



Capacidad predictiva del puntaje TREWS en la mortalidad a 24 horas y al alta, en pacientes valorados en el servicio de urgencias de la Fundación Santa Fe de Bogotá en el año 2023

Autores

Yurany Esperanza Rodriguez Torres

Sebastián Ricardo Andrade Moncada

Luis Fernando Rodríguez Gómez

Directores

German Alberto Devia Jaramillo

Danna Lesley Cruz Reyes

Trabajo presentado como requisito para optar por el título de especialista en Medicina de Emergencias

Bogotá - Colombia

2025

Capacidad predictiva del puntaje TREWS en la mortalidad a 24 horas y al alta, en pacientes
valorados en el servicio de urgencias de la Fundación Santa Fe de Bogotá
en el año 2023

Autores

Yurany Esperanza Rodriguez Torres
Sebastián Ricardo Andrade Moncada
Luis Fernando Rodríguez Gómez

Tutores

German Alberto Devia Jaramillo
Danna Lesley Cruz Reyes

Escuela de Medicina y Ciencias de la Salud
Especialización en Medicina de Emergencias
Universidad del Rosario

Bogotá - Colombia

2025

Identificación del proyecto

Institución académica: Universidad del Rosario

Dependencia: Escuela de Medicina y Ciencias de la Salud

Título de la investigación: Capacidad predictiva del puntaje TREWS en la mortalidad a 24 horas y al alta, en pacientes valorados en el servicio de urgencias de la Fundación Santa Fe de Bogotá en el año 2023

Instituciones participantes: Fundación Santa Fe de Bogotá

Tipo de investigación: Estudio de predicción diagnóstica

Investigador principal: Yurany Esperanza Rodríguez Torres

Investigadores asociados: Sebastián Ricardo Andrade Moncada y Luis Fernando Rodríguez Gómez

Asesor clínico o temático: German Alberto Devia Jaramillo

Asesor metodológico: Danna Lesley Cruz Reyes

CONTENIDO

1. Introducción	8
1.1 Planteamiento del problema	8
1.2 Justificación	10
2 Marco Teórico	11
2.1 Marco Conceptual	11
2.2 Estado del Arte	17
3 Pregunta de investigación	20
4 Objetivos	20
4.1 Objetivo general	20
4.2 Objetivos específicos	20
5 Formulación de hipótesis	21
6 Metodología	21
6.1 Tipo y diseño de estudio	21
6.2 Población	21
6.3 Criterios de inclusión y exclusión	21
6.4 Tamaño de muestra	22
6.5 Marco muestral	22
6.6 Diseño de la muestra	22
6.7 Variables de estudio	23
6.8 Técnicas, procedimientos e instrumentos de la recolección de datos	24
6.9 Plan de procesamiento de muestras biológicas	25
6.10 Análisis de datos	25
6.11 Alcances y límites de la investigación	26
7 Aspectos éticos	26
8 Resultados	28
8.1 Características demográficas y clínicas	29
8.2 Puntaje de TREWS en la población estudiada	30
8.3 Mortalidad en la cohorte de estudio	31
8.4 Punto de corte óptimo del puntaje de TREWS a 24 horas	31
8.5 Punto de corte óptimo del puntaje de TREWS al alta	32
8.6 Capacidad predictiva del puntaje TREWS para la mortalidad medida a 24 horas y al alta	33
8.7 Comparación de la capacidad de predicción del puntaje TREWS vs Triage	35

9	Discusión	37
10	Conclusión	41
11	Administración del proyecto	42
	<i>11.1 Presupuesto</i>	42
	<i>11.2 Cronograma</i>	42
12	Referencias	43

Resumen

Introducción: En los servicios de urgencias con alta demanda asistencial resulta fundamental contar con estrategias eficientes que permitan discriminar y categorizar el riesgo desde el momento del ingreso. La detección temprana de pacientes con alto riesgo de deterioro no solo reduce los tiempos de espera, sino que también agiliza la toma de decisiones y optimiza el uso de los recursos disponibles. En este escenario, los sistemas de alerta temprana (EWS) se han consolidado como herramientas sencillas y costo-efectivas para el monitoreo clínico. Actualmente, el NEWS es el más empleado debido a su destacado desempeño predictivo, aunque no fue diseñado específicamente para el triaje. En contraste, el TREWS fue desarrollado para aplicarse desde el ingreso a urgencias y ha demostrado un notable rendimiento en la predicción de mortalidad a 24 horas (AUC 0.906). No obstante, pese a su potencial, aún no se han publicado estudios que avalen su implementación en Latinoamérica.

Objetivo: Determinar la capacidad de predicción del puntaje TREWS (Triage in Emergency Department Early Warning Score) para riesgo de mortalidad, medida a 24 horas y al alta, en los pacientes mayores de 18 años valorados en el servicio de urgencias de la Fundación Santa Fe de Bogotá, entre enero y junio de 2023.

Métodos: Se realizó un estudio de predicción diagnóstica en el servicio de urgencias de la Fundación Santa Fe de Bogotá, entre enero y junio de 2023. Se incluyeron en el estudio pacientes adultos, mayores de 18 años, que acudieron al servicio de urgencias por cualquier motivo y fueron evaluados en la sala de triaje.

Resultados: Se analizó una muestra de 445 pacientes. La mortalidad fue de 2.9% en 24 horas y de 4.7% al alta. 51.6% eran mujeres con una mediana de edad de 67 años. El AUC del puntaje TREWS para la mortalidad a 24 horas fue de 0,856 (IC 95%: 0,708-1.00, $p < 0.001$); con un punto de corte de 6 puntos se obtuvo una sensibilidad de 84.6% (IC 95% 61.1-100%), una especificidad de 85.2% (IC 95% 81.7-88.6%), un VPP de 14.7% (IC 95% 5.99-23.3%), un VPN de 99.4% (98.6-100), un LR+ de 5.71 (4.13-7.9) y un LR- de 0.18 (0.05-0.65). El AUC del puntaje TREWS para la mortalidad al alta fue de 0,861 (IC 95%: 0,755-0,967 $p < 0.001$); igualmente con un punto de corte de 6 puntos se obtuvo una sensibilidad de 76.2% (IC 95% 55.6-96.8%), una especificidad de 86.1% (IC 95% 82.7-89.5%), un VPP de 21.3% (IC 95% 11.4-31.2%), un VPN de 98.6% (97.3-100), un LR+ de 5.48 (3.91-7.7) y un LR- de 0.28 (0.13-0.59).

Conclusión: Este estudio demuestra que el puntaje TREWS posee una notable capacidad predictiva para la mortalidad en pacientes desde el ingreso a urgencias, tanto a las 24 horas como al egreso. Un punto de corte de 6 se identificó como el umbral más confiable, al ofrecer el mejor equilibrio entre la detección temprana de pacientes en riesgo y la preservación de la capacidad para descartar desenlaces adversos. Además, su desempeño supera incluso la capacidad predictiva del puntaje NEWS reportada en la literatura.

Abstract:

Introduction: In emergency departments with high demand for care, it is essential to have efficient strategies in place to identify and categorize risk from the moment of admission. Early detection of patients at high risk of deterioration not only reduces waiting times, but also speeds up decision-making and optimizes the use of available resources. In this scenario, early warning systems (EWS) have established themselves as simple and cost-effective tools for clinical monitoring. Currently, NEWS is the most widely used due to its outstanding predictive performance, although it was not specifically designed for triage. In contrast, TREWS was developed for use from the moment of admission to the emergency department and has demonstrated remarkable performance in predicting 24-hour mortality (AUC: 0.906). However, despite its potential, no studies have yet been published to support its implementation in Latin America.

Objective: To determine the predictive capacity of the TREWS (Triage in Emergency Department Early Warning Score) score for mortality risk measured at 24 hours in patients over 18 years of age assessed in the emergency department of the Fundación Santa Fe de Bogotá between January and June 2023.

Methods: A diagnostic prediction study was conducted in the emergency department of the Fundación Santa Fe de Bogotá between January and June 2023. The study included adult patients over the age of 18 who visited the emergency department for any reason and were evaluated in the triage room

Results: A sample of 445 patients was analyzed. Mortality was 2.9% within 24 hours and 4.7% at discharge. 51.6% were women with a median age of 67 years. The AUC of the TREWS score for mortality at 24 hours was 0.856 (95% CI: 0.708-1.00, $p < 0.001$); with a cutoff point of 6 points, a sensitivity of 84.6% (95% CI 61.1-100%), a specificity of 85.2% (95% CI 81.7-88.6%), a PPV of 14.7% (95% CI 5.99-23.3%), a NPV of 99.4% (98.6-100), a LR+ of 5.71 (4.13-7.9) and a LR- of 0.18 (0.05-0.65). The AUC of the TREWS score for mortality at discharge was 0.861 (95% CI: 0.755-0.967 $p < 0.001$); Similarly, with a cut-off point of 6 points, a sensitivity of 76.2% (95% CI 55.6-96.8%), a specificity of 86.1% (95% CI 82.7-89.5%), a PPV of 21.3% (95% CI 11.4-31.2%), a NPV of 98.6% (97.3-100), a LR+ of 5.48 (3.91-7.7) and a LR- of 0.28 (0.13-0.59).

Conclusion: This study demonstrates that the TREWS score has remarkable predictive capacity for mortality in patients from the moment they are admitted to the emergency department, both at 24 hours and at discharge. A cut-off point of 6 was identified as the most reliable threshold, offering the best balance between early detection of patients at risk and preserving the ability to rule out adverse outcomes. Furthermore, its performance even exceeds the predictive ability of the NEWS score reported in the literature.

1. Introducción

1.1 Planteamiento del problema

A nivel mundial, los servicios de urgencias reciben un volumen creciente de pacientes, lo que genera una presión constante sobre su capacidad operativa y plantea retos significativos que comprometen la calidad de la atención y los desenlaces clínicos. Entre los factores más estrechamente asociados a un peor pronóstico se encuentran el tiempo de espera para la atención inicial y la prolongación de la estancia en urgencias, los cuales se han vinculado con un incremento de la mortalidad de hasta un 5% con un aumento en los costos por ingreso de aproximadamente un 1% (1,2). En Colombia, el monitoreo a los servicios de urgencias realizado por la defensoría del pueblo en 2020 puso en evidencia este fenómeno, encontrando un nivel de sobreocupación a nivel nacional del 142% (3). Lo que podría estar en relación con el incremento de la mortalidad evidenciado en el informe ASIS (*Así vamos en salud*), que para nuestro país, reporto entre 2005 y 2018 una tasa de mortalidad de 4,75 muertes por cada 1.000 personas (4).

Con el fin de optimizar el uso del tiempo y los recursos, en la actualidad los servicios de urgencias emplean sistemas de triaje estructurados en cinco niveles para clasificar pacientes desde su ingreso a urgencias. Estos requieren personal profesional, formación específica y la integración de variables objetivas, como los signos vitales, con un componente subjetivo que introduce una variabilidad considerable en su interpretación. Esta variabilidad ha sido documentada en una revisión sistemática reciente, en la que los estudios evaluados reportaron un índice kappa superior a 0,8 con valores de sensibilidad para predicción de mortalidad que rondan el 90% y con una tendencia al sobretriaje, es decir a sobre clasificar a algunos pacientes como de alta urgencia usando más recursos de los necesarios (5).

Sistemas de alerta temprana

Es en este contexto, que surgen los sistemas de alerta temprana, conocidos por su sigla en inglés como EWS (*Early Warning Scores*). Originalmente diseñados para identificar pacientes en riesgo de deterioro clínico en unidades de cuidado intensivo y hospitalización; no obstante, su uso se ha extendido progresivamente a múltiples escenarios, incluido el servicio de urgencias (6). Estos sistemas se fundamentan en la medición de signos vitales —como la frecuencia cardíaca, la presión arterial, la saturación de oxígeno, la temperatura y la puntuación en la Escala de Coma de Glasgow— y, en algunos casos, incorporan variables de laboratorio o datos demográficos para detectar de forma temprana cambios sutiles en el estado del paciente y anticipar eventos adversos (6).

Con el tiempo, se han desarrollado y actualizado nuevos EWS con el fin de obtener cada vez un mejor rendimiento diagnóstico. En 1997 Morgan y colaboradores proponen el primer EWS que utilizaba únicamente como variables los signos vitales (7), unos años después en 2001 surge el MEWS (*Modified Early Warning Score*) que mejora el área bajo la curva adicionando el estado de alerta y la edad mayor a 70 años como factor de riesgo (8), posteriormente el ViEWS (*Vital Early Warning Score*) creado en 2010 incluye el uso de oxígeno (9) y finalmente en 2012 el Royal College of Physicians propone el NEWS (*National Early Warning Score*) una recopilación de trabajos anteriores que estandariza la frecuencia de medición de los parámetros y los rangos de los mismos (10).

