gaussian Naive Bayes Classification of Blood Samples. *ijaimi*. 2023 Nov 30;1(2):115–23.
181. Elashmawi WH, Djellal A, Sheta A, Surani S, Aljahdali S. Machine Learning for Enhanced COPD Diagnosis: A Comparative Analysis of Classification Algorithms. *Diagnostics*. 2024 Dec 14;14(24):2822.
182. Mu J, Dai L, Liu JX, Shang J, Xu F, Liu X, et al. Automatic detection for epileptic seizure using graph-regularized nonnegative matrix factorization and Bayesian linear discriminate analysis. *Biocybernetics and Biomedical Engineering*. 2021 Oct;41(4):1258–71.
183. Uddin S, Haque I, Lu H, Moni MA, Gide E. Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Sci Rep*. 2022 Apr 15;12(1):6256.
184. Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez A. A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*. 2020 Sep;408:189–215.
185. Guido R, Ferrisi S, Lofaro D, Conforti D. An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review. *Information*. 2024 Apr 19;15(4):235.
186. Shi G, Chen Z, Zhang R. Efficient Support Vector Machine Toward Medical Data Processing. In: Yang XS, Sherratt S, Dey N, Joshi A, editors. *Proceedings of Seventh International Congress on Information and Communication Technology* [Internet]. Singapore: Springer Nature Singapore; 2023 [cited 2025 Mar 20]. p. 751–60. (Lecture Notes in Networks and Systems; vol. 448). Available from: https://link.springer.com/10.1007/978-981-19-1610-6_66
187. Qiu B, Su XH, Qin X, Wang Q. Application of machine learning techniques in real-world research to predict the risk of liver metastasis in rectal cancer. *Front Oncol*. 2022 Dec 20;12:1065468.
188. Montomoli J, Romeo L, Moccia S, Bernardini M, Migliorelli L, Berardini D, et al. Machine learning using the extreme gradient boosting (XGBoost) algorithm predicts 5-day delta of SOFA score at ICU admission in COVID-19 patients. *Journal of Intensive Medicine*. 2021 Oct;1(2):110–6.

189. Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms. *Artif Intell Rev.* 2021 Mar;54(3):1937–67.
190. Schonlau M, Zou RY. The random forest algorithm for statistical learning. *The Stata Journal: Promoting communications on statistics and Stata.* 2020 Mar;20(1):3–29.
191. Rhodes JS, Cutler A, Moon KR. Geometry- and Accuracy-Preserving Random Forest Proximities. *IEEE Trans Pattern Anal Mach Intell.* 2023 Sep 1;45(9):10947–59.
192. Hall TA, Caromi R, Souryal M, Wunderlich A. Reference Datasets for Training and Evaluating RF Signal Detection and Classification Models. In: 2019 IEEE Globecom Workshops (GC Wkshps) [Internet]. Waikoloa, HI, USA: IEEE; 2019 [cited 2025 Mar 13]. p. 1–5. Available from: <https://ieeexplore.ieee.org/document/9024532/>
193. Alawneh H, Hasasneh A. Survival Prediction of Children after Bone Marrow Transplant Using Machine Learning Algorithms. *IAJIT* [Internet]. 2024 [cited 2025 Apr 1];21(3). Available from: <https://iajit.org/upload/files/Survival-Prediction-of-Children-after-Bone-Marrow-Transplant-Using-Machine-Learning-Algorithms.pdf>
194. Kosiborod MN, Petrie MC, Borlaug BA, Butler J, Davies MJ, Hovingh GK, et al. Semaglutide in Patients with Obesity-Related Heart Failure and Type 2 Diabetes. *N Engl J Med.* 2024 Apr 18;390(15):1394–407.
195. Lucas MM, Schootman M, Laryea JA, Orcutt ST, Li C, Ying J, et al. Bias in Prediction Models to Identify Patients With Colorectal Cancer at High Risk for Readmission After Resection. *JCO Clin Cancer Inform.* 2024 Oct;(8):e2300194.
196. Yu Z, Li X, Zhao J, Sun S. Identification of hospitalized mortality of patients with COVID-19 by machine learning models based on blood inflammatory cytokines. *Front Public Health.* 2022 Nov 17;10:1001340.
197. Wan TK, Huang RX, Tulu TW, Liu JD, Vodencarevic A, Wong CW, et al. Identifying Predictors of COVID-19 Mortality Using Machine Learning. *Life.* 2022 Apr 6;12(4):547.
