

ORIGINAL

Modelos de riesgo para la predicción de mortalidad hospitalaria en ancianos con neumonía por COVID-19

Raúl López-Izquierdo^{1,2}, Tomás Ruiz Albi³, Jesús Francisco Bermejo-Martín^{4,5}, Raquel Almansa^{4,5}, Fátima Victoria Villafaña Sanz¹, Lucía Arroyo Olmedo³, Carolina Andrea Urbina Carrera³, Susana Sánchez Ramón^{1,6}, Francisco Martín-Rodríguez^{7,8}, Fernando Moreno Torrero³, Daniel Álvarez^{3,9,10}, Félix del Campo Matía^{3,6,9}

Objetivo. Los objetivos son comparar la utilidad pronóstica de tres escalas de gravedad (Pneumonia Severity Index: PSI; CURB-65 scale; Severity Community Acquired Pneumonia Score: SCAP) y diseñar un nuevo modelo predictivo de mortalidad hospitalaria en pacientes mayores de 75 años ingresados por neumonía por COVID-19.

Método. Estudio retrospectivo de pacientes mayores de 75 años ingresados por neumonía por COVID-19 desde el servicio de urgencias entre el 12 de marzo y el 27 de abril de 2020. Se recogieron variables demográficas (edad, sexo, institucionalización), clínicas (síntomas, comorbilidades, índice de Charlson) y analíticas (bioquímica en suero, gasometría, hematemetría, hemostasia). Se derivó un modelo de riesgo y se compararon las escalas de gravedad PSI, CURB-65 y SCAP para predecir la mortalidad intrahospitalaria por cualquier causa.

Resultados. Se incluyeron 186 pacientes, con una mediana de edad de 85 años (RIC 80-89), un 44,1% varones. La mortalidad fue del 47,3%. Las escalas PSI, CURB-65 y SCAP tuvieron un área bajo la curva (ABC) de 0,74 (IC 95% 0,64-0,82), 0,71 (IC 95% 0,62-0,79) y 0,72 (IC 95% 0,63-0,81), respectivamente. El modelo predictivo compuesto por la ausencia o presencia de síntomas (tos y disnea), comorbilidad (índice de Charlson) y datos analíticos (aspartato-aminotransferasa, potasio, urea y lactato-deshidrogenasa) tuvo un ABC de 0,81 (IC 95% 0,73-0,88).

Conclusión. Este estudio muestra que la escala PSI tiene una capacidad predictiva de mortalidad moderada, notablemente mejor que las escalas CURB-65 y SCAP. Se propone un nuevo modelo predictivo de mortalidad que mejora significativamente el rendimiento de estas escalas, siendo necesario verificar su validez externa.

Palabras clave: COVID-19. Pronóstico. Mortalidad. Escalas de gravedad. Modelo predictivo.

Risk models for predicting in-hospital mortality from COVID-19 pneumonia in the elderly

Objective. To compare the prognostic value of 3 severity scales: the Pneumonia Severity Index (PSI), the CURB-65 pneumonia severity score, and the Severity Community-Acquired Pneumonia (SCAP) score. To build a new predictive model for in-hospital mortality in patients over the age of 75 years admitted with pneumonia due to the coronavirus disease 2019 (COVID-19).

Methods. Retrospective study of patients older than 75 years admitted from the emergency department for COVID-19 pneumonia between March 12 and April 27, 2020. We recorded demographic (age, sex, living in a care facility or not), clinical (symptoms, comorbidities, Charlson Comorbidity Index [CCI]), and analytical (serum biochemistry, blood gases, blood count, and coagulation factors) variables. A risk model was constructed, and the ability of the 3 scales to predict all-cause in-hospital mortality was compared.

Results. We included 186 patients with a median age of 85 years (interquartile range, 80–89 years); 44.1% were men. Mortality was 47.3%. The areas under the receiver operating characteristic curves (AUCs) were as follows for each tool: PSI, 0.74 (95% CI, 0.64–0.82); CURB-65 score, 0.71 (95% CI, 0.62–0.79); and SCAP score, 0.72 (95% CI, 0.63–0.81). Risk factors included in the model were the presence or absence of symptoms (cough, dyspnea), the CCI, and analytical findings (aspartate aminotransferase, potassium, urea, and lactate dehydrogenase). The AUC for the model was 0.81 (95% CI, 0.73–0.88).

Conclusions. This study shows that the predictive power of the PSI for mortality is moderate and perceptibly higher than the CURB-65 and SCAP scores. We propose a new predictive model for mortality that offers significantly better performance than any of the 3 scales compared. However, our model must undergo external validation.

Keywords: COVID-19. Prognosis. Mortality. Severity scales. Predictive model.

Introducción

La infección por el nuevo coronavirus SARS-CoV-2 responsable de los primeros casos de COVID-19 en la

región de Wuhan (China) a finales del año 2019 se ha convertido en una pandemia de gran impacto sociosanitario¹. El número de pacientes afectados ha aumentado de forma progresiva, al igual que la mortalidad^{2,3}.

Filiación de los autores:

¹Servicio de Urgencias, Hospital Universitario Río Hortega, Valladolid, España.

²Departamento de Cirugía, Oftalmología, Otorrinolaringología y Fisioterapia, Facultad de Medicina, Universidad de Valladolid, España.

³Servicio de Neumología, Hospital Universitario Río Hortega, Valladolid, España.

⁴Grupo de Investigación Biomédica en Infección Respiratoria y Sepsis (Biosepsis) (BSAL), España.

⁵Centro de Investigación Biomédica en Red en Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, Madrid, España.

⁶Departamento de Medicina, Dermatología y Toxicología, Facultad de Medicina, Universidad de Valladolid, España

⁷Unidad Móvil de Emergencias, Gerencia de Emergencias Sanitarias de Castilla y León (SACYL), España.

⁸Centro de Simulación Clínica Avanzada, Facultad de Medicina, Universidad de Valladolid, Valladolid, España.

⁹Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Instituto de Salud Carlos III, Madrid, España.

¹⁰Grupo de Ingeniería Biomédica (GIB), Universidad de Valladolid, Valladolid, España.

Contribución de los autores:

Todos los autores han confirmado su autoría en el documento de responsabilidades del autor, acuerdo de publicación y cesión de derechos a EMERGENCIAS.

Autor para correspondencia:

Daniel Álvarez
Servicio de Neumología
Hospital Universitario Río Hortega
C/ Dulzaina, 2
47012 Valladolid, España

Correo electrónico:

dalvarezgo@saludcastillayleon.es

Información del artículo:

Recibido: 23-10-2020

Aceptado: 30-1-2021

Online: 9-6-2021

Editor responsable:

Francisco Javier Martín-Sánchez

Ello ha supuesto una alarma sanitaria que ha excedido a este ámbito, y se ha convertido en un problema social y económico, y ha evidenciado limitaciones de los recursos sanitarios en los momentos de mayor presión asistencial^{4,5}.

A pesar del escaso conocimiento que se tiene sobre esta nueva enfermedad, parece evidente que los pacientes que presentan una mala evolución lo hacen en la gran mayoría de los casos desarrollando a lo largo de los días un síndrome de distrés respiratorio agudo y un fallo multiorgánico asociado⁶. Se ha descrito una serie de factores de riesgo, tanto clínicos como analíticos, asociados a una mala evolución de estos pacientes⁷⁻⁹.

En los primeros meses de la pandemia, un hecho que se ha puesto de manifiesto es que la edad es un factor independiente de mortalidad¹⁰⁻¹³. Dada su vulnerabilidad, la atención a la población anciana en los servicios de urgencias hospitalarios (SUH) es de especial importancia, sobre todo en mayores de 75 años¹⁴, donde la mortalidad asociada al COVID-19 se incrementa de forma exponencial¹⁵. Esto ha hecho que, en la actualidad, la infección por el SARS-CoV-2 constituya una auténtica emergencia geriátrica¹⁶.

