

Validación externa de los modelos de predicción de mortalidad 4C y NEWS2 en COVID-19 en tres centros de Medellín, Colombia: Evaluando el impacto de la vacunación en el comportamiento a lo largo del tiempo.

Autores:

Paola Andrea Gallego Aristizabal^{1*}, Tania Paola Lujan Chavarría^{1*}, Sara Isabel Vergara Hernández², Federico Rincón Acosta², María Paula Sánchez Carmona², Paula Andrea Salazar Ospina², Carlos José Atencia Florez¹, Carlos Mario Barros Liñán^{1,4}, Fabián Jaimes¹.

1. Departamento de medicina interna, facultad de medicina, Universidad de Antioquia, Medellín, Colombia.
2. Facultad de medicina, Universidad de Antioquia, Medellín, Colombia.
3. Hospital universitario San Vicente Fundación, Medellín, Colombia.
4. Hospital Alma Mater, Medellín Medellín, Colombia.

**Los autores marcados con el asterisco contribuyeron por igual en la producción de esta investigación y la escritura de este manuscrito.*

Correspondencia dirigida a:

Paola Andrea Gallego Aristizabal: paolagallego383@gmail.com

ORCID:

Paola Andrea Gallego Aristizabal: <https://orcid.org/0000-0002-7488-0077>

Tania Paola Lujan Chavarría: <https://orcid.org/0000-0002-9551-4551>

Sara Isabel Vergara Hernández: <https://orcid.org/0009-0005-3265-269X>

Federico Rincón Acosta: <https://orcid.org/0009-0002-3462-3049>

María Paula Sánchez Carmona: <https://orcid.org/my-orcid?orcid=0009-0001-6666->

Paula Andrea Salazar Ospina: <https://orcid.org/0009-0006-6297-649X>

Carlos José Atencia Flórez: <https://orcid.org/0000-0002-4531-229X>

Carlos Mario Barros Liñán: <https://orcid.org/0000-0003-3297-2755>

Fabián Jaimes Barragán: <https://orcid.org/0000-0002-7315-5367>

Resumen:

Objetivos: validación externa de los puntajes 4C y NEWS2 para la predicción de mortalidad intrahospitalaria en pacientes con COVID-19 y evaluación del rendimiento operativo en dos períodos de tiempo definidos por el inicio del programa de vacunación en Colombia.

Diseño del estudio: cohorte retrospectiva.

Escenario: tres hospitales de alta complejidad de la ciudad de Medellín, Colombia, durante junio de 2020 a abril de 2022 (primer período antes del 1 de mayo de 2021, segundo período después del 1 de mayo de 2021).

Participantes: pacientes mayores o iguales de 18 años con diagnóstico confirmado de COVID-19 que ingresaron por el servicio de urgencias.

Desenlace: Mortalidad intrahospitalaria.

Resultados: en 1789 pacientes la mortalidad hospitalaria fue de 42,1% y la necesidad de ventilación mecánica invasiva de 30,9%. El puntaje de mortalidad 4C tuvo un área bajo la curva (AUC) de 0.75 (IC 95% 0.73 a 0.78) y un estadístico de Hosmer-Lemeshow con $p = 0,13$. El puntaje NEWS2 tuvo una AUC de 0.68 (0.66 - 0.71) y un estadístico de Hosmer-Lemeshow con $p = 0,23$. Para el puntaje 4C el

AUC del primer y segundo período fue de 0.77 (IC 95% 0.74 a 0.80) y 0.75 (IC 95% 0.71 a 0.78), respectivamente; mientras para el NEWS2 fue de 0.68 (IC 95% 0.65 a 0.71) y 0.69 (IC 95% 0.64 a 0.73), respectivamente. Los gráficos de calibración para ambos puntajes demostraron adecuada capacidad para predecir de manera exacta la mortalidad, aunque en el segundo período presentaron un desempeño menor.

Conclusiones: La escala de mortalidad 4C demostró ser una herramienta adecuada para predecir mortalidad intrahospitalaria en pacientes con COVID-19 en esta población latinoamericana. El rendimiento operativo durante los dos períodos de tiempo permaneció, lo cual evidencia su utilidad incluso con los cambios a través del tiempo de la infección por SARS-CoV2. Su uso puede contribuir a mejor gestión del riesgo, elecciones en el escenario de cuidado intrahospitalario y toma de decisiones basadas en la evidencia.

Introducción

La infección por SARS-CoV2 se ha expandido a todos los continentes, siendo declarada una pandemia por la Organización Mundial de la Salud (OMS) el 11 de marzo de 2020, amenazando la salud pública con miles de muertes diarias en todo el mundo, lo cual ha representado un desafío sustancial para los gobiernos, individuos y la sociedad(1,2). A tres años desde su inicio, fue declarada el fin de la emergencia sanitaria, sin embargo, es importante reconocer que la erradicación total del virus es poco probable: SARS-CoV2 mantiene una virulencia significativa y continúa circulando, lo que podría llevar a casos diarios y períodos de brotes (3).

La letalidad de la enfermedad por coronavirus 2019 (COVID-19, por sus siglas en inglés) en Colombia se ha estimado aproximadamente de 2.1%(4). Sin embargo, la presentación clínica y la evolución de la enfermedad en los pacientes son altamente variables, debido a la heterogeneidad en la respuesta inmunológica, la respuesta individual determinada genéticamente, las comorbilidades y la variante de SARS-CoV-2 implicada (5,6); lo que dificulta la clasificación y la determinación del riesgo de resultados adversos por parte de los médicos. La predicción del riesgo de muerte es esencial para una gestión eficiente como determinar la oportunidad de atención y seleccionar a aquellos que se beneficiarían de intervenciones que cambien el curso de la enfermedad (7). Durante la pandemia se han desarrollado numerosos modelos de predicción con fines pronósticos, con amplia heterogeneidad en el rendimiento en las cohortes de validación externa (8).