198. Valero-Bover D, Monterde D, Carot-Sans G, Cainzos-Achirica M, Comin-Colet J, Vela E, et al. Is Age the Most Important Risk Factor in COVID-19 Patients? The Relevance of Comorbidity Burden: A Retrospective Analysis of 10,551 Hospitalizations. *CLEP.* 2023 Jun;Volume 15:811–25.
199. 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 Dec;25(1):63.
200. Garrafa E, Vezzoli M, Ravanelli M, Farina D, Borghesi A, Calza S, et al. Early Prediction of In-Hospital Death of COVID-19 Patients: A Machine-Learning Model Based on Age, Blood Analyses, and Chest X-Ray Score [Internet]. 2021 [cited 2025 Apr 29]. Available from: <http://medrxiv.org/lookup/doi/10.1101/2021.06.10.21258721>
201. Tezza F, Lorenzoni G, Azzolina D, Barbar S, Leone LAC, Gregori D. Predicting in-Hospital Mortality of Patients with COVID-19 Using Machine Learning Techniques. *JPM.* 2021 Apr 24;11(5):343.
202. Romero Starke K, Reissig D, Petereit-Haack G, Schmauder S, Nienhaus A, Seidler A. The isolated effect of age on the risk of COVID-19 severe outcomes: a systematic review with meta-analysis. *BMJ Glob Health.* 2021 Dec;6(12):e006434.

203. Murri R, Lenkowicz J, Masciocchi C, Iacomini C, Fantoni M, Damiani A, et al. A machine-learning parsimonious multivariable predictive model of mortality risk in patients with Covid-19. *Sci Rep*. 2021 Oct 27;11(1):21136.
204. Subudhi S, Verma A, Patel AB, Hardin CC, Khandekar MJ, Lee H, et al. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. *npj Digit Med*. 2021 May 21;4(1):87.
205. Sánchez-Montalvá A, Álvarez-Sierra D, Martínez-Gallo M, Perurena-Prieto J, Arrese-Muñoz I, Ruiz-Rodríguez JC, et al. Exposing and Overcoming Limitations of Clinical Laboratory Tests in COVID-19 by Adding Immunological Parameters; A Retrospective Cohort Analysis and Pilot Study. *Front Immunol*. 2022 Jun 29;13:902837.
206. Chen H, Yang G, Long Y, Li C. Lymphocyte and Platelet Counts, as well as Interleukin-6 Levels, Predict Mortality of Coronavirus Disease 2019 (COVID-19): A Systematic Review and Meta-Analysis. Jamous R, editor. *Evidence-Based Complementary and Alternative Medicine*. 2021 Dec 20;2021:1–9.
207. Lai KL, Hu FC, Wen FY, Chen JJ. Lymphocyte count is a universal predictor of health outcomes in COVID-19 patients before mass vaccination: A meta-analytical study. *J Glob Health*. 2022 Sep 17;12:05041.
208. Ugalde A, Hellmann F, Homedes N. Inequity in access to vaccines: the failure of the global response to the COVID-19 pandemic. *SCOL*. 2022 Oct 1;18:e4190.
209. Rando HM, Lordan R, Lee AJ, Naik A, Wellhausen N, Sell E, et al. Application of Traditional Vaccine Development Strategies to SARS-CoV-2. Gilbert JA, editor. *mSystems*. 2023 Apr 27;8(2):e00927-22.
210. Silva-Pilipich N, Beloki U, Salaberry L, Smerdou C. Self-Amplifying RNA: A Second Revolution of mRNA Vaccines against COVID-19. *Vaccines*. 2024 Mar 17;12(3):318.
211. Chen P, Shi X, He W, Zhong G, Tang Y, Wang H, et al. mRNA vaccine-a desirable therapeutic strategy for surmounting COVID-19 pandemic. *Human Vaccines & Immunotherapeutics*. 2022 Jan 31;18(1):2040330.
212. Hussain A, Yang H, Zhang M, Liu Q, Alotaibi G, Irfan M, et al. mRNA vaccines for COVID-19 and diverse diseases. *Journal of Controlled Release*. 2022 May;345:314–33.
213. Bolourani S, Brenner M, Wang P, McGinn T, Hirsch JS, Barnaby D, et al. A Machine Learning Prediction Model of Respiratory Failure Within 48 Hours of Patient Admission for COVID-19: Model Development and Validation. *J Med Internet Res*. 2021 Feb 10;23(2):e24246.