Por tanto, en el contexto epidemiológico actual sería importante conocer qué herramientas se podrían utilizar para la valoración inicial de este grupo de pacientes. Desde hace años se han descrito una serie de escalas de gravedad que se utilizan de forma habitual para la valoración del riesgo de los pacientes diagnosticados de neumonía, entre las que destacan la Pneumonia Severity Index (PSI)¹⁷, la escala CURB-65¹⁸ o la Severity Community Acquired Pneumonia score (SCAP)¹⁹⁻²¹. Algunas de ellas (PSI y CURB-65) han sido recomendadas por las autoridades sanitarias españolas para su utilización en la infección por COVID-19²². Además, estas escalas resultan muy útiles entre la población anciana^{23,24}. Sin embargo, son escasos los trabajos que han evaluado su utilidad entre la población general diagnosticada de una infección por COVID-19, y se han empleado para valorar la necesidad de soporte respiratorio^{25,26} o la mortalidad²⁷. Los resultados son discordantes^{25,28}, por lo que realmente se desconoce su utilidad real, especialmente entre la población de mayor edad²⁹. El objetivo del presente trabajo es estudiar la utilidad pronóstica de estas escalas de gravedad, así como desarrollar un nuevo modelo predictivo en relación con la mortalidad hospitalaria por cualquier causa en los pacientes mayores de 75 años que requirieron ingreso hospitalario por neumonía por COVID-19.

Método

Estudio descriptivo retrospectivo de una muestra de conveniencia en un SUH. El estudio fue aprobado por el Comité de Ética de la Investigación con medicamentos (CEIm) del Área de Salud Valladolid Oeste y se desarrolló de acuerdo con los principios expresados en la declaración de Helsinki.

Se incluyeron de forma consecutiva pacientes mayores de 75 años que requirieron ingreso con diagnóstico de neumonía por COVID-19, confirmada mediante la reacción en cadena de la polimerasa (RT-PCR), desde el 12 de marzo de 2020 hasta el 27 de abril de 2020. Se definió como neumonía la presencia o la aparición de un nuevo infiltrado radiológico en comparación con radiografías previas si estas existían, bien mediante radiografía torácica interpretada por el médico de urgencias o mediante informe del servicio de radiología o tomografía computarizada (TC) de tórax en base al informe del servicio de radiología y clínica compatible. Se excluyeron a los pacientes sin confirmación mediante RT-PCR positiva a su llegada al SUH, aunque posteriormente se positivizara en el transcurso del ingreso hospitalario, así como a aquellos que se diagnosticaron de neumonía durante su estancia hospitalaria pero que no lo habían sido durante la valoración inicial en urgencias. Con respecto al diagnóstico, se tuvo en cuenta el diagnóstico al alta, siendo reevaluado por los investigadores. Todos los pacientes con datos perdidos en las variables necesarias para el cálculo de las escalas de gravedad fueron eliminados del estudio.

Se realizó una revisión de las historias clínicas digitalizadas de cada uno de los pacientes, registrándose las siguientes variables independientes: (i) variables sociodemográficas: edad, género e institucionalización; (ii) variables clínicas: sintomatología asociada (tos, fiebre, disnea, dolor torácico, anosmia, odinofagia, síntomas digestivos, mareos/síncope, síntomas digestivos), comorbilidades e índice de comorbilidad de Charlson; (iii) variables relacionadas con el cálculo de las escalas: PSI, CURB-65, SCAP; (iv) variables analíticas: recuento y porcentaje de células blancas, hematíes, hemoglobina, hematocrito, plaquetas, dímero D, índice de protrombina (International Normalized Ratio, INR), ácido láctico, urea, creatinina, glucosa, sodio, potasio, bilirrubina, alanina aminotransferasa (ALT/GPT), aspartato aminotransferasa (AST/GOT), procalcitonina, proteína C reactiva (PCR), albumina y lactato deshidrogenasa. La variable de resultado principal fue la mortalidad hospitalaria por cualquier causa recopilada por investigadores independientes.

Los datos recopilados fueron analizados mediante el programa estadístico SPSS Statistics 24.0 (IBM, Armonk, NY, EE.UU.), empleado para realizar un análisis descriptivo de los mismos, evaluar las diferencias entre las distribuciones de los grupos a estudio y diseñar el modelo predictivo, y MATLAB R2020a (The MathWorks, Natick, MA, EE.UU.), y se empleó para realizar la validación interna del modelo y estimar las métricas de rendimiento predictivo ajustadas. Además, se empleó el paquete *rms* de R (versión 6.1-0) para calcular la curva de calibración del modelo. Las variables cuantitativas continuas se describieron como media y desviación estándar (DE), en caso de mostrar una distribución normal, o como mediana y rango intercuartil (RIC) si la distribución no era normal, utilizando el test de Kolmogorov-Smirnov como prueba de normalidad. Las variables cualitativas se describieron mediante frecuencias absolutas y porcentajes (%). Para la comparativa de las variables cuantitativas, se utilizó la prueba t de Student para los valo-

res distribuidos normalmente y el test no paramétrico de Mann-Whitney para las variables que no cumplieron los criterios de normalidad. Se utilizó la prueba de ji cuadrado para el contraste de las variables cualitativas.

Para diseñar el modelo predictivo de mortalidad se utilizó regresión logística binaria. Se realizó un análisis multivariante, tomando como variable dependiente principal la mortalidad durante el ingreso hospitalario y como variables independientes un conjunto inicial de factores de riesgo compuesto por un total de 56 variables (socio-demográficas, síntomas, comorbilidades, analíticas). Todas las variables de entrada fueron estandarizadas (media nula y varianza unidad) para evitar diferencias en los rangos y ordenes de magnitud de los datos que pudieran afectar a la estimación de los valores de los coeficientes. Previamente, se redujo el número de predictores candidatos mediante técnicas de reducción de dimensionalidad. En primer lugar, se seleccionaron las variables que individualmente mostraron las diferencias estadísticamente más significativas entre los grupos de supervivientes (S) y no supervivientes (NS) ($p < 0,05$). A continuación, se realizó un análisis de colinealidad y de multicolinealidad en base a la matriz de correlaciones (ρ_{ij}) y al cálculo de la tolerancia (T) y del factor de inflación de la varianza (FIV) para cada predictor candidato a entrar en el modelo de regresión logística, respectivamente. Finalmente, se aplicó el algoritmo de selección automática de variables por pasos *forward-selection backward elimination* para determinar el subconjunto óptimo de predictores³⁰. Como umbrales de probabilidad de entrada y salida de variables del modelo durante el proceso de selección se emplearon 0,10 y 0,25, respectivamente.

Se realizó una validación interna del modelo mediante el procedimiento de remuestreo (*bootstrapping*) *bootstrap 0.632*³¹⁻³³. Para estimar el intervalo de confianza del 95% con una potencia estadística adecuada, se fijó un número de remuestreos de 1.000³⁴.

La capacidad predictiva de las escalas bajo estudio y del modelo propuesto se evaluó en términos de sensibilidad, especificidad, valores predictivos positivo y negativo, razón de verosimilitud positiva y negativa, precisión y área bajo la curva (ABC) de la característica operativa del receptor (ROC). En todas las métricas de rendimiento predictivo se corrigió el optimismo asociado a un potencial sobentrenamiento mediante *bootstrap 0.632*. Se utilizó la prueba no paramétrica de Kruskal-Wallis para el estudio comparativo de las distintas curvas ROC, realizando un análisis *post-hoc* mediante la corrección de Bonferroni en las comparativas dos a dos. Se ha considerado significativo un valor de $p < 0,05$. Se emplearon curvas de calibración para evaluar la tendencia del modelo hacia sobrestimaciones o infraestimaciones sistemáticas/extremas de las predicciones. Se realizó un análisis mediante curvas de decisión para comparar la utilidad clínica de los algoritmos bajo estudio en términos de beneficio neto³⁵. La descripción del diseño y de los resultados relativos a la evaluación del modelo predictivo se han realizado de acuerdo con las recomendaciones de la guía TRIPOD³⁶.

Resultados

Un total de 231 pacientes cumplían los criterios de inclusión en el periodo de estudio. De ellos, 45 fueron descartados por datos perdidos (Figura 1). Se incluyeron un total de 186 pacientes, con una edad mediana de 85 años (RIC 80-89), de los cuales un 44,1% eran varones. Un 2,2% de los pacientes ingresaron directamente en la unidad de cuidados intensivos (UCI) desde urgencias. El síntoma que más frecuentemente presentaban fue la disnea (66,7%), seguido de la fiebre (61,3%), la tos (47,9%) y los síntomas digestivos (19,4%). La comorbilidad más frecuente fue la hipertensión arterial (76,3%). La Tabla 1 recoge los datos demográficos, clínicos y analíticos de la población de estudio.