Actualmente, existen más de 100 sistemas de puntuación de alerta temprana, los cuales se han convertido en herramientas de uso cotidiano en la práctica clínica hospitalaria. Esta tendencia se consolidó aún más a partir de la pandemia por COVID-19, cuando se introdujo el NEWS 2 (*National Early Warning Score 2*); una actualización del sistema NEWS original. Esta versión incorporó variables como los niveles de saturación de oxígeno y la presencia de episodios de confusión aguda, lo que ha contribuido a que el NEWS 2 se consolide como uno de los sistemas de puntuación más utilizados y reconocidos en la actualidad (11). También resulta destacable la creación e implementación de sistemas de puntuación adaptados a poblaciones específicas. Tal es el caso del PEWS (*Pediatric Early Warning Score*) diseñado para la población pediátrica y del OBWS (*Obstetric Early Warning Score*) desarrollado específicamente para pacientes obstétricas. (12,13). Muchas de estos EWS han demostrado gran capacidad predictiva para desenlaces como mortalidad, por ejemplo, en una comparativa de seis de estas escalas frecuentemente usadas en urgencias, el NEWS demostró un AUC (Área bajo la curva) de 0.904 [IC 95%, 0.805 - 0.913] y el NEWS 2 de 0.901 [IC 95%, 0.892 - 0.910] siendo superiores al resto de escalas evaluadas (14).

A pesar del creciente rendimiento diagnóstico de los EWS persisten limitaciones relevantes cuando se aplican en el contexto de urgencias: en primer lugar, la mayoría fueron desarrollados en poblaciones hospitalizadas o en unidades de cuidado intensivo, entornos con una mayor prevalencia de enfermedad crítica que el servicio de urgencias (15). Además, su especificidad suele ser limitada, lo que incrementa el riesgo de sobretriaje y de asignar recursos a pacientes que no los requieren de forma inmediata. En cuanto a la predicción de mortalidad, suelen mostrar un mejor desempeño para estimar mortalidad a corto plazo —durante el mismo episodio de urgencias—, pero su capacidad predictiva disminuye para desenlaces a más largo plazo (6). Finalmente, la mayoría de estos sistemas han sido diseñados en contextos ajenos a Latinoamérica y Colombia, si bien algunos han sido validados localmente (16) no ha sido posible implementar de forma efectiva sistemas de alerta

automáticos —por ejemplo, integrados en la historia clínica electrónica— que se activen en tiempo real al ingresar los signos vitales del paciente (15).

Puntaje TREWS

Teniendo en cuenta estas limitaciones y enfocándose específicamente en la población atendida en urgencias, en 2019 Lee y colaboradores, en la República de Corea, desarrollaron un nuevo sistema de alerta temprana basado en las mismas variables empleadas por el conocido NEWS, pero asignando valores y ponderaciones diferentes a cada una (17). Para ello, utilizaron la base de datos NEDIS (*National Emergency Department Information System*), que recopila información de 146 servicios de urgencias. Con los datos de 81.520 pacientes, crearon el **TREWS** (*Triage in Emergency Department Early Warning Score*), una herramienta que, en su estudio original, mostró un rendimiento superior al de otras escalas similares para predecir la mortalidad a 24 horas en pacientes adultos ingresados a urgencias [AUC 0.906; IC 95%, 0.903 – 0.908] (17).

En junio de 2021, el TREWS fue evaluado en pacientes con infección confirmada por COVID-19, excluidos del estudio original de Lee, demostrando una efectividad comparable a la de la escala MEWS (18). Los resultados disponibles hasta la fecha sugieren que el puntaje TREWS es un mejor predictor de complicaciones tempranas y mortalidad que las escalas de alerta temprana actualmente utilizadas en los servicios de urgencias (17).

1.2 Justificación

El deterioro agudo de los pacientes suele estar precedido por cambios en los parámetros fisiológicos, si estos cambios se identifican precozmente se pueden implementar medidas que disminuyen las complicaciones e inclusive las tasas de mortalidad (11). Actualmente existen múltiples escalas de alerta temprana que se han modificado a través del tiempo con el fin de predecir con mayor exactitud la necesidad de intensificar la monitorización del paciente, brindando una mayor ventana de tiempo para actuar (19).

En este contexto, el TREWS surge como una herramienta diseñada específicamente para implementarse desde el ingreso a urgencias, capaz de identificar a los pacientes con alto riesgo de deterioro agudo. Su implementación promete brindar al personal de salud un soporte objetivo para la toma rápida de decisiones, facilitando la activación temprana de equipos de respuesta rápida, una clasificación más precisa de la prioridad asistencial y una asignación más eficiente de los recursos. Esto, en teoría, podría traducirse en mejoras sustanciales en los tiempos de atención, así como en la

reducción de complicaciones y mortalidad a corto y mediano plazo (17).

Hasta la fecha no se ha evaluado su desempeño predictivo en la población que ingresa a urgencias en servicios de alta complejidad en Colombia lo cual limita su aplicación práctica. Por esto, es necesario establecer el rendimiento predictivo y diagnóstico de esta herramienta para así ofrecer la posibilidad de tener un reconocimiento oportuno de pacientes críticos a su entrada al servicio de urgencias y potencialmente afectar desenlaces importantes como la mortalidad.

2 Marco Teórico

2.1 Marco Conceptual

¿Cómo funcionan los sistemas actuales de Triage?

Existen diferentes sistemas de triaje, cada uno clasificando pacientes con una cantidad variable de niveles de atención, desde los 2 niveles hasta la actual clasificación de 5 niveles que es basada en los sistemas de triaje implementados en Estados Unidos con el Grupo ESI (Emergency Severity Index) y el grupo de triaje de Manchester dependiente del Sistema Nacional de Salud (20).

El ESI fue desarrollado por dos emergenciólogos, Richard Wuerz y David Eitel en 1998, su funcionamiento se basa en la predicción del requerimiento de recursos necesarios para la atención de un paciente en urgencias y los signos vitales al ingreso. Este sistema identifica pacientes inestables o con alto riesgo de inestabilizarse; su mayor cualidad es tener un lenguaje común y de fácil entendimiento (20).

El sistema de triaje ESI se fundamenta en un algoritmo secuencial compuesto por cuatro preguntas. La primera es: ¿el paciente requiere una intervención inmediata para salvar su vida? Si la respuesta es afirmativa, se clasifica como nivel 1; si es negativa, se pasa a la segunda pregunta: ¿presenta una situación de alto estrés, alteración mental o confusión aguda? Una respuesta afirmativa lo ubica en nivel 2; si es negativa, se continúa con la tercera pregunta: ¿cuántos recursos se prevé que requerirá para su atención? Se considera recurso la realización de exámenes paraclínicos (sanguíneos o urinarios), estudios de imagen, valoraciones por especialistas o procedimientos. Con base en el número de recursos estimados y los signos vitales, el paciente se clasifica en los niveles 3, 4 o 5. Si en cualquier momento se detectan signos vitales anormales, el algoritmo permite retroceder y asignar un nivel superior de prioridad (20).

En Colombia, desde el año 2015, la Resolución 5596 implementó un sistema de triaje de cinco niveles adaptado a las necesidades del país, el cual establece tiempos máximos de atención para los casos de mayor prioridad. En este esquema, el nivel 1 exige atención inmediata, mientras que el nivel 2 debe ser atendido en un plazo máximo de 30 minutos desde su registro, dado que corresponde a condiciones con alto riesgo de deterioro rápido. Estos tiempos no aplican en situaciones de eventos con múltiples víctimas (21).

¿Qué son y cómo se desarrollaron los EWS?

De la necesidad de disminuir la mortalidad y los ingresos a unidad de cuidados intensivos (UCI) surgieron los sistemas de alerta temprana, cuya finalidad es que a través de variaciones en las constantes vitales se pueda detectar de manera temprana el deterioro clínico de un paciente; disminuyendo así la mortalidad a través de activación de sistemas y equipos de respuesta rápida. Los EWS también conocidos como PTTS (Physiologic track and trigger systems) son sistemas que permiten alertar a sistemas de respuesta rápida. Entre los cuales resaltan el CCOS (Critical Care Outreach Service) del Reino Unido y el MET en Australia, que han ganado popularidad al ser parte de una campaña que adelanto la OMS, International Partnership for Acute Care Safety (IPACS) (22). Así como el NEWS y NEWS 2, propiciados por el Royal College of Physicians de Reino Unido y que se han posicionado como dos de los modelos con mejor rendimiento diagnóstico a la hora de predecir mortalidad temprana, requerimiento de UCI e incluso detección de sepsis (23). Teniendo en cuenta sus características y el diseño de los puntajes se podrían dividir en varias categorías: las de parámetro sencillo, múltiple y sistemas ponderados. (24)

Dentro de la primera categoría está el sistema MET, diseñado en Australia, es uno de los más reconocidos; este permite la activación de un sistema de médicos en respuesta a unas variables o situaciones determinadas y busca disminuir el desenlace compuesto de paro cardíaco, muerte súbita e ingreso no planeado a UCI en los 6 meses siguientes a la activación del MET (25). En este tipo de EWS existe una lista de parámetros fijos y el sistema se activa cuando cualquiera de esos parámetros se cumple. Al haber demostrado menor capacidad de predicción de eventos adversos que los sistemas ponderados, los EWS de esta categoría se usan cada vez menos (25).

Los sistemas de múltiples variables son los que permiten una combinación de diferentes variables fisiológicas sin calcular un puntaje para activar un sistema de respuesta rápida. Son los menos utilizados sin embargo tienen la ventaja de poder activar un sistema de respuesta rápida sin realizar cálculos complejos. Bleyer et al. en un estudio realizado en Carolina del Norte analizaron variables

que se asociaron con un aumento en la mortalidad del 5% en la población mayor de 70 años, encontrando que la presencia de más de 3 signos vitales alterados se asociaba a un aumento de la mortalidad de hasta 33%, esto comparado con tener una variable alterada que reflejaba una mortalidad del 0,92%, siendo en ocasiones variables alteradas hasta 48 horas antes de la muerte. Al analizar las áreas bajo las curvas ROC comparando este planteamiento contra ViEWS y MEWS tenía valores similares (0.85, 0.87 y 0.86) (26).

En la tercera categoría se encuentran los puntajes más utilizados en la actualidad. Estos son los más complejos, ya que clasifican los signos vitales y otras variables clínicas, asignan una ponderación en puntos según el grado de anormalidad fisiológica. Su principal ventaja radica en que permiten ajustar el puntaje de riesgo de acuerdo con el parámetro alterado, reconociendo que algunas alteraciones implican un mayor riesgo de deterioro que otras. Sin embargo, este enfoque puede introducir errores derivados del cálculo del puntaje. Hasta la fecha, se han publicado más de 100 EWS de esta categoría basados en el trabajo original de Morgan et al., la mayoría de ellos adaptaciones específicas basadas en opinión de expertos por expertos para ajustarse a las necesidades de cada centro hospitalario (6).

El MEWS publicado en el año 2001, es una escala que toma como base los signos vitales y un parámetro adicional relacionado con el estado de alerta, el AVPU (Alert, Reactive to Voice, Reactive to Pain, Unresponsive); con un puntaje que iba desde 0 a 12 puntos calificaba como críticos a los pacientes con puntajes superiores a 5 puntos. En su estudio de publicación original se documentó que los pacientes hospitalizados con puntajes superiores a 5 puntos tenían un OR (Odds Ratio) de hasta 10,9 [IC 95%, 2.2 - 55.6] para ingreso a UCI y de 5,4 [IC 95%, 2.8 - 10.7] para mortalidad (8). Adicionalmente, detectaron que los pacientes que sobrepasan los 70 años tenían mayor riesgo de alcanzar los desenlaces de mortalidad o ingreso a UCI [OR 6.5, IC 95% 2.5 - 16.7] comparados con la población menor de 50 años, por lo que agregaron este factor adicional, mostrando así un aumento del AUC de 0.67 a 0.72. Sin embargo tenía la limitación de ser un estudio monocéntrico y con un número limitado de pacientes, en la época de la publicación de este EWS lo vieron como una herramienta de uso constante puesto que utilizaba variables de recolección rutinaria y que sería de utilidad para la clasificación de pacientes en el triaje evidenciando una relación directa entre puntajes altos y la mortalidad (8).

En el año 2006 el grupo de la Dra. Paterson en Edimburgo presento un EWS que desarrollo teniendo en cuenta las siguientes variables: frecuencia respiratoria, saturación de oxígeno, temperatura, presión sistólica, frecuencia cardiaca y nivel de conciencia. A este sistema se le denomino SEWS (Standarized Early Warning Score). En el diseño del estudio se realizó una muestra en dos periodos de 11 días en

dos meses separados, el primero sería el grupo control y para el segundo grupo se diseñó una tabla de fácil llenado que a su reverso contenía pautas de manejo inicial en los bajos niveles de atención. Ambos brazos del estudio debían incluir por lo menos 400 pacientes. Utilizaron una prueba de chi cuadrado para determinar la asociación de puntaje con mortalidad, mostrando que esta fue de hasta 8 veces mayor comparando puntajes de 0-3 con puntajes ≥ 4 lo que se correlaciono con una disminución de la mortalidad así: pre-SEWS 5.8% y post-SEWS 3.0% (P 0.046) (20).