214. Khuluq H, Yusuf PA, Perwitasari DA. Effectiveness of Machine Learning for COVID-19 Patient Mortality Prediction Using WEKA. *Glob Med Health Commun [Internet]*. 2023 Dec 25 [cited 2025 May 4];11(3). Available from: <https://ejournal.unisba.ac.id/index.php/gmhc/article/view/12119>
215. Bendavid I, Statlender L, Shvartsen L, Tepler S, Azullay R, Sapir R, et al. A novel machine learning model to predict respiratory failure and invasive mechanical ventilation in critically ill patients suffering from COVID-19. *Sci Rep*. 2022 Jun 22;12(1):10573.
216. Yin J, Yuan N, Huang Z, Hu Z, Bao Q, Shao Z, et al. Assessment of hypokalemia and clinical prognosis in Patients with COVID-19 in Yangzhou, China. Raghav A, editor. *PLoS ONE*. 2022 Jul 8;17(7):e0271132.

217. Nahkuri S, Becker T, Schueller V, Massberg S, Bauer-Mehren A. Prior fluid and electrolyte imbalance is associated with COVID-19 mortality. *Commun Med*. 2021 Nov 25;1(1):51.
218. Murdaca G, Di Gioacchino M, Greco M, Borro M, Paladin F, Petrarca C, et al. Basophils and Mast Cells in COVID-19 Pathogenesis. *Cells*. 2021 Oct 14;10(10):2754.
219. Várnai C, Palles C, Arnold R, Curley HM, Purshouse K, Cheng VWT, et al. Mortality Among Adults With Cancer Undergoing Chemotherapy or Immunotherapy and Infected With COVID-19. *JAMA Netw Open*. 2022 Feb 21;5(2):e220130.
220. Desai A, Gupta R, Advani S, Ouellette L, Kuderer NM, Lyman GH, et al. Mortality in hospitalized patients with cancer and coronavirus disease 2019: A systematic review and meta-analysis of cohort studies. *Cancer*. 2021 May;127(9):1459–68.
221. Bradley J, Sbaih N, Chandler TR, Furmanek S, Ramirez JA, Cavallazzi R. Pneumonia Severity Index and CURB-65 Score Are Good Predictors of Mortality in Hospitalized Patients With SARS-CoV-2 Community-Acquired Pneumonia. *Chest*. 2022 Apr;161(4):927–36.
222. Hassan S, Ramspek CL, Ferrari B, Van Diepen M, Rossio R, Knevel R, et al. External validation of risk scores to predict in-hospital mortality in patients hospitalized due to coronavirus disease 2019. *European Journal of Internal Medicine*. 2022 Aug;102:63–71.
223. Ilg A, Moskowitz A, Konanki V, Patel PV, Chase M, Grossestreuer AV, et al. Performance of the CURB-65 Score in Predicting Critical Care Interventions in Patients Admitted With Community-Acquired Pneumonia. *Annals of Emergency Medicine*. 2019 Jul;74(1):60–8.
224. Alwazzeah MJ, Subbarayalu AV, Abu Ali BM, Alabdulqader R, Alhajri M, Alwarthan SM, et al. Performance of CURB-65 and ISARIC 4C mortality scores for hospitalized patients with confirmed COVID-19 infection in Saudi Arabia. *Informatics in Medicine Unlocked*. 2023;39:101269.
225. Natanov D, Avihai B, McDonnell E, Lee E, Cook B, Altomare N, et al. Predicting COVID-19 prognosis in hospitalized patients based on early status. Moscona A, editor. *mBio*. 2023 Oct 31;14(5):e01508-23.
226. Jamshidi E, Asgary A, Tavakoli N, Zali A, Setareh S, Esmaily H, et al. Using Machine Learning to Predict Mortality for COVID-19 Patients on Day 0 in the ICU. *Front Digit Health*. 2022 Jan 13;3:681608.
227. Pigoga JL, Omer YO, Wallis LA. Derivation of a Contextually-Appropriate COVID-19 Mortality Scale for Low-Resource Settings. *Annals of Global Health*. 2021 Mar 26;87(1):31.
228. Ponce D, De Andrade LGM, Claire-Del Granado R, Ferreiro-Fuentes A, Lombardi R. Development of a prediction score for in-hospital mortality in COVID-19 patients with acute kidney injury: a machine learning approach. *Sci Rep*. 2021 Dec 24;11(1):24439.
229. Almenara Tejederas M, Luisa Serrano Salazar M, Rosa Melero Martin M, Molina-Andújar A, Ramos Terrada N, Poch Lopez De Briñas E, et al. MO328: Acute Renal Failure in COVID-19: Aetiology and Renal Evolution. AKI-COVID Registry of the Spanish Society of Nephrology. *Nephrology Dialysis Transplantation*. 2022 May 3;37(Supplement_3):gfac068.038.