El total de fallecidos durante el ingreso fue de 88 (47,3%) pacientes. Las variables demográficas y clínicas que se asociaron a mortalidad hospitalaria fueron la edad, la institucionalización, la tos, la disnea y el grado de comorbilidad según el índice de Charlson (Tabla 1). No se encontraron diferencias significativas en relación al estado de inmunosupresión previo al ingreso (tumores: S: 13,0% vs NS: 13,8%, $p = 0,875$; leucemia: S: 2,2% vs NS: 2,1%, $p = 0,983$; inmunosupresión: S: 4,3% vs NS: 2,1%, $p = 0,185$). Con respecto al tratamiento durante el ingreso, se administraron corticoides (S: 18,7% en vs NS: 30,0%; $p = 0,760$) y antibióticos (levofloxacino: S: 6,2% vs NS: 10,5%, $p = 0,292$; azitromicina: S: 68,1% vs NS: 59,0%, $p = 0,240$; ceftriaxona: S: 72,5% vs NS: 64,0%, $p = 0,220$), sin diferencias estadísticamente significativas entre los grupos de pacientes a estudio. En cuanto a las variables analíticas, la mortali-

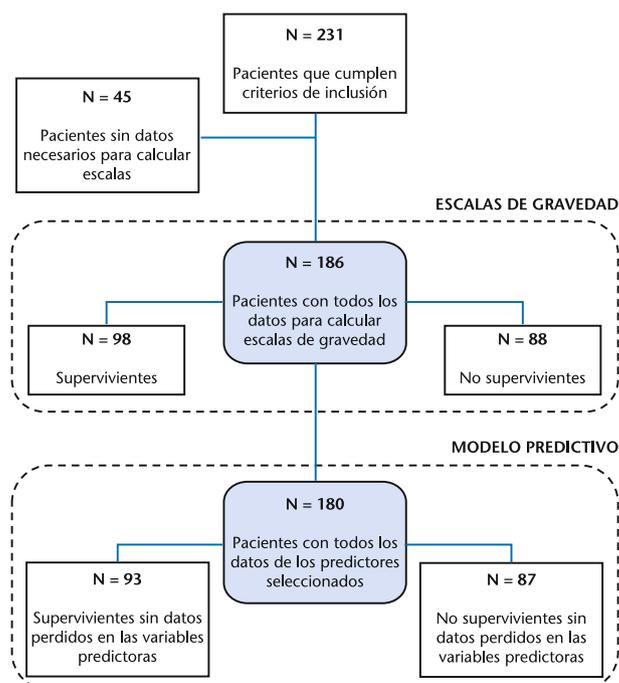


Figura 1. Diagrama de flujo de los pacientes analizados en el periodo de estudio.

Tabla 1. Características demográficas y clínicas de los grupos de población bajo estudio

	Total N = 186 n (%)	Supervivientes N = 98 n (%)	No supervivientes N = 88 n (%)	p
Edad (años) [mediana (RIC)]	85 [80-89]	84 [79-89]	87 [82-91]	0,001
Sexo (varones)	82 (44,1)	40 (40,8)	42 (47,7)	0,343
Institucionalización	101 (54,3)	44 (44,9)	57 (64,8)	0,007
Síntomas				
Tos	89 (47,9)	57 (58,2)	32 (36,4)	0,002
Fiebre	114 (61,3)	55 (56,1)	59 (67,1)	0,149
Disnea	124 (66,7)	55 (56,1)	69 (78,4)	0,002
Dolor torácico	5 (2,7)	3 (3,1)	2 (2,3)	0,731
Anosmia	4 (2,2)	4 (4,1)	0 (0,0)	0,054
Odinofagia/Síntomas catarrales	12 (6,5)	7 (7,1)	5 (5,7)	0,672
Síntomas digestivos	36 (19,4)	23 (23,5)	13 (14,8)	0,125
Síncope/Mareo	13 (7,0)	6 (6,1)	7 (8,0)	0,638
Síntomas dermatológicos	2 (1,1)	2 (2,0)	0 (0,0)	0,176
Comorbilidades				
Hipertensión arterial	142 (76,3)	76 (77,6)	66 (75,0)	0,683
Dislipemia	64 (34,4)	38 (38,8)	26 (29,6)	0,186
Infarto de miocardio	19 (10,2)	8 (8,2)	11 (12,5)	0,330
Insuficiencia cardíaca	21 (11,3)	8 (8,2)	13 (14,8)	0,155
Enfermedad arterial periférica	8 (4,3)	5 (5,1)	3 (3,4)	0,570
Enfermedad cerebrovascular	29 (15,6)	14 (14,3)	15 (17,1)	0,605
Deterioro cognitivo crónico	56 (30,1)	24 (24,5)	32 (36,4)	0,078
Enfermedad respiratoria crónica	30 (16,1)	16 (16,3)	14 (15,9)	0,938
Enfermedad tejido conectivo	4 (2,2)	1 (1,0)	3 (3,4)	0,262
Úlcera gastroduodenal	14 (7,5)	6 (6,1)	8 (9,1)	0,444
Hepatopatía crónica	3 (1,6)	0 (0,0)	3 (3,4)	0,324
Diabetes mellitus	44 (23,7)	20 (20,4)	24 (27,3)	0,234
Hemiplejía	6 (3,2)	3 (3,1)	3 (3,4)	0,893
Enfermedad renal crónica	35 (18,8)	16 (16,3)	19 (21,6)	0,359
Tumor	25 (13,4)	12 (12,3)	13 (14,8)	0,614
Leucemia	4 (2,2)	3 (3,1)	1 (1,1)	0,366
Linfoma	2 (1,1)	1 (1,0)	1 (1,1)	0,939
SIDA definido	1 (0,5)	0 (0,0)	1 (1,1)	0,290
Índice de Charlson	5 [4-6]	4 [3-6]	5 [4-7]	0,006
Valor de las escalas [mediana (RIC)]				
PSI	121 [102-147]	109 [88-135]	138 [117-164]	< 0,001
CURB-65	2 [2-3]	2 [1-3]	3 [2-3]	< 0,001
SCAP	15 [10-20]	10 [5-16]	16 [12-24]	< 0,001

PSI: Pneumonia severity index; SCAP: Severity Community Acquired Pneumonia score.

dad hospitalaria se asoció con presentar un menor valor de linfocitos, monocitos y albúmina, así como valores más elevados de INR, ácido láctico, creatinina, AST/GOT, potasio, PCR, procalcitonina, urea y LDH (Tabla 2).

En la Tabla 3 se muestra el análisis descriptivo de las variables involucradas en el cálculo de las escalas de gravedad PSI, CURB65 y SCAP en los grupos estudiados.

Durante el análisis multivariante con selección automática de características se descartaron 6 pacientes por falta de datos en las variables predictoras seleccionadas, por lo que el tamaño efectivo de la muestra empleada para diseñar el modelo fue de 180 pacientes (Figura 1). No se observaron indicios de colinealidad ($\rho_{ij} < 0,8$) ni de multicolinealidad ($T > 0,1$; $FIV < 5$) entre los predictores candidatos a entrar en el modelo. Las variables seleccionadas asociadas a la mortalidad intrahospitalaria fueron la disnea, tos, índice de Charlson, AST/GOT, potasio, urea y LDH. La influencia individual de cada predictor y los coeficientes del modelo se detallan en la

Tabla 4. La Figura 2 muestra las curvas de calibración (optimista y corregida) del modelo predictivo propuesto. El sesgo (*intercept*) de la curva es igual a 0,0557 (siendo 0 el valor ideal) mientras que la pendiente es 0,8926 (siendo 1 el valor objetivo). Dado que el sesgo es muy cercano a 0, nuestro modelo no muestra una tendencia hacia valores sobrestimados o infraestimados de las predicciones. De forma similar, el valor de la pendiente de la calibración sugiere que las estimaciones no son sistemáticamente demasiado extremas.

En relación con las escalas de gravedad bajo estudio, se observa que los pacientes que no sobrevivieron alcanzaron un valor mediano significativamente superior ($p < 0,001$) en las tres escalas analizadas frente al grupo de supervivientes, siendo la mortalidad significativamente mayor en los grupos de más gravedad ($p < 0,001$), independientemente de la escala utilizada (Tabla 5).