Prytherch en 2011 plantea el ViEWS (Vital Early Warning Score), un modelo basado en papel que no requería el uso de herramientas electrónicas para su aplicabilidad. Con esta herramienta se analizaron los desenlaces de los pacientes a las 24 horas de admisión con seguimiento hasta 120 horas. Cuando se adiciono como variable el uso de oxígeno suplementario se demostró una curva ROC con un AUC de 0.880, lo cual fue superior al demostrado previamente por MEWS + Edad que tenía un AUC de 0.850 (9).

Para el año 2012 el *Royal College of Physicians*, nota la proliferación de EWS y su variabilidad — con versiones distintas incluso entre hospitales—, así que propuso y financió la creación de un sistema unificado a nivel nacional. De esta iniciativa surgió el **NEWS** (*National Early Warning Score*), que incorpora seis parámetros fisiológicos registrados de forma rutinaria en la atención de pacientes hospitalizados: frecuencia respiratoria, saturación de oxígeno, temperatura, presión arterial sistólica, frecuencia cardíaca y nivel de consciencia según la escala AVPU. Además, añade un séptimo parámetro previamente incluido en el ViEWS: el uso de oxígeno suplementario (10). En el NEWS, cada signo vital se clasifica en categorías con puntajes asociados, cada signo vital puede dar entre cero y tres puntos; por ejemplo, una frecuencia cardíaca entre 51–90 lpm otorga 0 puntos, mientras que una frecuencia entre 91–110 lpm otorga 1 punto. Con la suma total se definen tres categorías de riesgo: bajo (0–4 puntos), intermedio (5–6 puntos; requiere valoración estricta) y alto (≥ 7 puntos; requiere valoración urgente y traslado a monitorización continua o UCI) (10). En 2015, esta escala fue reformulada para ajustar los puntajes de saturación de oxígeno en pacientes con falla respiratoria hipercápnica o con hipercapnia crónica, como aquellos con enfermedad pulmonar obstructiva crónica, originando así la versión NEWS 2 (11).

Puntaje TREWS

En 2020 Sang Bong Lee y sus colegas del hospital universitario de Gyeongsang en Corea del Sur deciden evaluar el uso del NEWS como predictor de mortalidad intrahospitalaria a 24 horas en pacientes que ingresan a urgencias, para esto usan la base de datos NEDIS (*National Emergency*

Department Information System), que recopila información de 146 servicios de urgencias a nivel nacional, de donde terminan usando datos de 81,520 pacientes (17). El estudio incluyó un análisis de regresión logística univariada para cada uno de los seis signos vitales contemplados en el NEWS. Con esos datos, identificaron puntos de corte óptimos para cada signo vital y asignaron nuevos pesos y puntuaciones a cada variable. De este proceso surgió el TREWS (Triage in Emergency Department Early Warning Score), un sistema adaptado a las características de los pacientes atendidos en urgencias (17). Un aspecto interesante es que, durante su desarrollo, decidieron excluir la variable de saturación periférica de oxígeno, al comprobar que su inclusión no aportaba mejoras relevantes en la capacidad predictiva del modelo. En la siguiente tabla se detalla el método de cálculo del puntaje TREWS:

Criterios de la Puntuación *	Puntajes asignados						
	3	2	1	0	1	2	3
Presión arterial sistólica (mmHg)	≤ 90	91 - 100	101 - 110	≥ 111	-	-	-
Frecuencia cardíaca (latidos/min)	≤ 40	-	-	≤ 110	111 - 120	≥ 121	-
Frecuencia respiratoria (respiraciones/min)	≤ 8	9-11	12 - 17	18 - 24	-	-	≥ 25
Temperatura corporal (°C)	≤ 35.0	35.1 - 36.0	-	≥ 36.0	-	-	-
Nivel de conciencia	-	-	-	A	-	-	V, P, I
Suministro de oxígeno	-	-	-	No	-	-	Sí
Edad (años)	≥ 71	-	41 - 70	≤ 40	-	-	-

* Puntaje de TREWS. Un puntaje > 4 puntos predicen mortalidad a 24 horas y necesidad de activar sistemas de alerta. Puntaje mínimo de 0 puntos y máximo de 21 puntos. A: Alert, V: Reactive to Voice, P: Reactive to Pain, U: Unresponsive. Fuente: Adaptado de Lee SB et al. Triage in Emergency Department Early Warning Score (TREWS) is predicting in-hospital mortality in the emergency department. *Am J Emerg Med.* 2020;38(2):203-10.

Para someter a prueba su nueva herramienta crean dos poblaciones de la misma cohorte, una para la derivación (n = 54.347 pacientes) y otra para validación (n = 27.173), en ambos grupos comparan el rendimiento del TREWS contra el NEWS, MEWS y REMS. Los autores definieron para el TREWS un punto de corte óptimo de > 4 puntos para predecir mortalidad a 24 horas en pacientes adultos que ingresan a urgencias. Este umbral se determinó tras analizar cada componente del NEWS con regresión logística univariada, asignar nuevos pesos y evaluar el rendimiento del puntaje total frente

al desenlace de mortalidad temprana. Con este corte, el TREWS mostró un AUC de 0.906 [IC 95% 0.903 - 0.908] para mortalidad a 24 horas, superando a otros EWS comparados en el mismo trabajo. En cuanto a la validez del puntaje:

- Validez de contenido: Como se mencionó previamente, los ítems que componen este puntaje corresponden a signos vitales, los mismos que han sido utilizados para la creación y validación de múltiples EWS previamente. Además, los elementos incluidos para desarrollar este puntaje en particular son los mismos presentes en el estándar actual o “estado del arte” de los EWS: el NEWS (17).
- Validez de criterio: El puntaje TREWS fue comparado, tanto en la cohorte de derivación como en la de validación, con el NEWS (considerado el Gold standard), así como con el MEWS y el REMS (17).
- Validez de constructo: La escala evalúa parámetros clínicos fisiológicos que representan el deterioro clínico y que, de forma individual, han demostrado capacidad para predecir mortalidad (6).
- Validez de apariencia: De forma subjetiva, la escala resulta sencilla de calcular y su resultado es clínicamente relevante, lo que favorece su adopción en la práctica clínica.
- Reproducibilidad: Al basarse en parámetros objetivos, medibles de forma estandarizada y registrados rutinariamente en la atención de pacientes en urgencias, se espera que el puntaje calculado por dos profesionales diferentes sea idéntico (17).
- Estado de validación en otros idiomas: Hasta la fecha, no se han publicado estudios de validación de esta escala en idiomas distintos al original (inglés). Aunque el presente trabajo no es un estudio de validación, se espera que aporte evidencia para reducir esta brecha.
- Validación en poblaciones distintas a la original: Al tratarse de un EWS reciente, no existen aún estudios de validación en poblaciones diferentes a la original. Además del estudio inicial, solo se ha encontrado otro trabajo, de carácter observacional, que evaluó su uso potencial en pacientes con COVID-19 en Turquía, demostrando utilidad predictiva con un AUC de 0.823 [IC 95% 0.764–0.882; $p < 0.001$] (18).

Otras herramientas con nombres similares

En la literatura existen otras escalas con nombres similares que no deben confundirse con el puntaje evaluado en este estudio. El TREW Score es una herramienta distinta, desarrollada y validada (27) para predecir el riesgo de sepsis en contextos clínicos específicos como la insuficiencia respiratoria y

otras patologías críticas (28,29). Por otra parte, el TEWS (Triage Early Warning Score) corresponde a un sistema de alerta temprana diseñado principalmente para su uso en triaje y clasificación inicial de pacientes en servicios de urgencias, con características y objetivos diferentes a los del presente trabajo (30).

2.2 Estado del Arte

¿Qué rendimiento en predicción de mortalidad tienen los EWS usados actualmente como referentes?

Tanto el NEWS como el NEWS 2 han mostrado un desempeño sólido para anticipar eventos adversos como la mortalidad, el ingreso a cuidados intensivos o combinaciones de ambos desenlaces. En 2021, Arévalo y colaboradores publicaron una revisión sistemática con metaanálisis que analizó 165.580 pacientes y confirmó que, entre las escalas de alerta temprana evaluadas, el NEWS era la más utilizada en la literatura disponible hasta ese momento y que, su capacidad para predecir mortalidad a 24 horas era superior a la de las demás escalas incluidas en el estudio [AUC 0.88, IC 95% 0.87 - 0.89; $p < 0,001$]; también mostró un rendimiento aceptable para estimar la probabilidad de hospitalización o ingreso a UCI (31). Respecto al NEWS 2, en 2023 se llevó a cabo un metaanálisis que reunió 30 estudios observacionales con un total de 185.835 pacientes atendidos en servicios de urgencias. Este trabajo estimó una sensibilidad del 81% [IC 95% 76 – 84], una especificidad del 81% [IC 95% 78– 84] y un AUC de 0.88 [IC 95% 0.85 - 0.90] para predecir mortalidad en las primeras 48 horas. No obstante, el rendimiento disminuyó cuando se evaluó la mortalidad a 30 días o al momento del alta hospitalaria (32).

¿Cambian los desenlaces clínicos al hacer este tipo de mediciones?

Una revisión sistemática en 2014 encontró que cuando se implementaba un EWS había una disminución de la mortalidad intrahospitalaria del 2.8% (19); dato previamente evidenciado en 2011 donde un estudio en población con trauma del alta complejidad observó una reducción de la mortalidad del 0,4% para hombres [IC 95% 0,003 - 0,81; mortalidad pre-MEWS: 1,8 %; mortalidad post-MEWS: 1,4 %) y del 1,5% para las mujeres [IC 95% 0.81 – 2.21; mortalidad pre-MEWS: 4,9 %; mortalidad post-MEWS: 3,4 %) (33). Otra variable influenciada por la implementación de escalas de alerta temprana es el tiempo de estadía, como se observa en un estudio de cohortes realizado en 2013 donde se documentó una reducción de 4,55 días en la estancia hospitalaria [IC 95%, 4.34 – 4.76] a 4,11 días [IC 95%, 3.92 – 4.30] (34). En este mismo estudio mostró una disminución de eventos

adversos (Reingreso a UCI en los 5 primeros días del egreso de UCI y número de pacientes sin intento de resucitación) con una reducción de 141 eventos a 78 en el periodo evaluado [IC 95%, 9.5 - 47.2] (34).

Hasta el momento no se han publicado estudios que evalúen si la aplicación del puntaje TREWS influye en los desenlaces clínicos. Sin embargo, los resultados previamente descritos sugieren que existe un potencial para que así ocurra.

¿Se ha evaluado el uso de EWS en el contexto local?

En Latinoamérica también se ha estudiado el uso de sistemas de alerta temprana. Un ejemplo de esto es un estudio publicado en el British Medical Journal en el 2021, en el cual se evaluó la sostenibilidad del uso del PEWS, una escala diseñada en 2005 en Inglaterra para detección temprana de pacientes pediátricos con riesgo de deterioro clínico (9). En este proyecto, al que llamaron el EVAT (Escala de Valoración de Alerta Temprana) se incluyeron 29 hospitales en diferentes países de Latinoamérica y se centraron en pacientes hospitalizados. Los resultados de este proyecto además de lograr la implementación del PEWS como parte de la práctica diaria del cuidado de esta población, lograron demostrar la sostenibilidad de su uso en hospitales de bajos recursos en el contexto latinoamericano (11). Existe también literatura sobre uso de escalas de alerta temprana en obstetricia con resultados prometedores y validación de uso en población latinoamericana y colombiana (12).

Más enfocados hacia la población adulta, existe en PubMed únicamente un estudio sobre escalas de alerta temprana en Latinoamérica. Se trata además de un estudio realizado en Bogotá, Colombia entre 2018 y 2019, en este caso se investigaron pacientes adultos que ingresaron a Urgencias en un hospital de tercer nivel de complejidad. Se propusieron validar el uso del NEWS como predictor de mortalidad en todos los pacientes que ingresaron a urgencias y lo compararon con el sistema de triaje del hospital, es este caso una escala de 5 niveles. Sus resultados arrojaron una curva ROC con un AUC de 0.916 [IC 95%, 0.88 - 0.95] demostrando un rendimiento diagnóstico similar y no inferior al triaje implementado por el hospital, con AUC de 0.88 [IC 95%, 0.85 - 0.91]. Sin embargo, hacen la aclaración de se debe aumentar el tamaño de la muestra. Aun así, es una escala con alta reproducibilidad y que puede ser utilizada en servicios de urgencias en población colombiana (35).