El puntaje de mortalidad del Consorcio de Caracterización Clínica del Coronavirus (4C) derivado durante la fase inicial de la pandemia (9), así como el puntaje NEWS2 (National Early Warning Score 2) (10–12) ambos desarrollados en el Reino Unido, han tenido una amplia aceptación y validación en el mundo. Ninguno de los dos se ha validado en población colombiana. Sumado a ello, las estrategias para la prevención y el manejo de COVID-19 han cambiado con el desarrollo de vacunas y las nuevas opciones de tratamiento (13), teniendo un impacto en la mortalidad, por lo que es importante investigar los cambios en el rendimiento de los puntajes a través del tiempo (7). La investigación y publicación de nuevos hallazgos sobre COVID-19 son fundamentales para apoyar la toma de decisiones basadas en evidencia que debe incluir a todas las poblaciones.

En este contexto, el objetivo de este estudio fue validar localmente los puntajes 4C y NEWS2 para la predicción de mortalidad intrahospitalaria en pacientes con COVID-19 confirmado por prueba de reacción de cadena polimerasa (PCR-RT) o antígeno para SARS-CoV2 admitidos por el servicio de urgencias; y comparar su desempeño en dos períodos de tiempo definidos por el inicio del programa de vacunación en Colombia.

Métodos

Diseño del estudio, escenario y participantes: estudio de validación retrospectivo multicéntrico en tres hospitales de alta complejidad en la ciudad de Medellín, Colombia, entre junio de 2020 hasta abril de 2022. Los criterios de inclusión fueron pacientes mayores o iguales de 18 años con

diagnóstico confirmado por prueba de antígeno para SARS-CoV2 o PCR-RT que ingresaron por el servicio de urgencias. Se excluyeron a los pacientes fallecidos en las primeras 24 horas de admisión o que fueran remitidos de otra institución en donde permanecieron por más de 24 horas.

Variables y fuentes de información: los datos para calcular los puntajes 4C y NEWS2 se obtuvieron de la historia clínica de ingreso con los primeros signos vitales registrados y los laboratorios tomados dentro de las primeras 24 horas. Cada modelo cuenta con diferentes variables estratificadas a las que se les asigna un puntaje y se determina el grupo de riesgo (**Tablas S1 y S2 del material suplementario**).

Desenlace primario: mortalidad intrahospitalaria con doble verificación de la historia clínica de egreso y del registro proporcionado por los hospitales.

Tamaño de la muestra: se determinó siguiendo las recomendaciones de presentación transparente de un modelo de predicción multivariable para pronóstico o diagnóstico individual (TRIPOD, por sus siglas en inglés) de 100 desenlaces por variable (14).

Análisis estadísticos y datos perdidos: Las variables cuantitativas con distribución normal se presentaron como medias con desviación estándar y las variables con distribución no normal se presentaron como medianas con rangos intercuartílicos. Para las variables cualitativas se usó la frecuencia absoluta y relativa expresada en porcentaje. Para cada puntaje se calcularon las métricas de desempeño sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo, así como las razones de probabilidad (LR), por medio del teorema de

Bayes, asumiendo como “prueba de referencia” el desenlace de mortalidad intrahospitalaria.

Se determinó la discriminación de los puntajes a través del área bajo la curva de características operativas del receptor (AUC) y la calibración a través de gráficos de calibración y pruebas de bondad de ajuste de Hosmer-Lemeshow (H-L). Se efectuó el análisis de dos periodos de tiempo, antes del 1 de mayo de 2021 y después del 1 de mayo de 2021, determinados por el tiempo de inicio de la vacunación en Colombia. Se hizo un análisis de casos completos y también con imputación de datos simple, asumiendo las variables faltantes como normales. Para el puntaje 4C en los pacientes que se registró la saturación con oxígeno suplementario al ingreso de urgencias, se asumió como saturación menor del 92% y para el puntaje NEWS2 se calculó el puntaje con la saturación de oxígeno registrada independiente del antecedente de enfermedad pulmonar obstructiva crónica. El estudio se informa de acuerdo con la guía (TRIPOD) (14).

El análisis estadístico se realizó en software STATA (StataCorp. 2019. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC).

Aspectos éticos: La investigación se realizó de acuerdo con la declaración de Helsinki de 2013, el protocolo de la investigación fue evaluado y aprobada por el comité de ética de las 3 instituciones participantes antes del inicio de la recolección de datos.

Resultados:

Participantes: se revisaron 3013 historias de las cuales 1789 cumplieron criterios de inclusión (**Figura 1**). La mortalidad hospitalaria fue de 42.1% (n = 754). La edad

media fue 64.9 años (desviación estándar (DE): 16.1), el 53,6% de los participantes fueron de sexo masculino y el 36,16% tenía al menos una comorbilidad. En la **Tabla 1** se muestran las características clínicas de los pacientes según su estado vital al alta. Los porcentajes de datos faltantes de cada variable de los modelos de predicción están disponibles en la **tabla S3 del suplemento**.

Desempeño del modelo: el AUC del puntaje 4C en la cohorte fue de 0.75 (IC 95% 0.73 a 0.78) (**Figura 2A**) y el de NEWS2 0.68 (IC 95% 0.66 a 0.71) (**Figura 2B**). El gráfico de calibración demostró una adecuada capacidad del puntaje 4C para predecir de manera exacta la mortalidad, aunque fue menor en los pacientes de muy alto y muy bajo riesgo, donde el modelo subestimó discretamente el riesgo de morir (**Figura 3A**). La calibración medida por la prueba de H-L tuvo un valor de p de 0,13. El puntaje NEWS2 presentó menor exactitud para predecir mortalidad en los pacientes de bajo riesgo, donde sobreestimó la ocurrencia de muerte (**Figura 3B**).

Impacto de cambios en el tiempo en la discriminación y calibración de las escalas

El AUC del puntaje 4C en el primer y segundo período propuestos fue de 0.77 (IC 95% 0.74 a 0.8) y 0.75 (IC 95% 0.71 a 0.78) respectivamente. Por su parte el AUC del puntaje NEWS2 en el primer período fue de 0.68 (IC 95% 0.65 a 0.71) y en el segundo período de 0.69 (IC 95% 0.64 a 0.73). Para el primer período el gráfico de calibración del puntaje 4C demostró adecuada capacidad para predecir de manera exacta la mortalidad, aunque en el segundo período la calibración presentó un desempeño menor, la prueba de H-L tuvo un valor de p de 0,51 y 0.13 para el primer

y segundo período respectivamente. Para el puntaje NEWS2 la calibración a través del gráfico de calibración se demostró sobreestimación del riesgo en el grupo de bajo y alto riesgo principalmente en el segundo periodo, medida por la prueba de H-L tuvo un valor de p de 0,81 para el primer período y 0.33 para el segundo período **(Tabla 2 y Figura S1 del suplementario)**.