La Tabla 6 muestra las métricas de rendimiento predictivo procedentes de la validación interna me-

Tabla 2. Variables derivadas de la analítica para los grupos de población bajo estudio

	Total N = 186 Mediana (RIC)	Supervivientes N = 98 Mediana (RIC)	No supervivientes N = 88 Mediana (RIC)	p
Leucocitos (x 10 ³ /microL)	6,90 (4,80-10,00)	6,90 (5,00-9,30)	6,85 (4,75-10,80)	0,927
Neutrófilos (x 10 ³ /microL)	5,40 (3,40-8,30)	5,30 (3,40-7,60)	5,70 (3,40-9,65)	0,272
Linfocitos (x 10 ³ /microL)	0,80 (0,60-1,10)	0,90 (0,60-1,20)	0,70 (0,60-1,00)	0,004
Monocitos (x 10 ³ /microL)	0,40 (0,30-0,60)	0,50 (0,40-0,70)	0,40 (0,30-0,50)	< 0,001
Hematíes (x 10 ⁶ /microL)	4,55 (4,03-4,89)	4,59 (4,20-4,86)	4,47 (3,92-4,90)	0,367
Hemoglobina (g/dL)	13,2 (11,9-14,2)	13,2 (12,0-14,2)	13,1 (11,7-14,3)	0,770
Hematocrito (%)	40,1 (36,5-43,3)	40,1 (36,9-43,2)	40,1 (35,3-43,9)	0,991
Plaquetas (x 10 ³ /microL)	175,5 (129,0-231,0)	179,5 (140,0-242,0)	173,5 (121,0-228,0)	0,229
Tiempo de protrombina (INR)	1,15 (1,07-1,37)	1,13 (1,06-1,25)	1,19 (1,09-1,47)	0,010
Fibrinógeno (mg/dL)	667 (579-790)	658 (582-796)	667 (576-788)	0,817
Dímero D (ng/mL)	636 (374-1.019)	587 (352-968)	685 (420-1689)	0,101
Bilirrubina total (mg/dL)	0,59 (0,45-0,78)	0,58 (0,46-0,82)	0,60 (0,45-0,74)	0,849
Calcio iónico (mg/dL)	4,53 (4,38-4,67)	4,55 (4,40-4,68)	4,53 (4,35-4,65)	0,365
Albumina (g/dL)	3,40 (3,20-3,68)	3,50 (3,30-3,70)	3,40 (3,10-3,60)	0,043
Ácido láctico (mmol/L)	1,60 (1,20-2,08)	1,50 (1,20-1,80)	1,90 (1,20-2,33)	0,004
Creatinina (mg/dL)	1,10 (0,84-1,52)	1,00 (0,76-1,40)	1,25 (0,90-1,88)	0,002
Glucosa (mg/dL)	126 (108-162)	122 (106-149)	132 (109-182)	0,061
ALT/GPT (U/L)	21,0 (15,0-31,0)	21,0 (15,0-30,0)	22,0 (15,5-32,0)	0,488
AST/GOT (U/L)	38,0 (28,0-51,2)	33,0 (25,2-41,0)	44,0 (32,0-58,0)	< 0,001
Potasio (mmol/L)	4,0 (3,7-4,4)	3,9 (3,7-4,2)	4,2 (3,8-4,5)	< 0,001
Sodio (mmol/L)	135 (132-139)	135 (132-138)	136 (132-140)	0,191
PCR (mg/L)	111 (58-182)	100 (39-153)	123 (83-195)	0,017
Procalcitonina (ng/mL)	0,15 (0,07-0,36)	0,10 (0,06-0,23)	0,23 (0,10-0,52)	< 0,001
Urea (mg/dL)	63,7 (43,9-93,2)	52,6 (38,2-81,5)	73,2 (55,4-104,3)	< 0,001
Lactato deshidrogenasa (U/L)	332 (254-462)	296 (237-371)	429 (306-494)	< 0,001

ALT/GPT: alanina aminotransferasa; AST/GOT: aspartato aminotransferasa; INR: International Normalized Ratio.

dante *bootstrap* 0.632 para el modelo propuesto y las distintas escalas de gravedad en relación con la mortalidad hospitalaria. En la Figura 3 se muestra la distribución de las estimaciones del estadístico ABC del modelo propuesto en los grupos de entrenamiento y test del procedimiento de validación interna *bootstrap* 0.632 empleado para ajustar (corregir optimismo en la estimación) cada una de las métricas de rendimiento. La utilización de los umbrales óptimos derivados de la curva ROC en lugar de los puntos de corte clásicos permitió obtener resultados más balanceados en términos del par sensibilidad-especificidad. De forma global, PSI mostró la mayor precisión, con un ABC de 0,74 (IC 95% 0,64-0,82, $p < 0,001$), seguido de la escala SCAP con un ABC de 0,72 (IC 95% 0,63-0,81, $p < 0,001$), mientras la escala CURB-65 presentó un ABC de 0,71 (IC 95% 0,62-0,79, $p < 0,001$) (Figura 4). En las comparaciones dos a dos, las curvas ROC de todas las escalas fueron significativamente diferentes entre sí ($p < 0,01$). El modelo predictivo basado en las variables seleccionadas automáticamente mostró la mayor capacidad predictiva con un ABC de 0,81 (IC 95% 0,73-0,88) (Figura 4) con diferencias estadísticamente significativas ($p < 0,01$) con respecto a las escalas de gravedad convencionales.

La Figura 5 muestra las curvas de decisión de los algoritmos a estudio. Se observa que el modelo propuesto muestra un beneficio neto superior a las escalas de gravedad clásicas evaluadas para todos los umbrales de probabilidad.

Discusión

Este estudio evalúa la utilidad de los factores de riesgo y las escalas PSI, CURB-65 y SCAP para la predicción de la mortalidad en los pacientes mayores de 75 años con neumonía por SARS-CoV-2. Los resultados revelan que estas escalas predicen la mortalidad hospitalaria con un rendimiento aceptable, si bien hay diferencias estadísticamente significativas entre ellas. La escala PSI presentó la mayor ABC ROC, mientras que SCAP y CURB-65 mostraron una menor capacidad predictiva. Por otra parte, el modelo de mortalidad hospitalaria obtenido a partir de los factores de riesgo de nuestra serie presentó una capacidad predictiva significativamente superior al resto de las escalas analizadas.

La mortalidad observada en el grupo de pacientes bajo estudio ha sido notablemente elevada (47,3%). La prevalencia de mortalidad descrita en la literatura es muy variable^{13,15,29}, dependiendo del grupo de edad estudiado, del país, del periodo de estudio de la mortalidad y si se evalúan pacientes ingresados¹³. Si bien ya se conocía que la edad por encima de 75 años resulta un factor de riesgo independiente de mortalidad entre los pacientes que acuden a los SUH¹⁴, es precisamente a partir de este grupo de edad en el que se ha producido una mayor mortalidad relativa y absoluta por el COVID-19¹⁵. Seguramente, porque es el colectivo que vive de forma mayoritaria en centros sociosanitarios, presentando mayor comorbilidad, fragilidad y discapacidad. Así, es en este grupo de edad en el que se ten-

Tabla 3. Análisis descriptivo de las variables involucradas en el cálculo de las escalas de gravedad PSI, CURB65 y SCAP en los grupos de población bajo estudio