¿Cómo se comporta el Triage de 5 niveles para predecir mortalidad en urgencias?

En una revisión sistemática se encontró que las herramientas de triaje como ESI y Manchester tenían una sensibilidad con amplio margen desde un 0.58 [IC 95% 0.48 - 0.68] hasta 0.88 [IC 95% 0.70 -

0.96] en adultos (36). En 2018 se publicó una revisión sistemática de la literatura que comparo los principales sistemas de triaje (ESI, Manchester y el sistema canadiense), demostró que estos sistemas tienen una alta sensibilidad acumulada de >90% para la predicción de la mortalidad intrahospitalaria en el servicio de urgencias. Sin embargo, esta sensibilidad desciende a menos del 80% cuando se evalúa la capacidad de identificar pacientes con enfermedad crítica y evaluada en diferentes subgrupos tiene una alta variabilidad, para casos de sepsis grave varía entre el 36 y 74% y para enfermedad coronaria varía entre el 44% y 92% (5).

Investigadores iraníes realizaron en 2022 un estudio donde se comparaba el sistema NEWS con el ESI para predicción de requerimiento de UCI y mortalidad intrahospitalaria. En esta cohorte de 1.048 pacientes se observó un AUC para el requerimiento de UCI de 0.8 [IC 95%, 0.71 - 0.89] para NEWS y de 0.88 [IC 95%, 0.81 - 0.95] para ESI. En relación con el desenlace de mortalidad las AUC fueron de 0.82 para NEWS [IC 95%, 0.82 - 0.91] y 0.91 [IC 95%, 0.85 - 0.96] para ESI (37).

¿Qué evidencia tiene el TREWS?

El resultado principal del estudio original de Lee en 2020 fue prometedor: el TREWS predice mortalidad intrahospitalaria con sensibilidad de 94.7% [IC 95%, 93.5% – 95.8%] y especificidad del 63.2% [IC 95%, 62.8% – 63.6%]. En el análisis ROC determinan un AUC de 0.906 [IC 95%, 0.903 - 0.908] si se toma como punto de corte ≥ 4 puntos para predecir mortalidad a 24 horas, que demostró mejor rendimiento con significancia estadística que el NEWS 0.878 [IC 95%, 0.875 - 0.881], MEWS 0.857 [IC 95%, 0.854 - 0.860] y REMS 0.834 [IC 95%, 0.831 - 0.837] (17).

Ahora bien, para Junio del 2025, la búsqueda en la literatura usando los términos MESH pertinentes (Early Warning Score, TREWS AND Emergency department) arroja únicamente una publicación más allá del estudio original en el que se creó la escala, se trata de un estudio de cohorte retrospectivo realizado en Turquía, publicado en Junio de 2021 que evaluó su uso para predecir mortalidad en pacientes con infección confirmada por COVID-19, población excluida en el estudio original de Lee (17). En esta nueva publicación, el análisis ROC encontró un AUC para la mortalidad medida a 28 días de 0.823 [IC 95% 0.764-0.882, $p < 0.001$] para el puntaje de TREWS, en el mismo artículo no se describe el AUC para mortalidad a 24 horas (18).

3 Pregunta de investigación

¿Cuál es la capacidad predictiva del puntaje TREWS (Triage in Emergency Department Early Warning Score) para el riesgo de mortalidad medida a 24 horas y al alta, en los pacientes mayores de 18 años valorados en el servicio de urgencias de la Fundación Santa Fe de Bogotá entre enero y Junio del 2023?

4 Objetivos

4.1 Objetivo general

Determinar la capacidad de predicción del puntaje TREWS (Triage in Emergency Department Early Warning Score) para riesgo de mortalidad medida a 24 horas y al alta, en los pacientes mayores de 18 años valorados en el servicio de urgencias de la Fundación Santa Fe de Bogotá entre enero y Junio del 2023

4.2 Objetivos específicos

1. Caracterizar demográfica y clínicamente la cohorte de estudio.
2. Establecer el puntaje de TREWS en los pacientes que ingresan al servicio de urgencias de la Fundación Santafé de Bogotá.
3. Determinar la tasa de mortalidad de la cohorte de estudio medida a 24 horas y al alta.
4. Identificar el punto de corte óptimo del puntaje de TREWS para predicción de la mortalidad medida a 24 horas y al alta.
5. Establecer la capacidad de predicción del puntaje TREWS (Valor predictivo y área bajo la curva) para riesgo de mortalidad medida a 24 horas y al alta en los pacientes mayores de 18 años valorados en el servicio de urgencias de la Fundación Santa Fe de Bogotá.
6. Establecer la capacidad de predicción del triaje ESI (Valor predictivo y área bajo la curva) para riesgo de mortalidad medida a 24 horas y al alta en los pacientes mayores de 18 años valorados en el servicio de urgencias de la Fundación Santa Fe de Bogotá
7. Comparar la capacidad de predicción del puntaje TREWS vs el Triaje ESI en la mortalidad medida a 24 horas y al alta en los pacientes mayores de 18 años valorados en el servicio de urgencias de la Fundación Santa Fe de Bogotá.

5 Formulación de hipótesis

El puntaje de TREWS ≥ 4 y el óptimo para el estudio tiene mayor capacidad predictiva de la mortalidad medida 24 horas que un puntaje TREWS ≤ 3 en pacientes mayores de 18 años valorados en el servicio de urgencias de la Fundación Santa Fe de Bogotá.

6 Metodología

6.1 Tipo y diseño de estudio

Estudio de predicción diagnóstica en una cohorte retrospectiva.

6.2 Población

Pacientes mayores de 18 años que ingresaron al servicio de urgencias de la Fundación Santafé de Bogotá entre Enero y Junio del 2023.

6.3 Criterios de inclusión y exclusión

6.3.1 Criterios de inclusión:

1. Pacientes mayores de 18 años
2. Que ingresen a urgencias por cualquier causa
3. Que sean valorados en el triaje
4. Que hayan firmado el consentimiento genérico de la institución para el uso de registros en la historia clínica

6.3.2 Criterios de exclusión:

1. Pacientes que sean trasladados a otras instituciones
2. Pacientes que soliciten egreso voluntario
3. Pacientes con indicación de egreso antes de completar 24 horas en la institución
4. Pacientes que ingresan en paro cardiorrespiratoria
5. Mujeres en estado de embarazo confirmado.

6.4 *Tamaño de muestra*

El tamaño de muestra fue calculado con el programa Epidat 3.1, utilizando el módulo “Tamaños de muestra y precisión para pruebas diagnósticas”. Se asumieron valores esperados de sensibilidad y especificidad del 85%, un nivel de confianza del 95% y una precisión absoluta del 10%. Con una razón no enfermos/enfermos de 8:1, se obtuvo un tamaño mínimo de 445 sujetos.

6.5 *Marco muestral*

Se utilizó como punto de partida la base de datos institucional generada automáticamente por el software de gestión de historias clínicas de la Fundación Santa fe de Bogotá, que exporta en formato Microsoft Excel versión 365, la información de todos los pacientes que consultaron durante el periodo definido. Se filtraron y eliminaron los pacientes menores de 18 años, los que se asignaron al servicio de ortopedia, los que se asignaron al servicio de obstetricia y aquellos que se les dio egreso por resolución de su patología antes de completar 24 horas en la institución.

6.6 *Diseño de la muestra*

Se realizó un estudio probabilístico con muestreo aleatorio simple sin reemplazamiento. Durante el periodo seleccionado el servicio de urgencias tuvo un total de 45.362 consultas en el área de triaje; de estos se excluyeron aquellos pacientes que fueron asignados a los servicios de ortopedia, pediatría y obstetricia, dejando un total de 26.629 pacientes. Posteriormente, se excluyeron aquellos casos con egreso antes de las 24 horas, obteniendo 8.415 pacientes potencialmente elegibles. De esta población de 8415 pacientes, registrados en una base de datos en Microsoft Excel versión 365 se seleccionó una muestra de 445 pacientes, usando la función `ALEATORIO(8415)+1`, se tomó la parte entera del número y se fueron seleccionando hasta completar el tamaño de la muestra. Por cada número elegido, se revisó la historia clínica correspondiente; si el paciente cumplía los criterios de inclusión y exclusión establecidos, era incluido en la muestra. En caso contrario, se seleccionaba un nuevo número y el proceso continuaba.

6.7 Variables de estudio

Tabla 1. Definición y Operacionalización de variables de estudio

Variable	Definición	Naturaleza	Escala	Unidades o categorías
DEPENDIENTES				
Mortalidad a 24 horas	Determinar si el paciente falleció en las primeras 24 horas posteriores al ingreso de urgencias	Cualitativa	Nominal dicotómica	<ul style="list-style-type: none"> ▪ Si ▪ No
Mortalidad al alta hospitalaria	Determinar si el paciente falleció antes del egreso hospitalario.	Cualitativa	Nominal dicotómica	<ul style="list-style-type: none"> ▪ Si ▪ No
INDEPENDIENTES				
Puntaje de TREWS	Puntaje de alerta temprana creado en 2019 que permite predecir la mortalidad intrahospitalaria en las primeras 24 horas de urgencias	Cualitativa	Ordinal	<ul style="list-style-type: none"> ▪ Puntaje del 0 - 20
Triage asignado al ingreso	Puntaje de triage (ESI) asignado al ingreso	Cuantitativa	Discreta	<ul style="list-style-type: none"> ▪ 1-5
VARIABLES INDEPENDIENTES PARA CALCULO DE ESCALA				
Presión arterial sistólica	Presión arterial sistólica del paciente al ingreso a urgencias	Cuantitativa	Discreta	<ul style="list-style-type: none"> ▪ mmHg
Frecuencia cardíaca	Frecuencia Cardíaca del paciente al ingreso a urgencias	Cuantitativa	Discreta	<ul style="list-style-type: none"> ▪ Latidos por minuto (lpm)
Frecuencia respiratoria	Frecuencia Respiratoria del paciente al ingreso a urgencias	Cuantitativa	Discreta	<ul style="list-style-type: none"> ▪ Respiraciones por minuto (rpm)
Temperatura	Temperatura del paciente al ingreso a urgencias	Cuantitativa	Continua	<ul style="list-style-type: none"> ▪ Grados centígrados (°C)
Nivel de consciencia	Nivel de consciencia del paciente al ingreso a urgencias	Cualitativa	Ordinal	<ul style="list-style-type: none"> ▪ Alerta (1) ▪ Verbal (2) ▪ Dolor (3) ▪ Ninguna (4)

Oxígeno suplementario	Determinar si el paciente requiere oxígeno suplementario al ingreso a urgencias	Cualitativa	Nominal dicotómica	<ul style="list-style-type: none"> ▪ Si ▪ No
VARIABLES DE CONFUSIÓN				
Sexo	Sexo del paciente	Cualitativa	Nominal	<ul style="list-style-type: none"> ▪ Masculino, ▪ Femenino
Edad	Número de años cumplidos	Cuantitativa	Discreta	<ul style="list-style-type: none"> ▪ Años cumplidos
Índice de Masa Corporal	División del peso sobre la talla elevada al cuadrado	Cuantitativa	Continua	<ul style="list-style-type: none"> ▪ Kg/m²
Índice de Comorbilidad de Charlson	Cálculo del índice de comorbilidad con antecedentes registrados al ingreso.	Cuantitativa	Discreta	<ul style="list-style-type: none"> ▪ 0-1 Punto Ausencia de comorbilidad ▪ 2 Puntos Comorbilidad baja ▪ > 3 puntos Comorbilidad alta.
DESCRIPTIVAS				
Diagnóstico al Alta	Diagnóstico de egreso asignado en el registro de la historia clínica	Cualitativa	Nominal	<ul style="list-style-type: none"> ▪ Trauma ▪ Infección ▪ Enfermedad Cardiovascular, ▪ Enfermedad Crónica no Transmisible ▪ Enfermedad Psiquiátrica ▪ Otras.
Requerimiento de UCI	Determinar si el paciente requirió estancia en UCI durante la hospitalización	Cualitativa	Nominal dicotómica	<ul style="list-style-type: none"> ▪ Si ▪ No

6.8 Técnicas, procedimientos e instrumentos de la recolección de datos

La recolección fue realizada por los investigadores, con el aval del respectivo comité de ética institucional, en estricto cumplimiento de los principios de confidencialidad. La información fue tomada a partir del registro en las historias clínicas electrónicas, extrayendo los datos de ingreso y los registros institucionales posteriores para documentar la mortalidad, el diagnóstico final y el desenlace clínico.