Otras métricas de desempeño:

El puntaje de mortalidad 4C demostró un buen desempeño en métricas clínicamente relevantes en su rango de valores, con un LR (-) de 0.063 (IC 95% 0.01 a 0.26) para un puntaje menor o igual a 4, con una sensibilidad de 99.7% (IC 95% 99 a 100); mientras en valores mayores de 15 puntos presentó una especificidad de 92.5% (IC 95% 90.7 a 94) y un LR (+) de 3.4 (IC 95% 2.7 a 4.4) **(Tabla 3)**. En el puntaje NEWS2 el punto de corte mayor o igual a 7 puntos presentó una especificidad de 60.4 (IC 95% 57.3 a 63.4) y un LR (+) de 1.53 (IC 95% 1.3 a 1.6) y para un puntaje mayor o igual a 5 se encontró una sensibilidad de 83.6% (IC 95% 80.7 - 86.1) y un LR (-) 0.53 (IC 95% 0.4 a 0.64) **tabla 4**.

La mortalidad total fue mayor en este estudio en el primer periodo (42.15 %) en comparación con el segundo periodo (31.3%) y la mortalidad observada a través de los grupos de riesgo del puntaje 4C fue superior a la reportada en la cohorte de desarrollo **(Tabla 5)**. De igual manera, en el puntaje NEWS2 la mortalidad total fue mayor a la reportada en la cohorte de desarrollo **(Tabla 6)**. Estratificada por grupos de riesgo se observaron diferencias principalmente en las categorías de riesgo bajo (27.93 %) y alto (52.66 %). La asociación entre las variables predictoras de cada

puntaje y la mortalidad se encuentran en la **tabla S4 y S5 del material suplementario**.

Análisis de sensibilidad: en el análisis de sensibilidad con datos imputados el rendimiento operativo de los puntajes permaneció (**tabla S6, figura S2 y figura S3 del material suplementario**).

Discusión

Este estudio es la primera validación colombiana de los puntajes de predicción de mortalidad intrahospitalaria 4C y NEWS2 en una cohorte multicéntrica de la ciudad de Medellín, Colombia. Se evaluó el rendimiento operativo de los puntajes a lo largo del tiempo considerando los cambios ocurridos durante la pandemia.

El puntaje de mortalidad 4C presentó una discriminación y calibración adecuadas, aunque con discreta subestimación en los extremos de alto y bajo riesgo de muerte. Estos resultados son similares a los encontrados en el estudio de desarrollo realizado en Reino Unido en el año 2020, donde se encontró en la cohorte de validación interna un AUC 0.76 (IC 95% 0.760 a 0.773), aunque se encontró una calibración excelente en todo el rango de riesgo. La mortalidad a través de los grupos de riesgo fue mayor en este estudio que en la cohorte de desarrollo, probablemente explicado por las características propias de la población.

El puntaje 4C también demostró un buen desempeño en métricas clínicamente relevantes, el punto de corte de 4 puntos presenta una alta sensibilidad, muy similar

a la cohorte de desarrollo (99.2%), y el punto de corte ≥ 15 puntos presenta una alta especificidad (92.5%), incluso mayor al de la cohorte de desarrollo (89.8%), lo que predice a los pacientes con muy alta probabilidad de morir intrahospitalariamente, que podría ser útil en la elección de escenarios de cuidado intrahospitalario como vigilancia en unidades de alta dependencia.

En otras cohortes de validación externa se han encontrado hallazgos similares para el puntaje de mortalidad 4C, un estudio también de Reino Unido realizado durante el segundo brote de la pandemia (agosto 2020 y febrero 2021) encontró una AUC 0.76, con una buena calibración. En los cuatro grupos de riesgo, la mortalidad correspondiente fue muy similar a la reportada en la cohorte de desarrollo (15). En París se reportó un AUC de 0.785 (IC 95% 0.775–0.795)(16). En EE. UU. se realizó validación externa en el conjunto de datos RECOVER en donde tuvo una discriminación comparable con AUC 0.786 (IC 95% 0.773 - 0.799) y una adecuada calibración (17).

El puntaje NEWS2 presentó una discriminación aceptable pero menor que la del 4C y en la calibración se evidencia una sobrestimación de la probabilidad de muerte en la categoría de riesgo bajo y alto. Hallazgos similares son reportados en un estudio de validación externa en paciente con COVID-19 en Inglaterra al inicio de la pandemia, encontrando que el modelo subestima el riesgo de mortalidad en todas las categorías de riesgo y presenta una pobre discriminación y un AUC de 0.65 (IC 95% 0.61 - 0.68) para predecir mortalidad intrahospitalaria (18). Otra cohorte de Reino Unido encontró un desempeño similar con AUC 0.65 (IC 95% 0.61 - 0.68)(11).

En este estudio, el punto de corte de 5 puntos presentó una sensibilidad del 83% y un valor predictivo negativo del 79% (19), menor a lo reportado en un estudio de Noruega que encontró una sensibilidad del 86.7 (IC 95% 59.5–98.3) y un valor negativo predictivo del 94.7 (IC 95% 83.0–98.5).

En este estudio se encontró que el puntaje de mortalidad 4C tiene mayor poder discriminatorio y mejor calibración que el puntaje NEWS2 lo que también coincide con la cohorte de desarrollo del puntaje 4C en donde se encontró mejor desempeño sobre otros 15 puntajes incluido el puntaje NEWS (IC 95% 0.774 vs 0.654)(9). Asimismo, en Bélgica se obtuvo resultados similares, donde se describe una AUC de 0.71 (IC 95% 0.58-0.83) para NEWS2 y una AUC de 0.80 (IC 95% 0.69-0.91) para el puntaje 4C en la predicción de la mortalidad a los 30 días (20).