	Inclusión en la escala			Total N = 186 n (%)	Supervivientes N = 98 n (%)	No supervivientes N = 88 n (%)	p
	PSI	CURB	SCAP				
Sexo (varones)	X			82 (44,1)	39 (42,4)	43 (45,8)	0,645
Edad (años) [mediana (RIC)]	X	X	X	85 (80-89)	84 (78-89)	87 (83-92)	0,002
Institucionalizado	X			101 (54,3)	38 (41,3)	63 (67,0)	0,0066
Nefropatía	X			35 (18,8)	15 (16,3)	20 (21,3)	0,386
Insuficiencia cardíaca	X			21 (11,3)	7 (7,6)	14 (14,9)	0,117
Enfermedad cerebrovascular	X			29 (15,6)	13 (14,1)	16 (17,0)	0,587
Hepatopatía	X			2 (1,1)	0 (0,0)	2 (2,1)	0,160
Neoplasia	X			25 (13,4)	12 (13,0)	13 (13,8)	0,875
Alteración del nivel de conciencia	X	X	X	56 (30,1)	14 (15,2)	42 (44,7)	< 0,001
Presión arterial sistólica (mmHg) [mediana (RIC)]	X	X	X	133 (120-147)	136 (120-150)	130 (120-145)	0,250
Presión arterial diastólica (mmHg) [mediana (RIC)]		X		70 (62-81)	71 (63-80)	69 (62-83)	0,856
Frecuencia respiratoria (respiraciones por minuto) [mediana (RIC)]	X	X	X	23 (16-28)	20 (15-26)	25 (20-32)	< 0,001
Frecuencia cardíaca (latidos por minuto) [mediana (RIC)]	X			88 (74-99)	88 (74-99)	88 (74-99)	0,983
Temperatura (°C) [mediana (RIC)]	X			36,4 (36,1-37,0)	36,5 (36,2-37,0)	36,3 (36,0-37,1)	0,386
pH [mediana (RIC)]	X		X	7,4 (7,4-7,5)	7,4 (7,4-7,5)	7,4 (7,4-7,5)	0,001
Hematocrito (%) [mediana (RIC)]	X			40,1 (36,5-43,3)	40,3 (37,0-43,1)	40,1 (35,3-44,1)	0,781
Sodio (mmol/L) [mediana (RIC)]	X			135 (132-139)	135 (132-138)	136 (132-140)	0,101
Urea (mg/dL) [mediana (RIC)]	X	X	X	64 (44-93)	52 (38-83)	73 (55-103)	< 0,001
PaO ₂ (mmHg) [mediana (RIC)]	X		X	63 (51-81)	65 (51-80)	58 (50-83)	0,490
PaO ₂ /FiO ₂ (mmHg) [mediana (RIC)]			X	262 (194-323)	281 (223-329)	243 (157-300)	0,046
Derrame pleural [mediana (RIC)]	X			18 (9,7)	7 (7,6)	11 (11,7)	0,345
Glucemia (mg/dL) [mediana (RIC)]	X			126 (108-162)	121 (106-150)	132 (110-178)	0,053
Anomalías Rx			X	108 (58,1)	53 (57,6)	55 (58,5)	0,901

Anomalías Rx: presencia de anomalías radiológicas en tórax; FiO₂: fracción inspirada de oxígeno; PaO₂: presión de oxígeno en sangre arterial; PSI: Pneumonia severity index; SCAP: Severity Community Acquired Pneumonia score.

drían que priorizar las medidas preventivas para disminuir las cifras de mortalidad alcanzadas en estos pacientes³⁷. Otro aspecto a tener en cuenta es que a la propia gravedad que supone la COVID-19 en esta población anciana se ha añadido la situación de sobrecarga sanitaria, con la consiguiente limitación de recursos asistenciales^{16,38}, siendo difícil precisar realmente la parte de responsabilidad de cada uno de estos factores.

Las escalas analizadas tienen una elevada sensibilidad (> 90%), particularmente al utilizar como punto de corte el valor que discrimina riesgo bajo vs riesgo moderado-alto, para detectar a los pacientes que realmente tienen un mayor riesgo de mortalidad. Sin embargo, la especificidad que nosotros hemos encontrado entre la población anciana es muy inferior a lo que se ha descrito en trabajos que incluye población más joven²⁷. La utilización de un umbral más óptimo para cada una de las escalas permite alcanzar un mejor balance entre sensibilidad y especificidad, mejorando notablemente la es-

pecificidad a expensas de una ligera disminución de la sensibilidad.

La escala PSI obtuvo el mejor comportamiento pronóstico, al presentar una mayor sensibilidad, valor predictivo negativo y precisión diagnóstica valorada mediante el ABC ROC. Estos hallazgos ya han sido descritos en estudios previos que han comparado de forma global el rendimiento pronóstico de las escalas PSI y CURB-65 entre los pacientes hospitalizados por COVID-19^{27,28}. Por otra parte, en un estudio en el que se comparaba la puntuación de la escala PSI entre la población joven frente a la anciana, se observó que los pacientes ancianos presentaban una mayor proporción de riesgo alto, datos que pueden compararse con nuestros resultados³⁹.

En relación con los valores de ABC ROC que hemos obtenido, estos son inferiores a lo descrito por otros autores cuando no tienen en cuenta la edad^{27,28}, lo que lleva a pensar que estas escalas se comportan mejor en la

Tabla 4. Modelo predictivo de mortalidad hospitalaria

Variables (N = 180)	Coeficientes ^a	OR	IC 95%		p
			Inferior	Superior	
Tos	$\beta_1 = -0,6582$	0,510	0,232	1,123	0,094
Disnea	$\beta_2 = 0,9045$	2,851	1,173	6,928	0,021
Índice de Charlson	$\beta_3 = 0,2429$	1,298	1,038	1,623	0,022
AST/GOT	$\beta_4 = 0,0310$	1,035	1,009	1,062	0,009
Potasio	$\beta_5 = 0,9984$	2,376	1,044	5,407	0,039
Urea	$\beta_6 = 0,0060$	1,007	0,996	1,017	0,204
LDH	$\beta_7 = 0,0035$	1,003	1,000	1,007	0,057

^aProbabilidad (mortalidad) = $1/(1 + e^{-(\beta_0 + \beta_1 \cdot \text{Tos} + \beta_2 \cdot \text{Disnea} + \beta_3 \cdot \text{Charlson} + \beta_4 \cdot \text{AST/GOT} + \beta_5 \cdot \text{Potasio} + \beta_6 \cdot \text{Urea} + \beta_7 \cdot \text{LDH})})$, siendo el término constante $\beta_0 = -8,6837$.

OR: odds ratio; IC: intervalo de confianza; %: porcentaje; AST/GOT: Aspartato aminotransferasa; LDH: lactato deshidrogenasa.

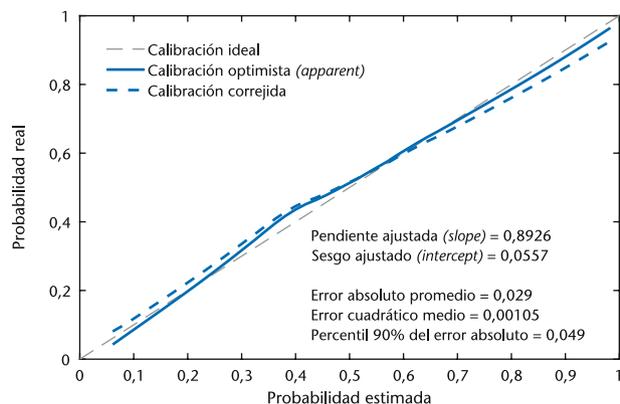


Figura 2. Curva de calibración del modelo predictivo y corrección mediante *bootstrap* para la escala desarrollada en el presente estudio.

población más joven, disminuyendo su capacidad pronóstica según aumenta la edad⁴⁰. Si estudiamos las curvas ROC de los trabajos que han valorado estas escalas en la neumonía adquirida en la comunidad entre la población general, observamos que se comportan de forma muy parecida a lo observado en el presente estudio y de la misma manera encuentran una ligera superioridad de la escala PSI con respecto a la escala CURB-65^{41,42}. En este sentido, la escala PSI involucra un número muy superior de variables, e incluye variables relacionadas con los antecedentes personales, analíticos y radiológicos, con una mayor complejidad que la CURB-65. En nuestro estudio, la escala CURB-65 presentó el mayor porcentaje de pacientes que fallecieron en el grupo pronóstico de riesgo bajo, lo que hace que esta escala pierda valor a la hora de decidir cuál es el lugar más adecuado para el tratamiento de estos pacientes⁴³.

Se propone emplear el modelo predictivo de mortalidad para identificar al grupo de pacientes mayores de 75 años con neumonía por COVID de mayor riesgo e instaurar un seguimiento más estrecho en aquellos rangos de probabilidad estimada en los que el modelo alcance un beneficio neto superior a las escalas existentes. En este sentido, analizando las curvas de decisión, es destacable que el modelo predictivo propuesto proporcionaría un beneficio neto superior a las escalas pronósticas en todos los umbrales de probabilidad. PSI, CURB-65 y SCAP muestran un comportamiento muy similar, con un beneficio neto superior a las estrategias "Ingreso de todos los pacientes" o "Ningún paciente es ingresado" únicamente en el rango de proba-

bilidades de 35% a 65%. Por el contrario, el modelo automático muestra un beneficio neto notablemente superior en todos los umbrales de probabilidad de mortalidad a partir del 20%.