El puntaje TREWS se calculó de forma automatizada en Microsoft Excel versión 365 mediante una fórmula previamente programada en la base de datos, siguiendo estrictamente el algoritmo y parámetros originalmente descritos por Lee et al. (17); se comprobó que la fórmula estuviera correctamente programada introduciendo datos simulados antes del inicio de la recolección de datos, comparando el cálculo manual vs el cálculo realizado por el programa. Las identificaciones registradas fueron codificadas para la protección de datos e identidad de los participantes. Esta base será almacenada en una carpeta en el sistema OneDrive compartida únicamente con el grupo investigador, con un tiempo de almacenamiento máximo de 5 años.

6.9 Plan de procesamiento de muestras biológicas

No se realizó la toma de ninguna muestra biológica.

6.10 Análisis de datos

Se hizo un análisis descriptivo de las variables cualitativas, evaluando la distribución de frecuencias. Para las variables continuas, se calcularon medidas de tendencia central y de dispersión, según el tipo de distribución. Luego, se utilizó la prueba de Shapiro-Wilk para determinar la normalidad de las distribuciones y según el resultado, aquellas con distribución normal se les aplicó media y desviación estándar y para las distribuciones no normales se utilizaron la mediana y los rangos intercuartílicos (IQ).

Para determinar la capacidad predictiva del puntaje de TREWS y el Triaje en la mortalidad medida a 24 horas y al alta se realizaron análisis de valores predictivos positivos y negativos, curvas ROC, medidas características operativas de sensibilidad, especificidad y razones de verosimilitud positiva y negativa con sus respectivos intervalos de confianza del 95%. Así mismo se calculó el puntaje TREWS y se obtuvo el punto de corte óptimo con la mayor suma de la sensibilidad y especificidad. Adicionalmente, se construyeron curvas de calibración con el fin de evaluar la concordancia entre las probabilidades predichas y los valores observados en la muestra. Finalmente se realizaron análisis con curvas ROC para comparar el puntaje TREWS y el Triaje en su capacidad de predecir mortalidad.

6.11 Alcances y límites de la investigación

Se identifican principalmente los siguientes potenciales sesgos y limitaciones:

- **Sesgo de Información:** Durante la recolección de datos se observó que algunos pacientes habían consultado en más de una ocasión al servicio de urgencias dentro del periodo de estudio, generando múltiples registros en la base de datos. En estos casos, se optó por conservar únicamente el primer ingreso registrado, excluyendo los duplicados para evitar sesgos por sobrerrepresentación de individuos. Cada caso excluido fue reemplazado siguiendo el mismo procedimiento de selección previamente descrito, garantizando así la integridad del muestreo.
- **Limitaciones:** Una limitación importante de este estudio es su validez externa, ya que la población analizada es altamente específica de un único centro. Esto restringe la capacidad de generalizar nuestros hallazgos a otras poblaciones o contextos clínicos.

Durante el estudio los participantes no tuvieron ninguna clase de beneficio. Mas aun, se espera que estos hallazgos puedan ser el fundamento para justificar y estandarizar el uso del puntaje de TREWS en los servicios de urgencias; permitiendo identificar de forma temprana a aquellos pacientes con alto riesgo de deterioro que requieren una atención oportuna.

7 Aspectos éticos

El estudio se realizó dentro de los principios éticos para las investigaciones médicas en seres humanos según la Declaración de Helsinki - 64ª Asamblea General, Fortaleza, Brasil, octubre 2013 (38). Se tuvo en cuenta las regulaciones locales del Ministerio de Salud de Colombia Resolución 8430 de 1993 en lo concerniente al Capítulo I “De los aspectos éticos de la investigación en seres humanos” (39).

La presente investigación es clasificada dentro de la categoría de riesgo mínimo puesto que se tomaron datos de historias clínicas, información no sensible y no habrá intervención en el manejo estos pacientes. Se limitará el acceso de los instrumentos de investigación únicamente a los investigadores según Artículo 8 de la Resolución 8430 de 1993 del Ministerio de Salud (39).

Será responsabilidad de los investigadores el guardar con absoluta reserva la información contenida en las historias clínicas y a cumplir con la normatividad vigente en cuanto al manejo de esta, reglamentados en los siguientes: Ley 100 de 1993, Ley 23 de 1981, Decreto 3380 de 1981, Resolución

008430 de 1993 y Decreto 1995 de 1999.

Todos los integrantes del grupo de investigación estarán prestos a dar información, previa solicitud razonable, sobre el estudio a entes organizados, aprobados e interesados en conocerlo siempre y cuando sean de índole académica y científica, preservando la exactitud de los resultados y haciendo referencia a datos globales y no a pacientes en particular.

El estudio se llevó a cabo en la Fundación Santa Fe de Bogotá, Colombia, y recibió la aprobación del comité de ética e investigación de la institución con el número de aprobación CCEI-15338-2023. Se mantendrá absoluta confidencialidad y se preservará el buen nombre institucional profesional.

Todos los pacientes ingresados en el servicio de urgencias firmaron un formulario de consentimiento genérico, en el que daban su consentimiento informado por escrito para utilizar y publicar sus historias clínicas con fines académicos y de investigación.

El estudio se realizará con un manejo estadístico imparcial y responsable.

No existe ningún conflicto de interés por parte de los autores del estudio que deba declararse.

Todos los autores han contribuido significativamente al trabajo presentado, ya sea en la concepción, el diseño del estudio, la ejecución, la adquisición de datos, el análisis, la interpretación o todas estas áreas. Participaron en la redacción, revisión o revisión crítica del artículo; dieron su aprobación final a la versión que se publicará; acordaron la revista a la que se ha enviado el artículo; y aceptaron ser responsables de todos los aspectos del trabajo.

8 Resultados

La cohorte de estudio la conformaron 445 pacientes tras verificar el cumplimiento de los criterios de inclusión y exclusión (Figura 1).

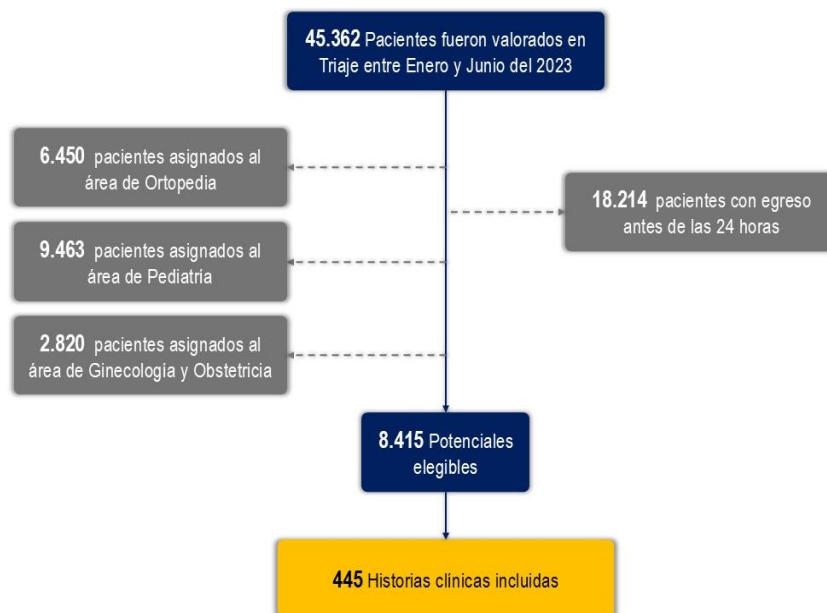


Figura 1. Selección de pacientes en la Fundación Santa Fe de Bogotá de enero a junio de 2023.

La distribución de los pacientes dentro de las distintas clasificaciones del triaje puede observarse en la Figura 2.

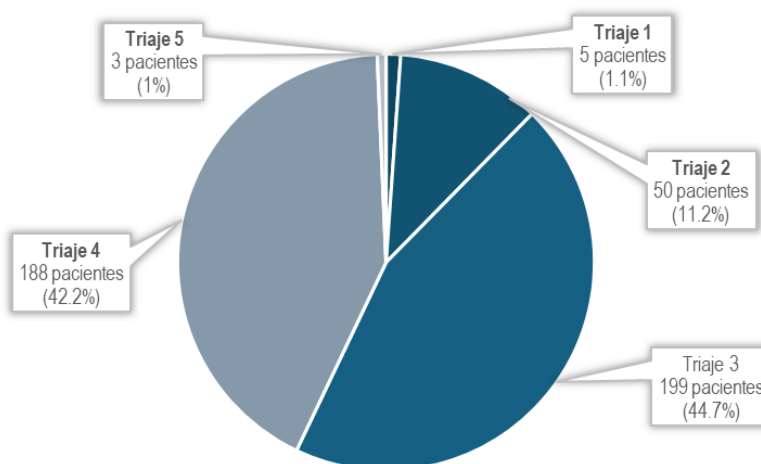


Figura 2. Distribución de la cohorte según la clasificación de Triaje

8.1 Características demográficas y clínicas

En cuanto a las características de los pacientes, el 51.6% eran mujeres, con una mediana de edad de 67 años, el diagnóstico principal al alta fue en un 35.7% relacionado con infecciones y en un 23.5% relacionado con descompensación de enfermedades crónicas; 48 pacientes requirieron traslado a unidad de cuidados intensivos representando el 10.7%. (Tabla 2).

Tabla 2. Características demográficas y clínicas de pacientes que ingresaron al servicio de urgencias de la Fundación Santafé de Bogotá, 2023

Variable	No sobreviviente (%)	Sobreviviente (%)	OR	P-valor
Total (445)	13 (2.9) * 21 (4.71)	432 (97.1) * 424 (95.2)		
Sexo (Masculino)	9 (69.2)	201 (46.5)	0.39 (0.10 - 1.25)	0.182
Edad (Mediana (IQ))	78 (55-89)	59 (42-75)		0.095
Diagnostico (%)				
Trauma (4)	1 (5)	17 (4)	NA	0.097
Infección (35.7)	3 (15)	156 (36.7)	4.82 (0.14-62.48)	
Cardiovascular (8.5)	3 (15)	35 (8.2)	1.11(0.03-14.72)	
Enfermedad crónica (23.5)	9 (45)	98 (23.1)	1.09(0.04-7.28)	
Otras (27.6)	4 (20)	119 (28)	3.72(0.11-48.30)	
Presión sistólica (Mediana (IQ))	99 (90 - 124)	124.5 (109 - 137)		<0.001
Presión diastólica (Mediana (IQ))	62(48-81)	75 (66-84)		0.123
Frecuencia cardiaca (Mediana (IQ))	99 (73-117)	85 (73-98.2)		0.207
Frecuencia respiratoria (Mediana (IQ))	23 (22-25)	18 (17-20)		0.007
Temperatura (Mediana (IQ))	36.4 (36-36.5)	36.5 (36.2-36.7)		0.158
Escala de conciencia (%)				
1	13 (65)	406 (95.5)	1.0 (NA)	<0.001
2	3 (15)	13 (3.1)	0.11 (0.02-0.91)	
3	2 (10)	4 (0.9)	0.03 (0.00-0.31)	
4	2 (10)	2 (0.5)	0.01 (0.00-0.19)	
Uso de oxígeno (%)	10 (19.6)	41 (80.3)	30.22 (8.67-145)	0.084
Índice de Charlson (Mediana (IQ))	5 (2-8)	2 (0-4)		0.012
Necesidad de UCI (NO) (%)	12 (3)	385 (96.9)	0.77 (0.03 – 4.08)	<0.001
Triage (%)				
1 (1.1)	1 (5)	4 (0.9)		<0.001
2 (11.2)	10 (50)	40 (9.4)		

3 (44.7)	8 (40)	191 (44.9)	
4 (42.2)	1 (5)	187 (44)	
5 (0.7)	0 (0)	3 (0.7)	
Triage (Mediana (IQ))	2 (2-3)	3 (3-4)	<0.001
TREWS (Mediana (IQ))	9 (8 - 12)	3 (2-4)	<0.001

* Sobrevida a 24 horas. IQ: Rangos intercuartílicos. OR: Odds Ratio

8.2 Puntaje de TREWS en la población estudiada

La Figura 3 muestra la distribución de la frecuencia del puntaje TREWS en la población analizada. Se observa una tendencia claramente asimétrica a la derecha, con una mayor concentración de pacientes en los puntajes bajos y una disminución progresiva hacia los valores más altos. El pico de frecuencia se presentó en el puntaje 3, con 122 pacientes, seguido del puntaje 2 con 88 pacientes y del puntaje 4 con 61 pacientes. En los extremos inferiores se registraron 34 pacientes con puntaje 0 y 28 pacientes con puntaje 1, mientras que a partir del puntaje 5 la frecuencia comenzó a descender de manera constante: 34 pacientes en el puntaje 5, 22 en el 6, 13 en el 7, y cifras menores entre los puntajes 8 y 11 (10 pacientes en los puntajes 8, 9 y 10; y 7 pacientes en el puntaje 11). En los puntajes superiores la frecuencia fue casi nula, con solo 1 a 2 pacientes entre los puntajes 12 y 15, y ausencia total en los puntajes de 16 a 20. Estos resultados evidencian que la mayoría de los pacientes evaluados presentaron un puntaje TREWS bajo, mientras que los valores altos fueron poco frecuentes en la muestra.