La implementación de los puntajes de predicción de riesgo de mortalidad 4C y NEWS2 continuaron siendo válidos a pesar de los cambios en la pandemia. Las características operativas del puntaje de mortalidad 4C y NEWS2 permanecieron en los periodos de tiempo evaluados, determinados principalmente a partir del inicio de la vacunación en Colombia, el linaje de mayor circulación B.1 a partir del segundo periodo y la publicación del estudio RECOVERY que determinó la efectividad de los esteroides el 25 de febrero de 2021. Hallazgos similares se reportaron para el puntaje 4C en un estudio realizado en Canadá, en donde se valoró el desempeño del puntaje por brotes epidémicos (3 periodos hasta junio de 2021); los valores AUC en los brotes fueron de 0.81 (IC 95% 0.76 - 0.86), 0.74 (IC 95% 0.69 - 0.80) y 0.76

(IC 95%0.69 - 0.83), respectivamente(21), sin embargo, pocos estudios han valorado el desempeño posterior al inicio de la vacunación.

Fortalezas y limitaciones:

Las fortalezas del estudio incluyen el escenario clínico, sin datos previamente validados en esta población colombiana, lo que es importante en enfermedades de interés público mundial como COVID-19 ya que disminuye la brecha de información e investigación entre países de bajos y altos ingresos. Adicionalmente se logró un número significativo de desenlaces por variable y se consideró el impacto potencial de los factores cambiantes en la pandemia que pueden modificar el desempeño de los modelos de predicción, teniendo en cuenta que gran parte de los estudios de desarrollo y validación de los modelos se llevaron a cabo al inicio de la pandemia.

Dentro de las limitaciones se encuentra el número de datos faltantes en variables que componen los modelos de predicción, principalmente Glasgow y temperatura; para abordar esta limitación se realizó imputación simple de datos, asumiendo las variables faltantes como normales y análisis de sensibilidad en donde el desempeño operativo de las pruebas fue similar. Otra limitación es el carácter retrospectivo del estudio que limita el control del registro de las variables predictoras y adicionalmente es probable que el desempeño de los puntajes cambie en escenarios de menor mortalidad.

Conclusiones e implicaciones:

Se validó en este estudio que la herramienta de predicción de mortalidad 4C es la más adecuada, en comparación con el NEWS2, para predecir mortalidad

intrahospitalaria en pacientes con COVID-19 en Medellín, Colombia. Además, se encontró que las características operativas del modelo pronóstico permanecen estables a pesar de los cambios en el tiempo de la pandemia por SARS CoV2, como la vacunación y las intervenciones terapéuticas. La implementación en la práctica clínica puede contribuir a una mejor gestión del riesgo, elecciones en el escenario de cuidado y toma de decisiones basadas en la evidencia.

Queda por explorar el efecto en la toma de decisiones y los resultados de los pacientes, lo cual debe ser evaluado en futuras investigaciones. Además, se debe explorar la contribución de los puntajes para definir a los mejores candidatos a intervenciones médicas específicas, como antivirales o nuevas terapias, lo que permitiría optimizar la eficiencia en la asignación de los recursos disponibles.

Bibliografía

1. Jin Y, Yang H, Ji W, Wu W, Chen S, Zhang W, et al. Virology, epidemiology, pathogenesis, and control of covid-19. *Viruses*. 2020;12(4).
2. Shi Y, Wang G, Cai X peng, Deng J wen, Zheng L, Zhu H hong, et al. An overview of COVID-19. *J Zhejiang Univ Sci B*. 2020 May 1;21(5):343–60.
3. Callaway E. Beyond Omicron- what’s next for COVID’s viral evolution. *Nature*. 2021;600.
4. Coronavirus Colombia [Internet]. Instituto nacional de salud. 2023 [cited 2023 Apr 14]. Available from: <https://www.ins.gov.co/Noticias/paginas/coronavirus.aspx>
5. Pereira NL, Ahmad F, Byku M, Cummins NW, Morris AA, Owens A, et al. COVID-19: Understanding Inter-Individual Variability and Implications for Precision Medicine. Vol. 96, *Mayo Clinic Proceedings*. Elsevier Ltd; 2021. p. 446–63.
6. Flisiak R, Rzymiski P, Zarębska-Michaluk D, Ciechanowski P, Dobrowolska K, Rogalska M, et al. Variability in the Clinical Course of COVID-19 in a Retrospective Analysis of a Large Real-World Database. *Viruses*. 2023 Jan 3;15(1).
7. Jones A, Pitre T, Junek M, Kapralik J, Patel R, Feng E, et al. External validation of the 4C mortality score among COVID-19 patients admitted to hospital in Ontario, Canada: a retrospective study. *Sci Rep*. 2021 Dec 1;11(1).
8. Harish keerthi. Validation of parsimonious prognostic models for patients infected with COVID-19. *BMJ Health an care informatics*. 2020;

9. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Development and validation of the 4C Mortality Score. *The BMJ*. 2020 Sep 9;370.
10. 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. *The BMJ*. 2022;
11. Scott LJ, Tavaré A, Hill EM, Jordan L, Juniper M, Srivastava S, et al. Prognostic value of National Early Warning Scores (NEWS2) and component physiology in hospitalised patients with COVID-19: A multicentre study. *Emergency Medicine Journal*. 2022;
12. Royal College of Physicians. National Early Warning Score (NEWS) 2 Standardising the assessment of acute-illness severity in the NHS [Internet]. 2017. Available from: www.rcplondon.ac.uk
13. Baba H, Ikumi S, Aoyama S, Ishikawa T, Asai Y, Matsunaga N, et al. Statistical Analysis of Mortality Rates of Coronavirus Disease 2019 (COVID-19) Patients in Japan Across the 4C Mortality Score Risk Groups, Age Groups, and Epidemiological Waves: A Report From the Nationwide COVID-19 Cohort. *Open Forum Infect Dis*. 2023 Jan 4;10(1).
14. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med*. 2015 Jan 6;162(1):W1–73.
15. Knight SR, Gupta RK, Ho A, Pius R, Buchan I, Carson G, et al. Prospective validation of the 4C prognostic models for adults hospitalised with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol. *Thorax*. 2022 Jun 1;77(6):606–15.
16. Lombardi Y, Azoyan L, Szychowiak P, Bellamine A, Lemaitre G, Bernaux M, et al. External validation of prognostic scores for COVID-19: a multicenter cohort study of patients hospitalized in Greater Paris University Hospitals. *Intensive Care Med*. 2021 Dec 1;47(12):1426–39.
17. Gordon AJ, Govindarajan P, Bennett CL, Matheson L, Kohn MA, Camargo C, et al. External validation of the 4C Mortality Score for hospitalised patients with COVID-19 in the RECOVER network. *BMJ Open*. 2022 Apr 21;12(4).
18. Bradley P, Frost F, Tharmaratnam K, Wootton DG. Utility of established prognostic scores in COVID-19 hospital admissions: Multicentre prospective evaluation of CURB-65, NEWS2 and qSOFA. *BMJ Open Respir Res*. 2020 Dec 7;7(1).
19. Bradley P, Frost F, Tharmaratnam K, Wootton DG. Utility of established prognostic scores in COVID-19 hospital admissions: Multicentre prospective evaluation of CURB-65, NEWS2 and qSOFA. *BMJ Open Respir Res*. 2020 Dec 7;7(1).
20. Yildiz H, Castanares-Zapatero D, Hannedouche C, Vandermeersch D, Pothen L, Yombi JC. Prospective validation and comparison of COVID-GRAM, NEWS2, 4C mortality score, CURB-65 for the prediction of critical illness in COVID-19 patients. Vol. 53, *Infectious Diseases*. Taylor and Francis Ltd.; 2021. p. 640–2.
21. Jones A, Pitre T, Junek M, Kapralik J, Patel R, Feng E, et al. External validation of the 4C mortality score among COVID-19 patients admitted to hospital in Ontario, Canada: a retrospective study. *Sci Rep*. 2021 Dec 1;11(1).
22. Corfield AR, Lees F, Zealley I, Houston G, Dickie S, Ward K, et al. Utility of a single early warning score in patients with sepsis in the emergency department. *Emergency Medicine Journal*. 2014;31(6):482–7.

Tabla 1. Características clínicas según estado vital al alta

Variable	Total, n: 1789	Vivos n: 1035	Muertos n: 754
Edad (media ± DE)	64,9 ± 16,1	61,03 ± 16,4	70,26 ± 13,9
Sexo masculino (n, %)	959 (53,6)	540 (53,4)	419 (53,9)
Malignidad* (n, %)	150 (8,4)	78 (7,7)	72 (9,3)
VIH* (n, %)	16 (0,9)	13 (1,3)	3 (0,4)
Diabetes mellitus* (n, %)	462 (25,9)	252 (24,8)	210 (27,0)
Enfermedades del tejido conectivo* (n,%)	91 (5,1)	47 (4,6)	44 (5,7)
Enfermedad neurológica* (n, %)	93 (5,2)	53 (5,2)	40 (5,2)
Demencia* (n, %)	107 (6,0)	50 (4,9)	57 (7,3)
Enfermedad hepática* (n, %)	34 (1,9)	16 (1,6)	16 (2,1)
Enfermedad renal* (n, %)	123 (6,9)	61 (6,0)	62 (8,0)
Enfermedad pulmonar* (n, %)	223 (12,5)	110 (10,8)	113 (14,5)
Enfermedad cardíaca* (n, %)	193 (10,8)	96 (9,5)	97 (12,5)
Obesidad** (n, %)	505 (28,2)	266 (26,2)	239 (30,8)
IMC (kg/m ²) (media ± DE)	28,7 ± 12,8	28,5 ± 10,5	28,9 ± 14,6
Temperatura C° (media ± DE)	36,9 ± 2,13	36,9 ± 0,9	36,9 ± 0,9
Saturación % (media ± DE)	84 ± 12,9	86,8 ± 10	81,1 ± 15,1
PAD (mmHg) (media ± DE)	92 ± 13,6	77 ± 12	73 ± 14,3

PAS (mmHg) (media ± DE)	127 ± 23,7	127 ± 22,9	127 ± 24,8
FR (/min) (media ± DE)	25,1 ± 8,2	23,7 ± 7,2	27
FC (/min) (media ± DE)	95,3 ± 18,6	95,5 ± 18,5	94,9 ± 18,9
Glasgow 15 (n (%))	1503 (84,0)	890 (86,0)	613 (81,3)
Glasgow <15 (n (%))	160 (8,9)	32 (3,1)	64 (8,5)
BUN (mg/dL) (media ± DE)	25,5 ± 18,6	25,5 ± 18,6	30,6 ± 20,8
PCR (mg/dL) (media ± DE)	12,4 ± 13,7	12,9 ± 14,1	11,7 ± 12,9
PO2 (mmHg) (media ± DE)	73,1 ± 26,9	73,1 ± 26,9	66,3 ± 23,6
PCO2 (mmHg) (media ± DE)	32,8 ± 10,1	32,8 ± 10,1	32,7 ± 10,6
Ferritina ng/mL (IQR)	1313 (752 – 2199)	1100 (154-1984)	1526 (238-2694)
Troponina pg/ml (RIQ)	302 (2,5-396)	151 (2,5 -140)	485 (4,77-752)
Dímero D ng/ml (RIQ)	2247 (0,65 – 2446)	2631 (0,6 – 2192)	2102 (0,69 - 679)
LDH U/L (RIQ)	440 (289 – 703)	363 (202-540)	537 (265-840)
VMI (%)	553 (30,9)	122 (11,7)	430 (57)
TVP (n (%))	41 (2,3)	40 (3,9)	17 (2,2)
TEP (n (%))	69 (3,8)	24 (2,3)	29 (3,8)
Bacteriemia (%)	572 (31,9)	257 (41,7)	315 (2,3)

VIH virus de inmunodeficiencia humana, IMC índice de masa corporal, PAD presión arterial diastólica, PAS presión arterial sistólica, FR frecuencia respiratoria, FC frecuencia cardiaca, BUN nitrógeno ureico en Sangre, PCR proteína C reactiva, PO2 presión arterial de oxígeno, PCO2 presión parcial de dióxido de carbono, LDH lactato deshidrogenasa, VMI ventilación mecánica invasiva, TEP tromboembolismo pulmonar, TVP trombosis venosa profunda.

*Comorbilidades definidas por índice de Charlson.

**Obesidad clínicamente definida, DE desviación estándar, RIQ rango intercuartílico.

Tabla 2. Área bajo la curva (AUC) y prueba de Hosmer-Lemeshow para mortalidad en pacientes con COVID-19 por período de estudio.