En nuestro modelo de predicción de mortalidad, las variables incluidas han sido la ausencia de tos, presencia de disnea, la elevación de LDH, GOT, urea y potasio, así como el grado de comorbilidad valorado mediante el índice de Charlson. Alguno de estos factores ya ha sido descrito previamente, como es la elevación de la LDH, que se ha asociado con un aumento de la mortalidad en otros modelos predictivos del COVID-19⁴⁴. Con respecto al aumento de AST/GOT, el hecho de presentar un fallo hepático en el contexto de una clínica respiratoria hace que estos pacientes con mayor riesgo presenten una disfunción orgánica desde el primer momento de la atención sanitaria⁴⁵. En los pacientes con infección por COVID-19, la alterción del potasio se ha asociado a un peor pronóstico. La hiperpotasemia a la llegada de los pacientes a los SUH se ha descrito como un factor de mal pronóstico y de ingreso hospitalario⁴⁶, que se ha atribuido a un alto recambio celular incluso entre los pacientes sin fallo renal agudo⁴⁷.

La mayor limitación de nuestro estudio es el pequeño tamaño de la muestra analizada. El número de eventos por variable (número de pacientes que fallecen respecto al número de variables/factores independientes) es bajo, lo que puede comprometer la estabilidad del modelo y la generalización de los resultados. Aunque se ha implementado un procedimiento de selección de características para reducir el número de variables y cumplir con el rango mínimo recomendado de 5-10 eventos por variable, estudios recientes aconsejan que todos los factores incluidos en el diseño del estudio sean considerados variables candidatas y no solo aquellas que componen el modelo final resultado de un proceso de selección⁴⁸, lo que dificulta alcanzar tamaños muestrales óptimos. También es importante indicar que la selección de predictores en base a su significación estadística puede ser causa de sobrentrenamiento⁴⁹, siendo preferible basarse en el conocimiento previo del contexto y problema a estudio. En este sentido, se han tomado medidas para paliar el potencial efecto del sobre-entrenamiento asociado al pequeño tamaño muestral, como el empleo de un algoritmo de moderada/baja complejidad y flexibilidad (regresión logística frente a, por ejemplo, una red neuronal)⁴⁹ o la validación interna mediante remuestreo^{50,51}. Trabajos

Tabla 5. Distribución de pacientes en los grupos de riesgo (bajo vs intermedio-alto) y mortalidad hospitalaria para las escalas de gravedad bajo estudio

Escala	Total N = 186		Supervivientes N = 98		No supervivientes N = 88	
	Bajo n (%)	Intermedio-alto n (%)	Bajo n (%)	Intermedio-alto n (%)	Bajo n (%)	Intermedio-alto n (%)
PSI (≥ 91)	34 (18,3)	152 (81,7)	28 (28,6)	70 (71,4)	6 (6,8)	82 (93,2)
CURB65 (≥ 2)	34 (18,3)	152 (81,7)	27 (27,6)	71 (72,4)	7 (8,0)	81 (92,0)
SCAP (≥ 10)	40 (21,5)	146 (78,5)	33 (33,7)	65 (66,3)	7 (8,0)	81 (92,0)

PSI: Pneumonia severity index; SCAP: Severity Community Acquired Pneumonia score.

Tabla 6. Características de las distintas escalas de gravedad como predictoras de la mortalidad hospitalaria en pacientes con neumonía COVID-19

Escala	Umbral	Se (%)	Sp (%)	VPP (%)	VPN (%)	LR+	LR-
PSI (N = 186)	≥ 117 ^a	71,0 (51,8-89,0)	63,6 (45,8-82,0)	64,0 (52,3-77,3)	71,8 (58,7-86,2)	2,07 (1,40-3,42)	0,45 (0,20-0,71)
	≥ 91 ^b	93,2 (86,2-99,2)	28,5 (18,1-39,9)	53,9 (44,3-63,6)	82,5 (65,5-98,0)	1,31 (1,11-1,57)	0,25 (0,03-0,55)
CURB65 (N = 186)	≥ 3 ^a	65,9 (54,1-77,4)	69,4 (57,5-80,5)	65,9 (54,2-77,9)	69,4 (58,5-80,0)	2,24 (1,48-3,50)	0,50 (0,32-0,69)
	≥ 2 ^b	92,0 (84,8-99,1)	27,6 (17,0-39,1)	53,2 (43,6-62,3)	79,3 (61,3-97,2)	1,28 (1,08-1,52)	0,30 (0,04-0,65)
SCAP (N = 186)	≥ 16 ^a	64,4 (44,2-84,9)	65,5 (43,9-84,1)	63,0 (50,4-76,7)	67,6 (55,8-79,8)	1,99 (1,31-3,26)	0,54 (0,31-0,77)
	≥ 10 ^b	92,1 (84,7-99,1)	33,5 (22,2-44,3)	55,4 (45,4-64,7)	82,5 (67,7-97,6)	1,39 (1,16-1,70)	0,25 (0,03-0,51)
Modelo predictivo desarrollado en el presente estudio ^a (N = 180)		75,5 (62,5-87,8)	76,3 (62,8-88,5)	75,0 (62,2-86,7)	77,0 (65,3-88,0)	3,49 (2,06-6,40)	0,32 (0,17-0,50)

^aUmbral óptimo derivado de la curva ROC.

^bUmbral para la diferenciación de las categorías riesgo bajo vs riesgo intermedio-alto.

Los valores de cada métrica se proporcionan como el valor medio e IC 95% derivados de un proceso de remuestreo.

Se: sensibilidad; Sp: especificidad; VPP: valor predictivo positivo; VPN: valor predictivo negativo; LR+: razón de verosimilitud positiva; LR-: razón de verosimilitud negativa.

previos han reportado que la validación interna mediante *bootstrap* de modelos de regresión logística construidos en muestras pequeñas (eventos por variable ≤ 10) permite obtener estimaciones óptimas⁵⁰, si bien es cierto que las métricas de rendimiento predictivo pueden tender a ser optimistas cuando el número de eventos por variable es igual o menor que 5 y el ABC ROC ≤ 0,71⁴⁹. Una importante limitación adicional es el carácter retrospectivo del estudio, llevado a cabo en un único SUH. Son necesarios estudios multicéntricos y prospectivos para poder validar externamente tanto el modelo propuesto como las escalas de gravedad en pacientes ancianos con neumonía por COVID-19 dentro del ámbito de las urgencias hospitalarias. De forma similar, en este estudio se ha analizado una muestra de conveniencia tomada en un periodo de tiempo condicionado por las consecuencias de la pandemia. Aunque se incluyeron en el estudio pacientes consecutivos sin realizar ningún tipo de selección que pudiera dar lugar a un sesgo no deseado y se tomaron medidas para minimizar el efecto del sobrentrenamiento, el análisis de una muestra mayor incrementaría la capacidad de generalización del modelo y la

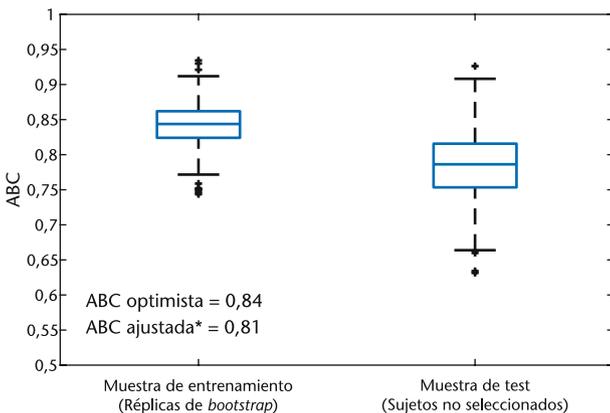


Figura 3. Distribución de los valores de ABC ROC en las muestras obtenidas por remuestreo para el cálculo y corrección del optimismo de acuerdo a la técnica de validación interna *bootstrap* 0.632.