230. Aktar S, Talukder A, Ahamad MdM, Kamal AHM, Khan JR, Protikuzzaman Md, et al. Machine Learning Approaches to Identify Patient Comorbidities and Symptoms That Increased Risk of Mortality in COVID-19. *Diagnostics*. 2021 Jul 31;11(8):1383.
231. Xu W, Sun NN, Gao HN, Chen ZY, Yang Y, Ju B, et al. Risk factors analysis of COVID-19 patients with ARDS and prediction based on machine learning. *Sci Rep*. 2021 Feb 3;11(1):2933.
232. Batra R, Whalen W, Alvarez-Mulett S, Gomez-Escobar LG, Hoffman KL, Simmons W, et al. Multi-omic comparative analysis of COVID-19 and bacterial sepsis-induced ARDS. Yount JS, editor. *PLoS Pathog*. 2022 Sep 19;18(9):e1010819.
233. Nilsson B, Persson B, Eriksson O, Fromell K, Hultström M, Frithiof R, et al. How the Innate Immune System of the Blood Contributes to Systemic Pathology in COVID-19-Induced ARDS and Provides Potential Targets for Treatment. *Front Immunol*. 2022 Mar 8;13:840137.
234. Qian G, Fang H, Chen A, Sun Z, Huang M, Luo M, et al. A hub gene signature as a therapeutic target and biomarker for sepsis and geriatric sepsis-induced ARDS concomitant with COVID-19 infection. *Front Immunol*. 2023 Sep 26;14:1257834.
235. Fralick M, Colacci M, Munshi L, Venus K, Fidler L, Hussein H, et al. Prone positioning of patients with moderate hypoxaemia due to covid-19: multicentre pragmatic randomised trial (COVID-PRONE). *BMJ*. 2022 Mar 23;e068585.
236. Qassim SH, Chemaitelly H, Ayoub HH, Coyle P, Tang P, Yassine HM, et al. Population immunity of natural infection, primary-series vaccination, and booster vaccination in Qatar during the COVID-19 pandemic: an observational study. *eClinicalMedicine*. 2023 Aug;62:102102.
237. Diani S, Leonardi E, Cavezzi A, Ferrari S, Iacono O, Limoli A, et al. SARS-CoV-2—The Role of Natural Immunity: A Narrative Review. *JCM*. 2022 Oct 25;11(21):6272.
238. Spinardi JR, Srivastava A. Hybrid Immunity to SARS-CoV-2 from Infection and Vaccination—Evidence Synthesis and Implications for New COVID-19 Vaccines. *Biomedicines*. 2023 Jan 27;11(2):370.
239. Soy M, Keser G, Atagündüz P, Tabak F, Atagündüz I, Kayhan S. Cytokine storm in COVID-19: pathogenesis and overview of anti-inflammatory agents used in treatment. *Clin Rheumatol*. 2020 Jul;39(7):2085–94.
240. Rabaan AA, Al-Ahmed SH, Muhammad J, Khan A, Sule AA, Tirupathi R, et al. Role of Inflammatory Cytokines in COVID-19 Patients: A Review on Molecular Mechanisms, Immune Functions, Immunopathology and Immunomodulatory Drugs to Counter Cytokine Storm. *Vaccines*. 2021 Apr 29;9(5):436.
241. Marinetti I, Jdanov D, Grigoriev P, Klüsener S, Janssen F. Effects of the COVID-19 pandemic on life expectancy and premature mortality in the German federal states in 2020 and 2021. Dávila-Cervantes CA, editor. *PLoS ONE*. 2023 Dec 21;18(12):e0295763.
242. Wang H, Paulson KR, Pease SA, Watson S, Comfort H, Zheng P, et al. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. *The Lancet*. 2022 Apr;399(10334):1513–36.
243. Paganuzzi M, Nattino G, Ghilardi GI, Costantino G, Rossi C, Cortellaro F, et al. Assessing the heterogeneity of the impact of COVID-19 incidence on all-cause excess

Referencias

- mortality among healthcare districts in Lombardy, Italy, to evaluate the local response to the pandemic: an ecological study. *BMJ Open*. 2024 Feb;14(2):e077476.
244. De Jong VMT, Rousset RZ, Antonio-Villa NE, Buenen AG, Van Calster B, Bello-Chavolla OY, et al. Clinical prediction models for mortality in patients with covid-19: external validation and individual participant data meta-analysis. *BMJ*. 2022 Jul 12;e069881.