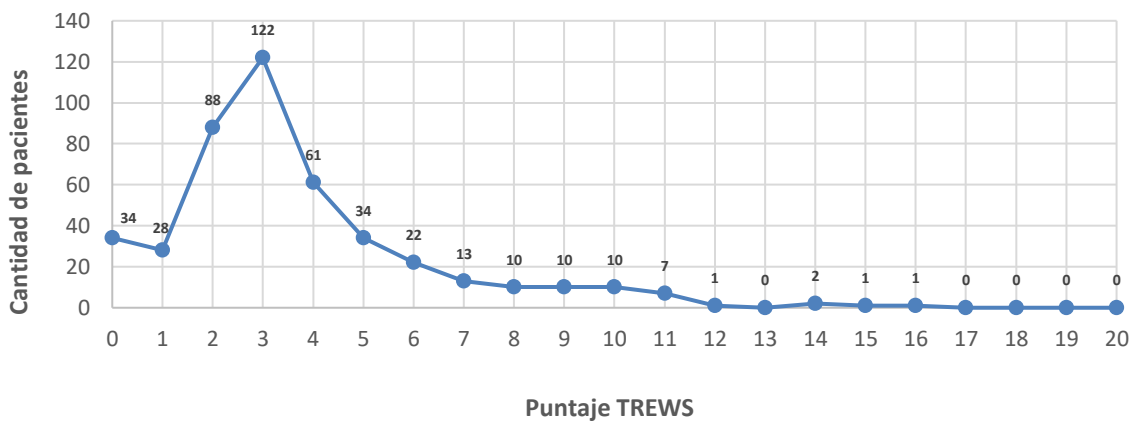


Figura 3. Frecuencia del puntaje de TREWS calculado en la muestra

8.3 Mortalidad en la cohorte de estudio

La mortalidad intrahospitalaria en la muestra fue del 4,7% (21 pacientes), de los cuales el 2,9% (13 pacientes) fallecieron en las primeras 24 horas tras el ingreso a urgencias. No se encontraron diferencias significativas en la distribución por sexo (59,2% vs. 46,5% masculino; $p = 0,182$). Los pacientes que no sobrevivieron eran de mayor edad (78 [58–89] vs. 59 [42–75] años), aunque esta diferencia no alcanzó significancia estadística ($p = 0,095$). En cuanto al diagnóstico inicial, en el grupo de los fallecidos, fueron más frecuentes las infecciones (38,5%) y las enfermedades crónicas (23,5%), mientras que en los sobrevivientes predominaron los cuadros respiratorios (31,5%) y cardiovasculares (8,3%), sin diferencias estadísticamente significativas ($p = 0,097$).

En el análisis de parámetros clínicos, los pacientes que fallecieron ingresaron con una presión arterial sistólica significativamente más baja (99 [90–124] vs. 124,5 [109–137] mmHg; $p < 0,001$), mientras que no hubo diferencias relevantes en la presión diastólica (62 [48–81] vs. 75 [60–84] mmHg; $p = 0,123$). Así mismo, presentaron frecuencias cardíacas ligeramente más altas (99 [73–117] vs. 85 [73–98] lpm; $p = 0,207$) y una frecuencia respiratoria significativamente mayor (23 [22–25] vs. 18 [17–20] rpm; $p = 0,007$). La temperatura corporal y el estado de conciencia no mostraron diferencias significativas entre los grupos.

Respecto al soporte respiratorio, se observó que un menor porcentaje de no sobrevivientes requirió oxígeno al ingreso (10% vs. 41%; $p = 0,084$). El índice de Charlson fue significativamente más alto en este grupo (5 [2–8] vs. 2 [0–4]; $p = 0,012$), lo que refleja una mayor carga de comorbilidad y en cuanto a la necesidad de ingreso a UCI fue más frecuente entre los fallecidos (30% vs. 0,7%; $p < 0,001$).

Finalmente, el puntaje TREWS fue significativamente más alto en los pacientes fallecidos en comparación con los sobrevivientes (9 [8–12] vs. 3 [2–4]; $p < 0,001$).

8.4 Punto de corte óptimo del puntaje de TREWS a 24 horas

Al establecer el punto de corte óptimo para el puntaje TREWS, definido como aquel con la mayor suma de sensibilidad y especificidad, se observó que los valores de 5,5 y 7 ofrecían el mismo resultado global en dicha suma. No obstante, el punto de corte de 5,5 se asoció a un perfil con mayor sensibilidad, mientras que el de 7 mostró un mayor peso en la especificidad. Teniendo en cuenta, además, que en la literatura se ha descrito un umbral de > 4 puntos, se realizaron análisis de medidas operativas en los diferentes valores de corte para evaluar de manera integral el desempeño del puntaje en la predicción de mortalidad.

Tabla 3. Sensibilidad, especificidad, valores predictivos y razones de verosimilitud con sus respectivos intervalos de confianza del 95% para el puntaje de TREWS a diferentes puntos de corte en la mortalidad medida a las 24 horas.

Umbral	Sensibilidad %	Especificidad %	VPP %	VPN %	LR+	LR-
4.0	84.6 (61.1-100)	63.2 (58.5-67.8)	6.47 (2.48-10.5)	99.3 (98.1-100)	2.30 (1.77-2.9)	0.24 (0.07-0.87)
5.0	84.6 (61.1-100)	77.3 (73.2-81.4)	10.1 (3.98-16.2)	99.4 (98.4-100)	3.73 (2.79-5.0)	0.20 (0.06-0.71)
5.5	84.6 (61.1-100)	85.2 (81.7-88.6)	14.7 (5.99-23.3)	99.4 (98.6-100)	5.71 (4.13-7.9)	0.18 (0.05-0.65)
6.0	84.6 (61.1-100)	85.2 (81.7-88.6)	14.7 (5.99-23.3)	99.4 (98.6-100)	5.71 (4.13-7.9)	0.18 (0.05-0.65)
7.0	76.9 (50.1 – 100)	90.1 (87.1-92.9)	18.9 (7.39-30.3)	99.2 (98.2-100)	7.73 (5.1-11.7)	0.26 (0.09-0.69)

VPN: Valor predictivo negativo, VPP: Valor predictivo positivo, LR+: razón de probabilidad positiva, LR-: razón de probabilidad negativa

En el comportamiento del puntaje de TREWS en relación con la mortalidad a 24 horas para distintos puntos de corte, se observa que la sensibilidad se mantiene alta y constante en los umbrales de 4 a 6 (84,6%), con un leve descenso en el punto de corte de 7 (76,9%). La especificidad, en cambio, aumenta progresivamente desde 63,2% en el umbral 4 hasta alcanzar 90,1% en el umbral 7, lo cual refleja un mejor desempeño para discriminar a los pacientes sin mortalidad a medida que se eleva el punto de corte. Los valores predictivos positivos fueron bajos en todos los escenarios, aunque tienden a incrementarse desde 6,5% en el umbral 4 hasta 18,9% en el umbral 7, mientras que los valores predictivos negativos se mantuvieron consistentemente altos ($\geq 99\%$), lo que sugiere que la escala es útil principalmente para descartar mortalidad. En cuanto a las razones de verosimilitud, el LR+ se incrementó de forma marcada con puntos de corte más altos (de 2,30 a 7,73), indicando mayor utilidad para confirmar el desenlace en los valores extremos; en contraste, el LR- se mantuvo bajo en todos los casos (0,18 - 0,26), apoyando la capacidad de la herramienta para excluir el riesgo de mortalidad (Tabla 3).

8.5 Punto de corte óptimo del puntaje de TREWS al alta

En relación con el comportamiento de los diferentes puntos de corte estudiados del puntaje de TREWS en relación con la mortalidad a alta, se observa que la sensibilidad disminuye de forma progresiva a medida que aumenta el umbral, pasando de 80,9% en el punto de corte de 4 a 71,4% en el umbral de 7. En contraste, la especificidad mejora de forma consistente, aumentando de 63,9% a 91,0% en los mismos umbrales. Los valores predictivos positivos, aunque bajos en general, se incrementan desde 10,0% en el umbral de 4 hasta 28,3% en el de 7, lo que indica mayor probabilidad de mortalidad en quienes superan los puntos de corte más altos. El valor

predictivo negativo se mantiene estable y muy elevado en todos los escenarios ($\geq 97\%$). Respecto a las razones de verosimilitud, el LR+ se incrementa de forma importante a medida que se eleva el punto de corte (2,24 en umbral 4 frente a 7,97 en umbral 7), lo que indica una mayor capacidad para confirmar el desenlace en umbrales más altos. Por el contrario, el LR- se mantuvo bajo y estable en todos los casos (0,28–0,31), lo que respalda la utilidad del TREWS para excluir el desenlace cuando el resultado es negativo (Tabla 4).

Tabla 4. Sensibilidad, especificidad, valores predictivos y razones de verosimilitud con sus respectivos intervalos de confianza del 95% para el puntaje de TREWS a diferentes puntos de corte en la mortalidad medida al alta.

Umbral	Sensibilidad %	Especificidad %	VPP %	VPN %	LR+	LR-
4.0	80.9 (61.7-100)	63.9 (59.2-68.6)	10.0 (5.2-14.8)	98.5 (96.9-100)	2.24 (1.76-2.3)	0.30 (0.12-0.72)
5.0	76.2 (55.6-96.8)	78.1 (74.0-82.1)	14.7 (7.58-21.8)	98.5 (97.0-100)	3.47 (2.58-4.5)	0.30 (0.14-0.66)
5.5	76.2 (55.6-96.8)	86.1 (82.7-89.5)	21.3 (11.4-31.2)	98.6 (97.3-100)	5.48 (3.91-7.7)	0.28 (0.13-0.59)
6.0	76.2 (55.6-96.8)	86.1 (82.7-89.5)	21.3 (11.4-31.2)	98.6 (97.3-100)	5.48 (3.91-7.7)	0.28 (0.13-0.59)
7.0	71.4 (49.7-93.1)	91.0 (88.2-93.9)	28.3 (15.2-41.3)	98.5 (97.1 -100)	7.97 (5.31-12)	0.31 (0.16-0.62)

VPN: Valor predictivo negativo, VPP: Valor predictivo positivo, LR+: razón de probabilidad positiva, LR-: razón de probabilidad negativa

8.6 Capacidad predictiva del puntaje TREWS para la mortalidad medida a 24 horas y al alta

El análisis de las curvas ROC muestra que el puntaje de TREWS tiene una alta capacidad predictiva tanto la mortalidad en 24 horas con un AUC de 0,856 (IC 95%: 0,708 - 1.00) como la mortalidad previa al egreso con un AUC de 0,861 (IC 95%: 0,755 - 0,967) (Figura 4).

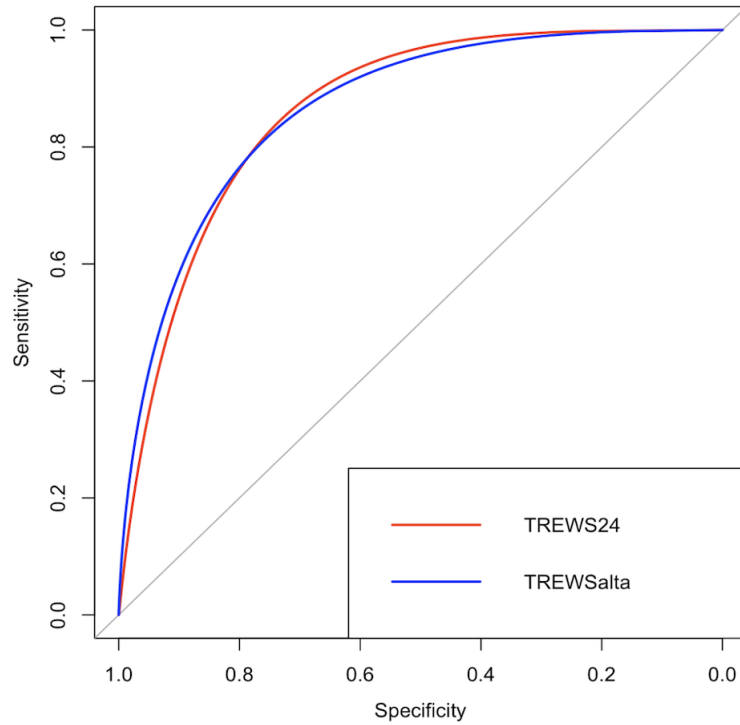
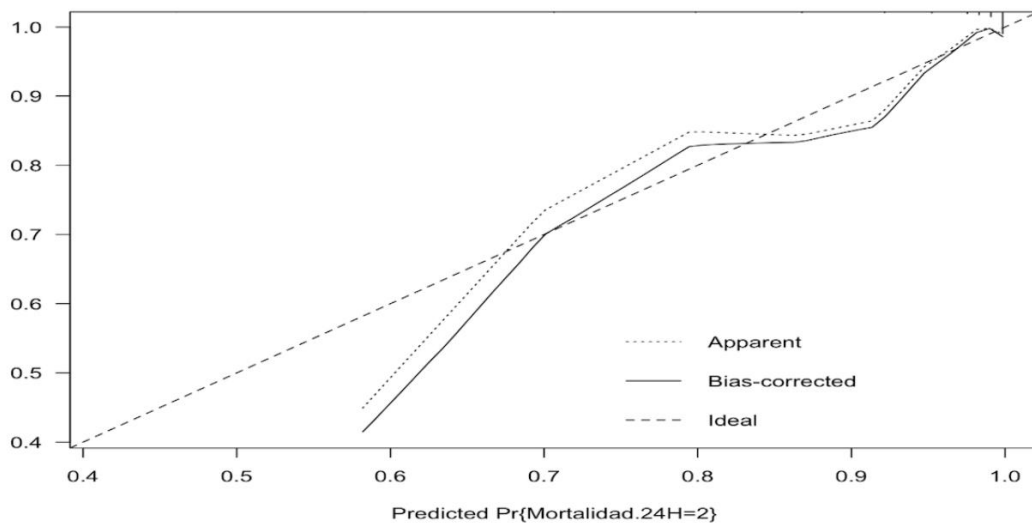