*Corregido con *bootstrap* 0.632: (0,368 x ABC ROC aparente) + (0,632 x ABC ROC test).

potencia estadística de los resultados. También es necesario tener en cuenta el efecto de los datos perdidos. Aunque la presencia de datos perdidos en la población de estudio fue residual (0,8%), el modelo predictivo fue desarrollado con 6 instancias menos que el número de pacientes con que se evaluaron las escalas de gravedad, lo que podría influir en la ganancia de rendimiento y beneficio neto observados. Desde un punto de vista metodológico, se eligió la regresión logística como método de modelado por su probada utilidad y alto rendimiento en el contexto de los modelos predictivos clínicos⁵². Sin embargo, es conocido que la regresión logística asume una relación lineal entre las variables predictoras y la transformación *logit* de la variable dependiente, lo que no puede asegurarse particularmente en el caso de las variables clínicas continuas. Por ello, en próximos estudios sería necesario evaluar otros

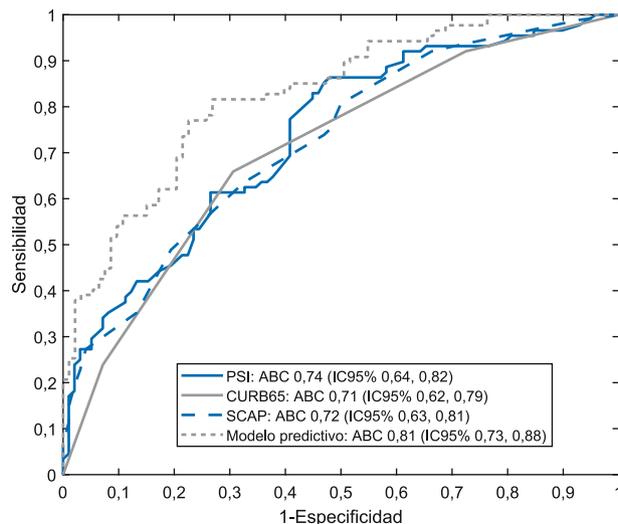


Figura 4. Curvas ROC de las escalas de gravedad y del nuevo modelo automático propuesto como predictores de la mortalidad hospitalaria en pacientes con neumonía por COVID-19 en el ingreso.

El análisis *post-hoc* arrojó diferencias estadísticamente significativas entre todos los pares de curvas aplicando la corrección de Bonferroni para múltiples comparaciones.

PSI: Pneumonia Severity Index; SCAP: Severity Community Acquired Pneumonia score.

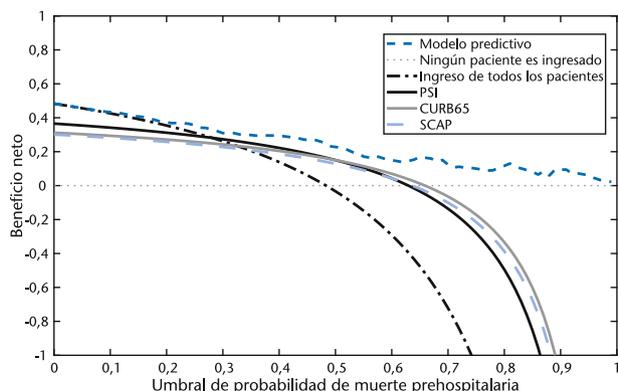


Figura 5. Curvas de decisión del modelo predictivo de mortalidad y de las escalas de gravedad.

PSI: Pneumonia Severity Index; SCAP: Severity Community Acquired Pneumonia score.

métodos de regresión no lineal, como *multivariate adaptive regression splines* (MARS) o *fraction polynomials*. Respecto a la selección automática de variables por pasos, es importante indicar que, a pesar de ser ampliamente utilizada, presenta ciertas limitaciones, especialmente cuando el tamaño de la muestra es pequeño: el subconjunto óptimo puede ser inestable, los coeficientes de la regresión pueden tomar valores extremos y el modelo seleccionado puede mostrar sobrentrenamiento⁵³. Partiendo de una muestra de estudio mayor, sería recomendable aplicar una selección por pasos hacia atrás empleando como regla de parada el Akaike Information Criterion para minimizar el sobrentrenamiento, mientras que otros métodos de selección basados en el conocimiento del contexto de estudio e independientes del algoritmo de clasificación, como los basados en la información mutua, también podrían resultar en modelos con mayor capacidad de generalización. Finalmente, es necesario destacar que únicamente se incluyeron en el estudio pacientes que precisaron ingreso hospitalario. Además, solo se evaluó la mortalidad intrahospitalaria, excluyendo la mortalidad que haya tenido lugar una vez que el paciente fue dado de alta.

En conclusión, las escalas convencionales PSI, CURB-65 y SCAP pueden ayudar a los profesionales a valorar el pronóstico de los pacientes ancianos mayores de 75 años diagnosticados de neumonía por COVID-19 en los SUH. Sin embargo, es posible superar la capacidad predictiva de estas escalas combinando mediante un modelo automático multivariante que incluya la ausencia/presencia de síntomas (tos, disnea), grado de comorbilidad (índice de Charlson) y variables analíticas (AST/GOT, potasio, urea, LDH). Estas herramientas, junto con el juicio clínico del profesional sanitario, permitiría abordar el manejo de estos pacientes de forma más eficaz desde su llegada al sistema sanitario, determinando desde el primer momento el riesgo vital para poder así programar y valorar su atención, ya que hoy en día son el grupo de población más vulnerable al COVID-19.

Conflicto de intereses: Los autores declaran no tener conflictos de in-

tereses en relación con el presente artículo.

Financiación: Este estudio ha sido parcialmente financiado por la Gerencia Regional de Salud de Castilla y León mediante el proyecto GRS COVID 09/A/20, por la Consejería de Educación de la Junta de Castilla y León al amparo de la "Convocatoria de expresiones de interés para la financiación de proyectos de investigación sobre el SARS-COV-2 y la enfermedad COVID-19 con cargo al fondo COVID-19 del Instituto de Salud Carlos III" y por el CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III, cofinanciado con fondos FEDER. Daniel Álvarez es receptor de una ayuda Ramón y Cajal (RYC2019-028566-I) del Ministerio de Ciencia e Innovación-Agencia Estatal de Investigación, cofinanciada por el Fondo Social Europeo. La investigación desarrollada por Félix del Campo fue parcialmente financiada a través de la ayuda "Intensificación de la actividad investigadora" (INT/M/15/20) de la Gerencia Regional de Salud de Castilla y León.

Responsabilidades éticas: Todos los autores han confirmado el mantenimiento de la confidencialidad y respeto de los derechos de los pacientes en el documento de responsabilidades del autor, acuerdo de publicación y cesión de derechos a EMERGENCIAS. Todos los participantes firmaron el consentimiento informado antes del inicio del estudio. El estudio fue aprobado por el Comité de Ética de la Investigación con medicamentos (CEIm) del Área de Salud Valladolid Oeste y se desarrolló de acuerdo con los principios expresados en la declaración de Helsinki.

Artículo no encargado por el Comité Editorial y con revisión externa por pares.

Bibliografía

- Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from patients with pneumonia in China, 2019. *N Engl J Med.* 2020;382:727-33.
- Coronavirus Update (Live): 41,093,940 Cases and 1,130,531 Deaths from COVID-19 Virus Pandemic - Worldometer n.d. (Consultado 21 Octubre 2020). Disponible en: <https://www.worldometers.info/coronavirus/>
- Johns Hopkins Coronavirus Resource Center (Consultado 21 Octubre 2020). Disponible en: <https://coronavirus.jhu.edu/>
- González del Castillo J, Cánora Lebrato J, Zapatero Gaviria A, Barba Martín R, Prados Roa F, Marco Martínez J. Epidemia por COVID-19 en Madrid: crónica de un reto. *Emergencias.* 2020;32:191-3.
- Emanuel EJ, Persad G, Upshur R, Thome B, Parker M, Glickman A, et al. Fair allocation of scarce medical resources in the time of Covid-19. *N Engl J Med.* 2020;382:2049-55.
- Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive Care Med.* 2020;46:846-8.
- Yan L, Zhang H-T, Gonçalves J, Xiao Y, Wang M, Guo Y, et al. An interpretable mortality prediction model for COVID-19 patients. *Nat Mach Intell.* 2020;14:283-8.
- Du RH, Liang LR, Yang CQ, Wang W, Cao TZ, Li M, et al. Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: A prospective cohort study. *Eur Respir J.* 2020;55:2000524.
- Zhang L, Yan X, Fan Q, Liu H, Liu X, Liu Z, et al. D-dimer levels on admission to predict inhospital mortality in patients with Covid-19. *J Thromb Haemost.* 2020;18:1324-9.
- Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet.* 2020;28:1054-62.
- Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, et al. Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. *J Infect.* 2020;81:e16-25.
- Rivera-Izquierdo M, Valero-Ubierna MC, R-del Amo JL, Fernández-García MA, Martínez-Díaz S, Tahery-Mahmoud A, et al. Sociodemographic, clinical and laboratory factor son admission associated with COVID-19 mortality in hospitalized patients: A retrospective observational study. *Plos One.* 2020;15:e0235107.
- Sahid Z, Kalayanamitra R, McClafferty B, Kepro D, Ramgobin D, Patel R, et al. COVID -19 and older adults: what we know. *J Am Geriatr Soc.* 2020;68:926-9.
- Freund Y, Vicent-Cassy CH, Bloom B, Riou B, Ray P, APHP Emergency Database Study Group. Association between age older than 75 years and exceeded target waiting times in the emergency Department: A multicenter cross-sectional survey in the Paris metropolitan area, France. *Ann Emerg Med.* 2013;62:449-56.
- Martín-Sánchez F, Del Toro E, Cardassay E, Valls Cabo A, Cuesta F, Vígara M, et al. Clinical presentation and outcome across age cate-