Puntaje	Total	Primer periodo*	Segundo periodo*
AUC (IC 95%)			
Mortalidad 4C	0.75 (0.73 - 0.78)	0.77 (0.74 - 0.80)	0.75 (0.71 - 0.78)
NEWS2	0.68 (0.66 - 0.71)	0.68 (0.65 - 0.71)	0.69 (0.64 - 0.73)
Prueba de Hosmer-Lemeshow (estadístico <i>p</i>)			
Mortalidad 4C	0,13	0,51	0,13
NEWS2	0,68	0,82	0,33

*Primer periodo: antes del 1 de mayo de 2021, segundo periodo: después del 1 de mayo de 2021.

Tabla 3. Métricas de rendimiento de la puntuación de mortalidad 4C para confirmar la mortalidad en diferentes valores de corte.

Punto de corte	No (%) de pacientes	Sensibilidad (%) (IC95%)	Especificidad (%) (IC95%)	VPP (%) (IC95%)	VPN (%) (IC95%)	LR (+) (IC95%)	LR (-) (IC95%)
≥15	275 (15.3)	26.1 (23- 29.4)	92.5 (90.7 - 94)	71.6 (65.9 -76.9)	63.2 (60.7- 65.6)	3.4 (2.7- 4.4)	0.79 (0.7 - 0.8)
≥9	1205 (67)	86.9 (84.2 - 89.2)	46.9 (43.8 - 50)	54.4 (51.5 -57.2)	83 (79.8 - 86)	1.63 (1.5-1.7)	0.28 (0.2- 0.3)
≥4	1744 (97)	99.7 (99 - 100)	4.15 (3.02 -5.56)	43.1 (40.8 - 45.5)	95.6 (84 - 99)	1.04 (1.0-1.0)	0.063 (0.01-0.26)

LR(+): Likelihood ratio positivo, LR(-): Likelihood ratio negativo, VPP: Valor predictivo positivo, VPN: Valor predictivo negativo No: Número.

Tabla 4. Métricas de rendimiento de la puntuación de NEWS2 para confirmar la mortalidad en diferentes valores de corte.

Punto de corte	No (%) de pacientes	Sensibilidad (%) (IC95%)	Especificidad (%) (IC95%)	VPP (%) (IC95%)	VPN (%) (IC95%)	LR (+) (IC95%)	LR (-) (IC95%)
≥7	886 (48,41)	60.5 (56.9 - 64)	60.4 (57.3 - 63.4)	52.7 (49.3 - 56)	67.7 (64.6 - 70.7)	1.53 (1.3 - 1.6)	0.65 (0.5 - 0.72)
≥5	1345 (75)	83.6 (80.7 - 86.1)	30.9 (28.1 - 33.8)	46.8 (44.1 - 49.5)	72.1 (67.6 - 76.2)	1.21 (1.1-1.27)	0.53 (0.4 - 0.64)

LR(+): Likelihood ratio positivo, LR(-): Likelihood ratio negativo, VPP: Valor predictivo positivo, VPN: Valor predictivo negativo No: Número

Tabla 5. Mortalidad de pacientes con enfermedad por COVID-19 por período, estratificadas por grupo de riesgo de mortalidad 4C y comparada con la cohorte de desarrollo.

Grupos de riesgo puntaje mortalidad 4C	Cohortes			
	Primer periodo	Segundo periodo	Total	Reino Unido (Cohorte de desarrollo) *
Todos los grupos	39.91 % (439/1100)	45.72% (315/ 689)	42.15 % (754/1789)	31.3% (18155/57924)
Bajo (0-3)	0 % (0/27)	11.11% (2/18)	4.44 % (2/45)	1.5% (65/4224)
Intermedio (4-8)	15.65% (54/ 345)	22.16% (43/194)	18 % (97/ 539)	9.4% (1237/13166)
Alto (9-14)	46.99% (265/ 564)	52.73% (193/366)	49.25 % (458/930)	34% (9976/29755)
Muy alto (15-21)	73.17% (120/164)	69.37 % (77/111)	71.64 % (197/275)	64% (6877/10679)

*Datos reproducidos de (7)

Tabla 6. Mortalidad de pacientes con enfermedad por COVID-19 por período, estratificadas por grupo de riesgo del puntaje NEWS2 y comparada con la cohorte de desarrollo.

Grupos de riesgo puntaje NEWS2	Cohortes			
	Primer periodo	Segundo periodo	Total	Cohorte de desarrollo*
Todos los grupos	39.91% (439/1100)	45.72% (315/ 689)	42.15 % (754/1,789)	14,8%
Bajo (0-4)	25.87% (74/ 286)	31.65% (50/158)	27.93 % (124/444)	5.5 %
Medio (5-6)	35.93% (106/295)	36.96% (68/184)	36.33 % (174/479)	11.3%
Alto (≥7)	49.90% (259/519)	56.77% (197/ 347)	52.66% (456/866)	27.03%

*Datos reproducidos de (10,20)