Figura 4. Curvas ROC para la capacidad predictiva del puntaje TREWS para mortalidad a 24 horas (Línea roja) y mortalidad al alta (Línea azul). (AUC a las 24 horas: 0,856, IC 95%: 0,708-1.00, AUC al alta: 0,861, 95% CI: 0,755-0,967. $P < 0.001$)

Adicionalmente, se trazaron curvas de calibración para ilustrar las predicciones de muerte de la puntuación tanto al alta como a las 24 horas (Figura 5).



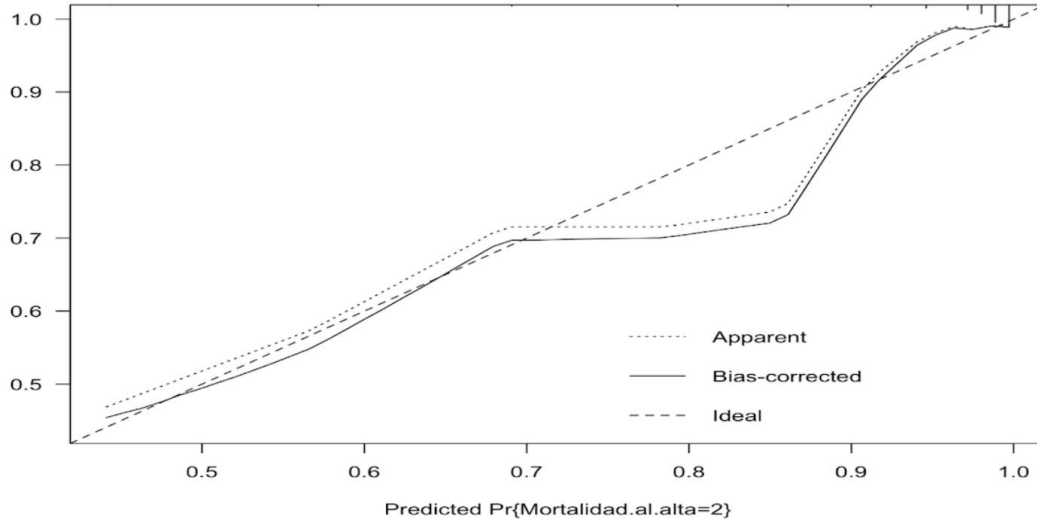


Figura 5. Gráficos de calibración de las escalas evaluadas

8.7 Comparación de la capacidad de predicción del puntaje TREWS vs Triage

Por último, se evaluó la predicción y características operativas del puntaje TREWS y el nivel de triaje clínico (de 1 menor gravedad a 5 mayor gravedad según el Triage ESI usado de forma estándar en la Fundación Santa Fe de Bogotá) asignado al ingreso. El análisis mostro que por cada aumento de 1 unidad en el nivel de triaje (es decir, menor urgencia clínica), el TREWS disminuye en promedio 0.11 puntos, y cuando Triage es 0, el valor estimado de TREWS es 3.8. El valor de la pendiente es negativo por lo que se interpreta que existe una asociación inversamente proporcional, entre las dos variables. Así mismo, aunque el modelo fue significativo estadísticamente ($p < 0.001$), el R^2 bajo (0.17) sugiere que el nivel de triaje explica solo una fracción moderada de la variabilidad del TREWS, indicando que otros factores no medidos podrían influir en el puntaje.

Finalmente, al comparar la capacidad predictiva del puntaje TREWS con el triaje clínico, se encontró que TREWS mostró un AUC ligeramente superior tanto para mortalidad a las 24 horas como para mortalidad al alta, aunque estas diferencias no fueron estadísticamente significativas ($p = 0.556$ y $p = 0.38$, respectivamente). (Figura 6 y Figura 7)

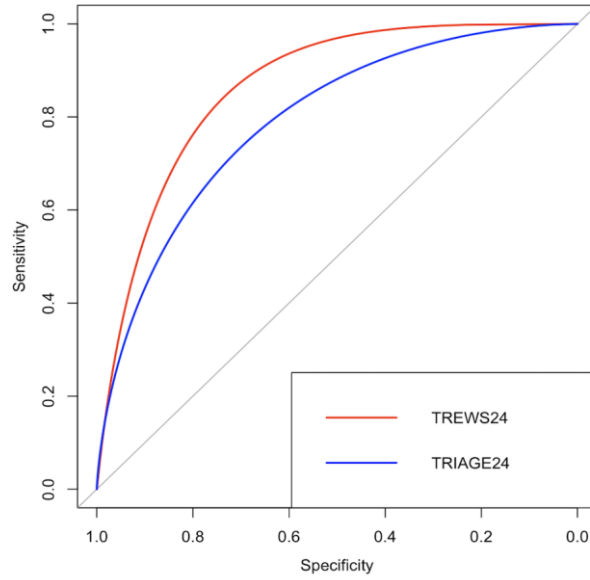


Figura 6. Curvas ROC comparativa para la mortalidad a 24 horas entre el puntaje TREWS (Línea roja) y el triage (Línea azul).

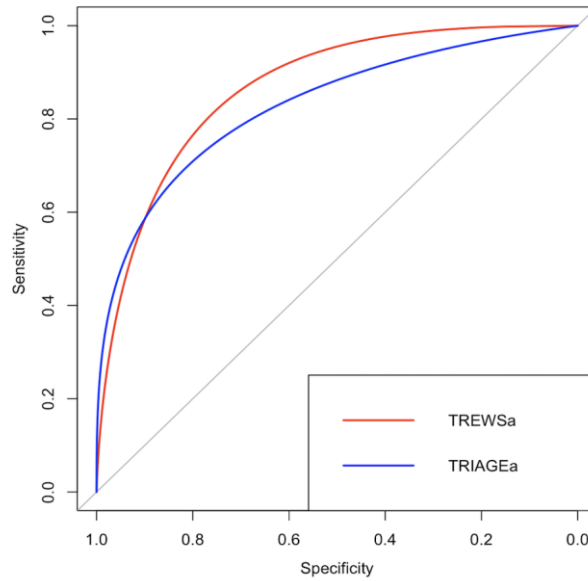


Figura 7. Curvas ROC comparativa para la mortalidad al alta entre el puntaje TREWS (Línea roja) y el triage (Línea azul).

9 Discusión

Se optó por un diseño analítico de cohorte retrospectiva para este estudio de predicción diagnóstica, ya que ofrece ventajas prácticas en términos de reducción de tiempo y costos, y facilita el acceso a una población amplia en la que el desenlace de interés —la mortalidad a 24 horas— ocurre con baja frecuencia (40). Este enfoque resulta adecuado para evaluar, en nuestra población de pacientes mayores de 18 años que ingresaron al servicio de urgencias de la Fundación Santa Fe de Bogotá (centro de alta complejidad) entre enero y junio de 2023, una escala desarrollada en un contexto distinto, explorando así su utilidad local como base para estudios prospectivos posteriores. El muestreo se efectuó mediante un diseño aleatorio simple sin reemplazo, con el objetivo de preservar la independencia de las observaciones y minimizar el sesgo de selección. Para evitar la sobrerrepresentación de usuarios frecuentes de urgencias, la unidad de análisis se definió en el nivel del paciente, reteniéndose únicamente la primera consulta registrada y excluyéndose las subsecuentes. En los casos en que se identificaron entradas duplicadas, estas fueron reemplazadas por otro paciente seleccionado mediante el mismo algoritmo de aleatorización establecido en el protocolo.

Se extrajo la información de la historia clínica electrónica del episodio de ingreso, incluida la mortalidad y se excluyeron los pacientes con estancia inferior a 24 horas, al no ser posible medir el desenlace objetivo. El impacto de estas pérdidas se considera limitado: en estudios poblacionales, la mortalidad temprana en los primeros siete días de los pacientes que salen de urgencias antes de 24 horas se ha estimado en torno al 0,12% (41). Para controlar errores de digitación, se implementó un cálculo automático del puntaje TREWS en Excel y se verificó su correcto funcionamiento mediante simulación. Durante la recolección de datos no hubo cambios en el protocolo que alteraran los pacientes incluidos en el análisis.

Los resultados de este estudio demuestran que la puntuación TREWS tiene un buen desempeño en la predicción de la mortalidad tanto a las 24 horas, con un área bajo la curva (AUC) de 0,856 (IC 95%: 0,708-1.00, $p < 0.001$), así como en la predicción de la mortalidad intrahospitalaria, con un AUC de 0,861 (IC 95%: 0,755-0,967 $p < 0.001$). Estos datos reflejan una buena capacidad predictiva indicando que el puntaje de TREWS puede identificar adecuadamente a los pacientes con mayor riesgo de fallecer en ambos momentos del seguimiento realizado, sustentado en la similitud de ambas curvas ROC que sugiere una consistencia en el poder predictivo del modelo, independientemente del horizonte temporal.

Específicamente en el estudio de la mortalidad a 24 horas, se encontró que los puntos de corte de 5,5 y 7 ofrecieron los mejores desempeños, alcanzando la mayor suma entre sensibilidad y especificidad.

No obstante, el umbral de 5,5 mostró un mejor balance para identificar tempranamente a los pacientes con riesgo de mortalidad con una sensibilidad de 85,7% y una especificidad de 86,7%. Para este mismo punto de corte, el elevado VPN respalda la utilidad del puntaje TREWS como una herramienta confiable para descartar el riesgo de mortalidad temprana. En contraste, el bajo VPP observado probablemente se relaciona con la baja prevalencia de muerte súbita en el servicio de urgencias, más que con una limitación del puntaje.

En relación con la mortalidad al alta, los umbrales de 5,5 y 6 fueron más consistentes, manteniendo una sensibilidad de 76,2% con una especificidad de 86,1%. Estos valores, combinados con un LR+ mayor a 5 y un VPN cercano al 98%, consolidan a este umbral como el mejor para la predicción de este desenlace en el seguimiento intrahospitalario.

El rendimiento del puntaje TREWS tanto a 24 horas como al alta hospitalaria muestra que el punto de corte de 5,5–6 ofrece el mejor equilibrio entre sensibilidad y especificidad, con valores superiores al 75% en ambos parámetros y un VPN consistentemente alto. Estos hallazgos respaldan la elección de este umbral como el más confiable para predecir el riesgo de mortalidad en escenarios clínicos. Sin embargo, dado que 5,5 no corresponde a un valor posible dentro de la escala, y considerando que en 5 se mantiene la sensibilidad, pero se disminuye la especificidad, se sugiere 6 como el punto de corte más adecuado para la práctica clínica, puesto que conserva un balance apropiado entre identificar oportunamente a los pacientes en riesgo y minimizar la pérdida de capacidad para descartar desenlaces adversos.

En este estudio también se comparó el desempeño del puntaje TREWS con el sistema de triaje clínico ESI. Aunque ambas herramientas mostraron capacidades predictivas similares para mortalidad a 24 horas y al alta, TREWS presentó un rendimiento diagnóstico superior, aunque no estadísticamente significativo. Esta diferencia, aunque sutil, sugiere que el TREWS podría aportar valor adicional como herramienta complementaria al triaje tradicional en contextos de alta complejidad o en pacientes con triajes intermedios, donde la estratificación del riesgo puede ser más ambigua. Así mismo, su fuerte correlación inversa negativa indica que ambas herramientas no son redundantes, sino que evalúan dimensiones diferentes del estado clínico del paciente.