- gories among patients with COVID-19 admitted to a Spanish Emergency. *Eur Geriatr Med.* 2020;11:829-41.
- 16 Bonanad C, García Blas S, Tarazona Santa Balbina FJ, Díez Villanueva P, Ayesta A, Sanchis Forte J, et al. Coronavirus: la emergencia geriátrica de 2020. Documento conjunto de la Sección de Cardiología Geriátrica de la Sociedad Española de Cardiología y la Sociedad Española de Geriatria y Gerontología. *Rev Esp Cardiol.* 2020;73:569-76.
 - 17 Fine MJ, Auble T, Yealy D, Hanusa B, Weissfel L, Singer D, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med.* 1997;336:243-50.
 - 18 Lim WS, Van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community-acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax.* 2003;58:377-82.
 - 19 Brown S, Dean N. Defining and Predicting Severe Community-Acquired Pneumonia (SCAP). *Curr Opin Infect Dis.* 2010;234:158-64.
 - 20 España PP, Capelastegui A, Gorordo I, Esteban C, Oribe M, Ortega M, et al. Development and validation of a clinical prediction rule for severe community-acquired pneumonia. *Am J Respir Crit Care Med.* 2006;174:1249-56.
 - 21 Yandiola PPE, Capelastegui A, Quintana J, Díez R, Gorordo I, Bilbao A, et al. Prospective comparison of severity scores for predicting clinically relevant outcomes for patients hospitalized with community-acquired pneumonia. *Chest.* 2009;135:1572-9.
 - 22 Documento técnico. Manejo en urgencias del COVID-19. Ministerio de Sanidad (Consultada 15 Octubre 2020). Disponible en: https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov-China/documentos/Manejo_urgencias_pacientes_con_COVID-19.pdf
 - 23 Ochoa Gondar O, Vila Córcoles A, Rodríguez Blanco T, De Diego Cabanes C, Salsench Serrano E, Hospital Guardiola I. Utilidad de la escala de severidad modificada CRB75 en el manejo del paciente anciano con neumonía adquirida en la comunidad. *Aten Primaria.* 2013;45:208-15.
 - 24 Chen JH, Chang SS, Lui J, Chan R, Wu J, Wang W, et al. Comparison of clinical characteristics and performance of pneumonia severity score and CURB-65 among younger adults, elderly and very old subjects. *Thorax.* 2010;65:e971-7.
 - 25 Su Y, Tu GW, Ju MJ, Yu SJ, Zheng JL, Ma GG, et al. Comparison of CRB-65 and quick sepsis-related organ failure assessment for predicting the need for intensive respiratory or vasopressor support in patients with COVID-19. *J Infect.* 2020;81:647-79.
 - 26 Nguyen Y, Corre F, Honsel V, Curac S, Zarrouk V, Fantin B, et al. Applicability of the CURB-65 pneumonia severity score for outpatient treatment of COVID-19. *J Infect.* 2020;81:e96-8.
 - 27 Satici C, Demirkol MA, Altunok ES, GURSOY B, Alkan M, Karmat S, et al. Performance of Pneumonia Severity Index and CURB-65 in Predicting 30-day mortality in patients with COVID-19 *Int J Infect Dis.* 2020;98:84-9.
 - 28 Fan G, Tu C, Zhou F, Liu A, Wang Y, Song B, et al. Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study. *Eur Respir J.* 2020;56:2002113.
 - 29 Wang L, He W, Yu X, Hu D, Bao M, Liu H, et al. Coronavirus disease 2019 in elderly patients: Characteristics and prognostic factors based on 4-week follow-up. *J Infect.* 2020;80:639-45.
 - 30 Hosmer DW, Lemeshow S. *Applied Logistic Regression.* Nueva York: John Wiley & Sons; 1989.
 - 31 Witten IH, Frank E, Hall MA. *Data Mining Practical Machine Learning Tools and Techniques;* Burlington, MA, EE.UU.: Morgan Kaufmann/Elsevier; 2011.
 - 32 Gutiérrez-Tobal GC, Alonso-Álvarez ML, Álvarez D, del Campo F, Terán-Santos J, Hornero R. Diagnosis of pediatric obstructive sleep apnea: Preliminary findings using automatic analysis of airflow and oximetry recordings obtained at patients' home. *Biomed Signal Process Control.* 2015;18:401-7.
 - 33 Álvarez D, Alonso-Álvarez ML, Gutiérrez-Tobal GC, Crespo A, Kheirandish-Gozal L, Hornero R, et al. Automated screening of children with obstructive sleep apnea using nocturnal oximetry: An alternative to respiratory polygraphy in unattended settings. *J Clin Sleep Med.* 2017;13:693-702.
 - 34 Efron B, Tibshirani RJ. *An Introduction to the Bootstrap.* Nueva York, NY, EE.UU.: Chapman and Hall; 1994.
 - 35 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* 2006;26:565-74.
 - 36 Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. *Ann Intern Med.* 2015;162:55-63.
 - 37 Carmel S, Anson O, Levin M. Emergency department utilization: A comparative analysis of older— adults, old and old— old patients. *Aging Clin Exp Res.* 1990;2:387-93.
 - 38 Segrelles-Calvo G, de Granda-Orive JJ, López-Padilla D, Zamora García E. Limitación terapéutica en pacientes ancianos: reflexiones a propósito del COVID-19. *Arch Bronconeumol.* 2020;56:677-9.
 - 39 Liu K, Chen Y, Lin R, Han K. Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. *J Infect.* 2020;80:e14-8.
 - 40 Chen JH, Chang SS, Lui J, Chan R, Wu J, Wang W, et al. Comparison of clinical characteristics and performance of pneumonia severity score and CURB-65 among younger adults, elderly and very old subjects. *Thorax.* 2010;65:971-7.
 - 41 Loke Y, Shing Kwok C, Niruban A, Myint P. Value of severity scales in predicting mortality from community-acquired pneumonia: systematic review and meta-analysis. *Thorax.* 2010;65:884-90.
 - 42 Ewig S, de Roux A, Bauer T, García E, Mensa J, Niederman M, et al. Validation of predictive rules and indices of severity for community acquired pneumonia. *Thorax.* 2004;59:421-7.
 - 43 Ji D, Zhang D, Xu J, Chen Z, Yang T, Zhao P, et al. Prediction for progression risk in patients with COVID-19 pneumonia: the CALL score. *Clin Infect Dis.* 2020;71:1393-9.
 - 44 Liang W, Liang H, Ou L, Chen b, Chen A, Li C, et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *JAMA Intern Med.* 2020;180:1081-9.
 - 45 Pei G, Zhang Z, Peng J, Liu L, Zhang C, Yu C, et al. Renal Involvement and early prognosis in patients with COVID-19 pneumonia. *J Am Soc Nephrol.* 2020;31:1157-65.
 - 46 Gupta A, Madhavan MV, Sehgal K, Nair N, Mahajan S, Sehrawat TS, et al. Extrapulmonary manifestations of COVID-19. *Nat Med.* 2020;26:1017-32.
 - 47 Singer AJ, Thode HC, Peacock WF. Rapid correction of hyperkalemia is associated with reduced mortality in ED patients. *Am J Emerg Med.* 2020;38:2361-4.
 - 48 Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ.* 2020;368:m441.
 - 49 Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. *BMC Medicine.* 2019;17:230.
 - 50 Steyerberg EW, Harrell Jr FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol.* 2001;54:774-81.
 - 51 Smith GCS, Seaman SR, Wood AM, Royston P, White IR. Correcting for optimistic prediction in small data sets. *Am J Epidemiol.* 2014;180:318-24.
 - 52 Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of ma-

- chine learning over logistic regression for clinical prediction models. *J Clin Epidemiol.* 2019;110:12-22.
- 53 Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J.* 2014;35:1925-31.