Figura 1: Diagrama de flujo de inclusiones y exclusiones del estudio

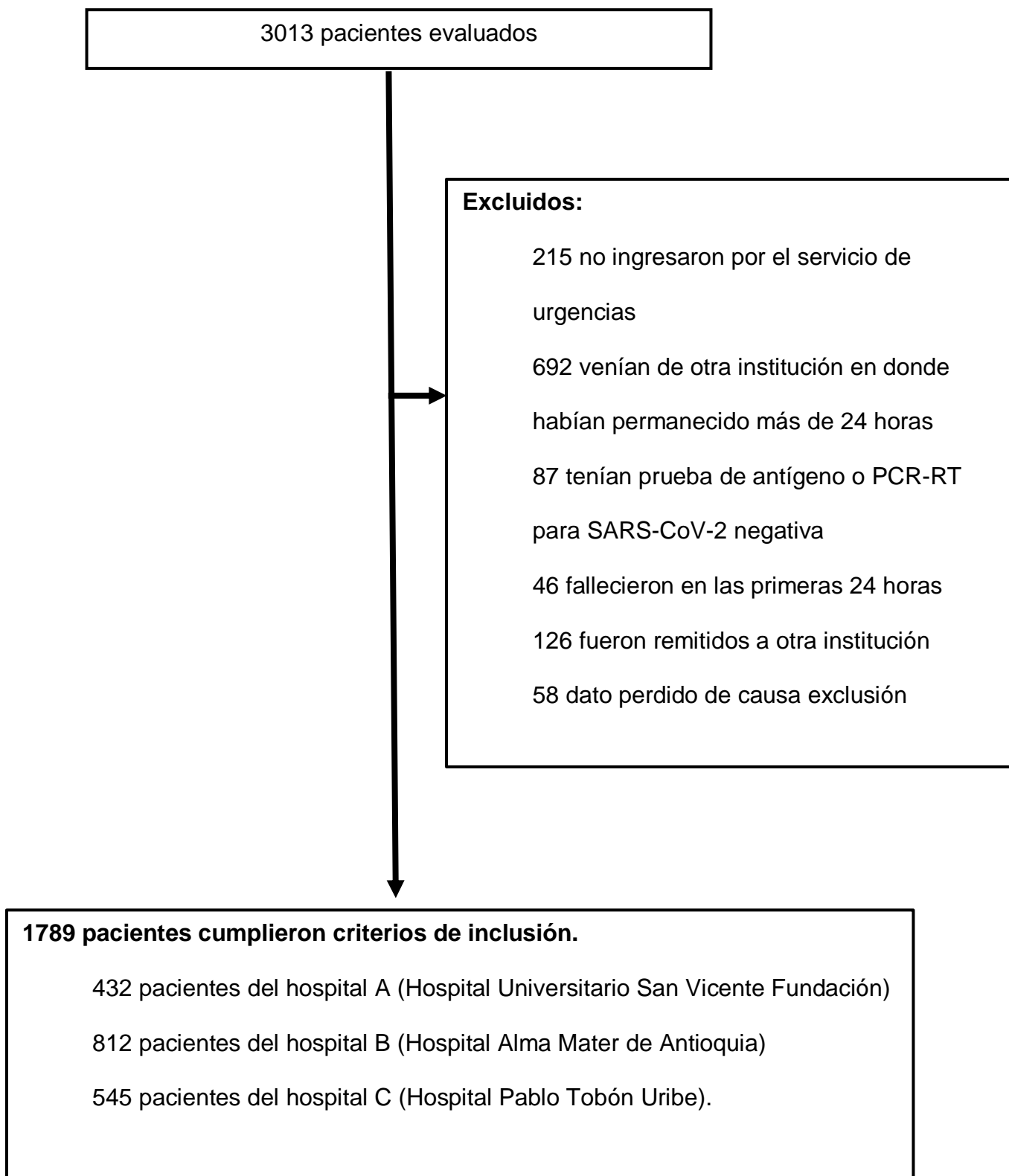
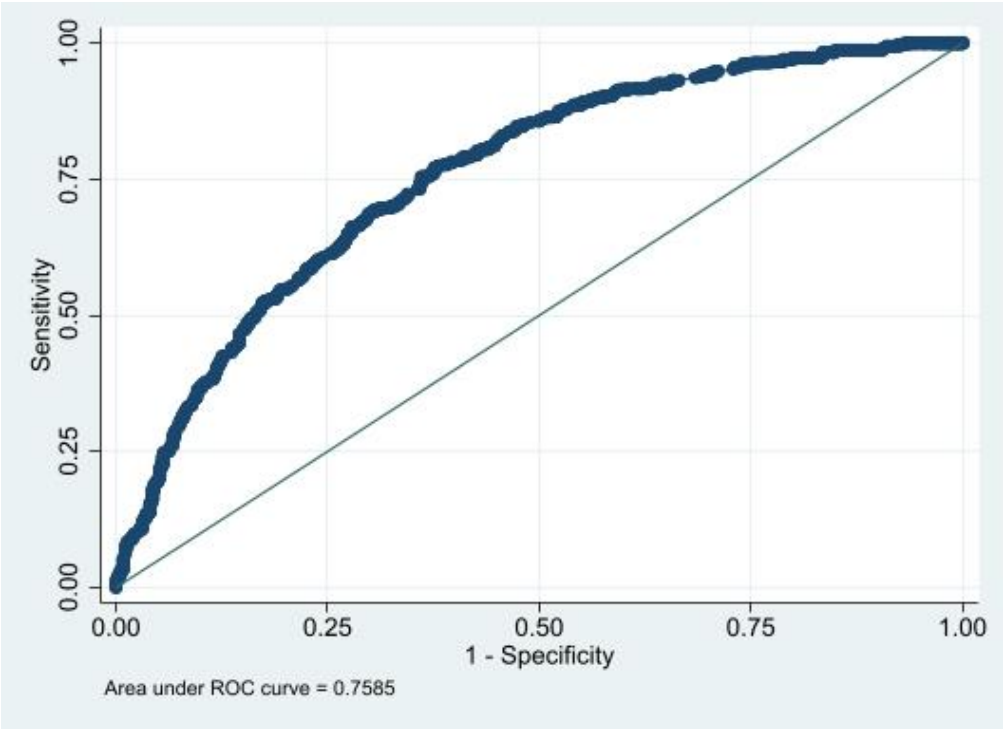
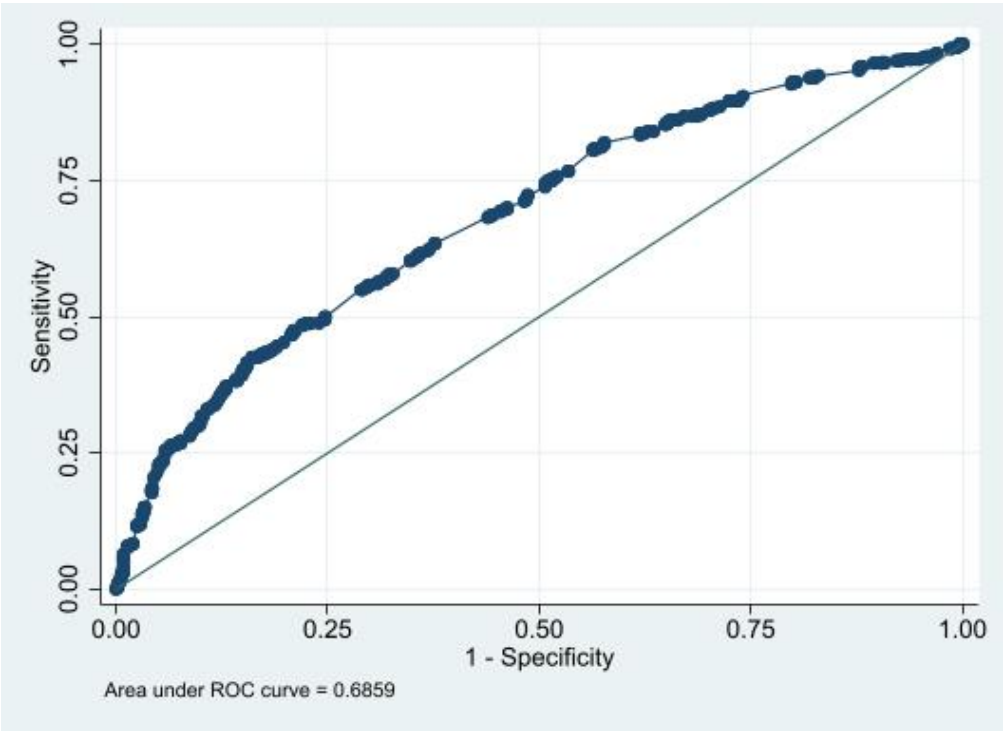


Figura 2: Curva de Característica Operativa del Receptor para mortalidad en pacientes con COVID-19.

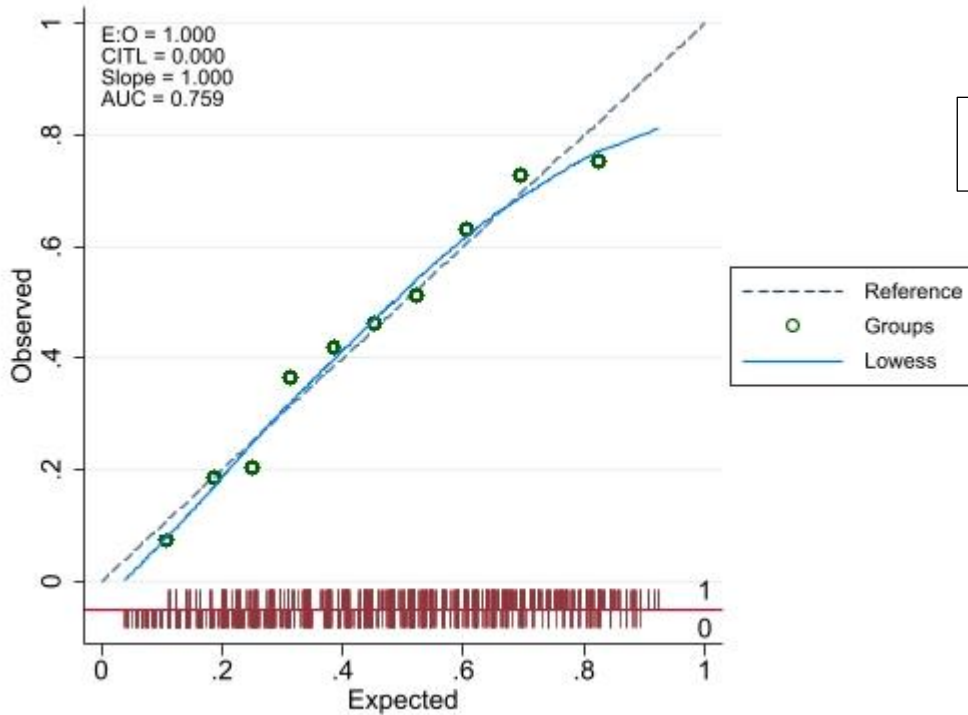


A. Puntuación de mortalidad 4C.

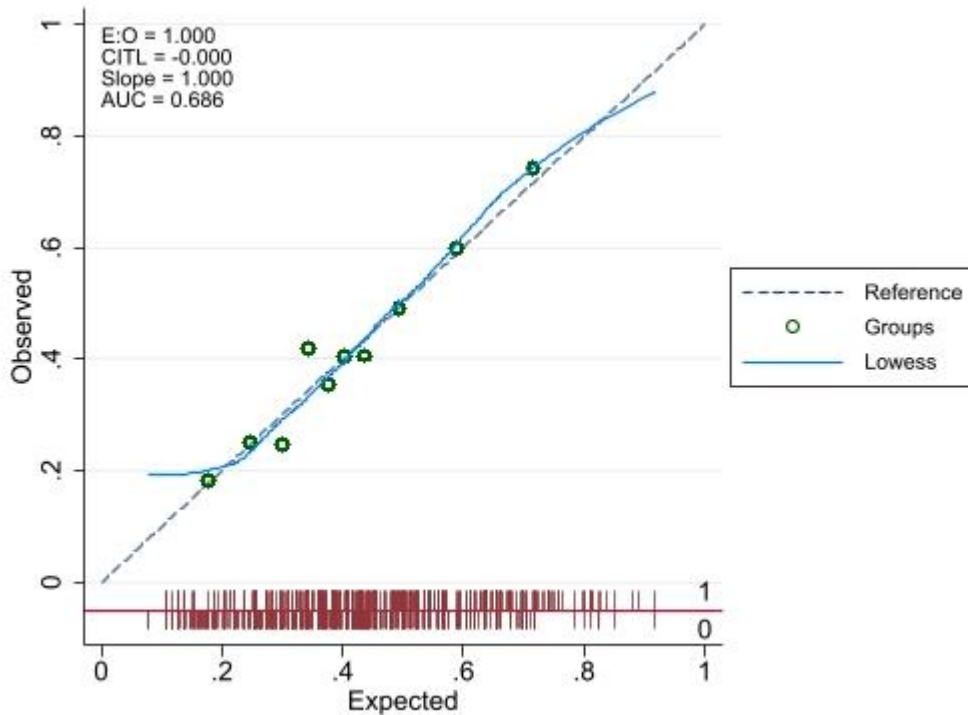


B. Puntuación NEWS2

Figura 3: La probabilidad de resultado prevista frente a la observada se muestra como gráficos de calibración



A. Puntuación de mortalidad 4C.



B. Puntuación NEWS2