Los sistemas de alerta temprana han ganado lugar en el ámbito hospitalario por su capacidad de detectar el deterioro clínico en los pacientes mediante el seguimiento de sus signos vitales (6). Variados puntajes como el NEWS, el NEWS 2, el RAPS y el SEWS han sido estudiados para determinar cuál es el más preciso para predecir el riesgo de muerte o ingreso a la UCI en las primeras 24 horas desde el ingreso a un hospital; encontrando que el NEWS es el más preciso con un AUC de

0.904 (14). No obstante, el punto de corte del puntaje de NEWS y su consecuente AUC varía ampliamente dependiendo del estudio, cambiando de puntos de corte entre 3, 4 y hasta 6 con AUC que varía desde el 0.881 hasta el 0.901 (14,16,42).

Estas variaciones en la precisión del puntaje NEWS pueden deberse a que su diseño originalmente fue pensado para la identificación temprana de pacientes con riesgo de deterioro agudo durante la hospitalización, no desde su ingreso a urgencias y principalmente, el desenlace primario evaluado fue el requerimiento de UCI (43,44). Actualmente el Colegio Real de Médicos en su documento de estandarización de la evaluación de la gravedad de las enfermedades agudas en el sistema nacional de salud menciona que el principal uso del puntaje de NEWS es el seguimiento de los pacientes durante su estancia hospitalaria y que ‘podría’ usarse en el abordaje inicial de pacientes que ingresan a urgencias (45).

En este contexto, en 2019 se desarrolló el puntaje TREWS, derivado del NEWS, como el primer y único sistema de alerta temprana diseñado específicamente para predecir la mortalidad a 24 horas desde el primer contacto del paciente con el hospital, es decir, desde la sala de triaje (17). Este es un instrumento diseñado con el objetivo de discriminar y categorizar el riesgo de los pacientes, permitiendo una identificación temprana de aquellos con mayor probabilidad de deterioro. Su utilidad radica en que, al aplicarse oportunamente facilita la toma de decisiones clínicas y optimiza recursos asistenciales (17). En este sentido y hasta donde se tiene conocimiento, este estudio representa la primera aplicación del puntaje TREWS en una población adulta en Colombia, lo que lo convierte en una contribución novedosa que aporta conocimiento y evidencia local sobre su utilidad clínica.

Al contrastar el desempeño del puntaje NEWS reportado en la literatura con la capacidad predictiva del puntaje TREWS identificada en este estudio, se observa que, para la mortalidad a 24 horas, el NEWS con un punto de corte recomendado ≥ 5 alcanza una sensibilidad entre 71–77% y una especificidad de 60–66%, con un valor predictivo negativo elevado ($>98\%$), pero un valor predictivo positivo limitado ($<10\%$) (14,46). En comparación, los resultados de este estudio demuestran que el puntaje TREWS con un punto de corte de 5.5, mostró un rendimiento superior, con sensibilidad del 84.6% (IC 95% 61.1-100%) y especificidad del 85,2% (IC 95% 81.7-88.6%), manteniendo igualmente un VPN muy alto (99.4%, IC 95% 98.6-100%) pero acompañado de un VPP más favorable (14.7%, IC 95% 5.99-23.3%). Esto sugiere que, aunque ambas escalas son herramientas útiles para estratificación temprana del riesgo, el TREWS ofrece un mejor balance entre sensibilidad y especificidad, lo que lo posiciona como una alternativa más precisa para identificar pacientes con riesgo de mortalidad en las primeras 24 horas.

En la mortalidad al alta, el puntaje NEWS en una revisión sistémica de 2021 con 107,008 pacientes presento un rendimiento moderado con una sensibilidad de 71%, una especificidad de 60% y un AUC de 0.70, acompañado de un VPP limitado (8–19%) pero un VPN alto (97–99%) (47). En comparación la escala TREWS muestra un desempeño más sólido, con un punto de corte de 5.5 alcanzó un AUC de 0.861 con una sensibilidad de 76.2% (IC 95% 55.6-96.8%) y especificidad de 86.1% (IC 95% 82.7-89.5%), además de mejorar el VPP (21.3%, IC 95% 82.7-89.5%) manteniendo un VPN igualmente elevado (98.6%, IC 95% 97.3-100%).

Cabe destacar, que si bien Lee y sus colaboradores, en 2019, en el estudio original de creación del puntaje TREWS recomiendan un punto de corte de 4 para la predicción de mortalidad (17); este estudio encontró que este umbral mostró un rendimiento subóptimo, con buena sensibilidad, pero a expensas de una baja especificidad, lo que limita su utilidad clínica al generar un número elevado de falsos positivos. En contraste, los puntos de corte de 5.5 y 6 alcanzaron un balance mucho más favorable tanto a 24 horas como al alta, con sensibilidades y especificidades superiores al 75% y un VPN cercano al 99%.

Asimismo, aunque comparables, los hallazgos de este estudio en relación con la curva ROC para la predicción de mortalidad a 24 horas (AUC 0.856, IC 95% 0.708–1.00; $p < 0.001$) resultan inferiores a los reportados por Lee et al., donde la puntuación TREWS alcanzó un AUC de 0,906 (IC 95%: 0.903–0.908) (17) . Esta diferencia, podría explicarse a partir de las divergencias demográficas que afectan las características poblacionales, los perfiles de cuidado en urgencias, la heterogeneidad en las estrategias de atención y la carga de comorbilidades de los pacientes en nuestro estudio. Hipótesis que se ve respaldada por características relevantes entre los pacientes fallecidos como una mediana de edad significativamente mayor, así como por un mayor índice de comorbilidad y alteraciones relevantes en parámetros fisiológicos al ingreso. No obstante, es relevante destacar que la puntuación TREWS ha mostrado resultados variables y versátiles en su aplicación y utilidad clínica en poblaciones específicas, como en pacientes con COVID-19, con resultados igualmente prometedores (48).

La principal fortaleza de este estudio radica en que constituye la primera aplicación del puntaje TREWS en el área de triaje de un servicio de urgencias en un país de ingresos bajos y medios como Colombia, lo que lo convierte en un punto de partida valioso para futuros estudios multicéntricos en contextos similares. Sin embargo, el análisis se realizó en un único hospital de alta complejidad, lo que limita la generalización de los resultados a instituciones con condiciones estructurales, operativas o epidemiológicas distintas. En este sentido, la inclusión de una muestra mayor y más heterogénea en investigaciones posteriores permitiría no solo fortalecer la validez externa y reducir la amplitud de

los intervalos de confianza de las características operativas, sino también mejorar la precisión del puntaje sin modificar sustancialmente el valor actual de los hallazgos. De hecho, las curvas de calibración mostraron dificultades en la estimación del riesgo en valores extremos, con tendencia a subestimar en puntajes bajos y sobreestimar en puntajes altos, lo que podría estar asociado a la homogeneidad de la muestra. Este hallazgo refuerza la necesidad de estudios con poblaciones más diversas que contribuyan a un ajuste más robusto.

En resumen, un puntaje TREWS > 6 se asocia con un incremento significativo del riesgo de mortalidad; sin embargo, su limitada capacidad confirmatoria obliga a interpretar con cautela los resultados positivos. Aunque el TREWS demuestra una notable solidez como herramienta de exclusión, su rendimiento predictivo podría potenciarse mediante su integración con otras escalas clínicas o con el juicio profesional, sentando las bases para la construcción de modelos combinados de riesgo y protocolos de toma de decisiones en urgencias.

10 Conclusión

En conclusión, este estudio demuestra que el puntaje TREWS presenta un desempeño sólido y consistente para la predicción de mortalidad tanto a 24 horas como al alta hospitalaria en pacientes atendidos en un servicio de urgencias de alta complejidad, superando en varios parámetros al NEWS y mostrando un mejor balance entre sensibilidad y especificidad. Estos hallazgos respaldan la utilidad clínica del TREWS como herramienta complementaria al triaje convencional, particularmente en escenarios donde la estratificación del riesgo puede resultar ambigua. El punto de corte de 6 es el más apropiado para la práctica clínica, al ofrecer una mejor capacidad predictiva puesto que conserva un balance apropiado entre identificar oportunamente a los pacientes en riesgo y minimizar la pérdida de capacidad para descartar desenlaces adversos. Si bien las limitaciones metodológicas propias de un estudio unicéntrico y con población relativamente homogénea restringen la generalización de los resultados, este trabajo constituye la primera aproximación en Colombia, aportando evidencia local y abriendo la puerta a estudios multicéntricos y prospectivos. Para su implementación efectiva, recomendamos aplicarlo rutinariamente durante el triaje inicial en servicios de urgencias de alta complejidad, integrándolo con protocolos institucionales claros que faciliten la activación rápida de equipos de respuesta inmediata o derivación anticipada a áreas críticas. Se requiere automatización del cálculo mediante sistemas informáticos (sistema de historia clínica institucional), junto con capacitación continua del personal clínico para su interpretación adecuada.

11 Administración del proyecto

11.1 Presupuesto

Rubro	Valor	Fuente
Investigadores	\$ 6.000.000	Recursos propios
Equipos existentes (2 Computadores)	\$ 3.000.000	Recursos propios
Transporte	\$ 500.000	Recursos propios
Papelería	\$ 600.000	Recursos propios
Publicaciones y patentes	\$ 4.000.000	Recursos propios
TOTAL	\$ 14.100.000	100% Recursos propios

11.2 Cronograma

Actividad / Semestre	I	II	III	IV	V	VI	VII	VIII
Realización Protocolo de Investigación	x	x						
Sometimiento del protocolo al comité técnico científico			x					
Piloto de formatos de recolección de información				x				
Recolección de información					x	x		
Tabulación de los datos							x	
Análisis de los datos							x	
Redacción de informe final								x
Entrega Artículo final								x

12 Referencias

1. Singer AJ, Thode Jr HC, Viccellio P, Pines JM. The Association Between Length of Emergency Department Boarding and Mortality. *Acad Emerg Med.* 2011;18(12):1324–9.
2. Sun BC, Hsia RY, Weiss RE, Zingmond D, Liang LJ, Han W, McCreath H ASM. Effect of Emergency Department Crowding on Outcomes of Admitted Patients. *Ann Emerg Med.* 2013;61(6):605–11.
3. Bolívar Vargas MC, Valdivieso MHA, Robayo EAP, González DP, Ariza HP, Lizarazo LMG. Informe Defensorial – Derecho a la Salud Monitoreo a los servicios de urgencias de IPS en Colombia. *Defensoria Pueblo.* 2020;145.
4. Ministerio de Salud y Protección Social. Análisis de la Situación de Salud ASIS- Colombia 2020. *Minist Salud Prot Soc- Colomb.* 2020;98.
5. Hinson JS, Martinez DA, Cabral S, George K, Whalen M, Hansoti B, et al. Triage Performance in Emergency Medicine: A Systematic Review. *Ann Emerg Med.* Elsevier Inc.; 2019;74(1):140–52. DOI: 10.1016/j.annemergmed.2018.09.022
6. Guan G, Lee CMY, Begg S, Crombie A, Mnatzaganian G. The use of early warning system scores in prehospital and emergency department settings to predict clinical deterioration: A systematic review and meta-analysis. *PLoS ONE.* 2022;17(3):e0265559. DOI: 10.1371/journal.pone.0265559
7. Morgan RJM, Williams F WM. An early warning scoring system for detecting developing critical illness. *Clin Intensive Care.* 1997;8(100).
8. Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified early warning score in medical admissions. *QJM - Mon J Assoc Physicians.* 2001;94(10):521–6. DOI: 10.1093/qjmed/94.10.521
9. Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS-Towards a national early warning score for detecting adult inpatient deterioration. *Resuscitation.* Elsevier Ireland Ltd; 2010;81(8):932–7. DOI: 10.1016/j.resuscitation.2010.04.014
10. Royal College of Physicians (London). National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. Vol. 17, Report of a working party 2012.
11. Royal College of Physicians. National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity in the NHS. Vol. 17, Updated report of a working party 2017.
12. Agulnik A, Malone S, Puerto-Torres M, Gonzalez-Ruiz A, Vedaraju Y, Wang H, et al. Reliability and validity of a Spanish-language measure assessing clinical capacity to sustain Paediatric Early Warning Systems (PEWS) in resource-limited hospitals. *BMJ Open.* 2021;11(10):1–10. DOI: 10.1136/bmjopen-2021-053116
13. Paternina-Caicedo A, Miranda J, Bourjeily G, Levinson A, Dueñas C, Bello-Muñoz C, et al. Performance of the Obstetric Early Warning Score in critically ill patients for the prediction of

- maternal death. *Am J Obstet Gynecol*. Elsevier Inc.; 2017;216(1):58.e1-58.e8. DOI: 10.1016/j.ajog.2016.09.103
14. Covino M, Sandroni C, Della Polla D, De Matteis G, Piccioni A, De Vita A, et al. Predicting ICU admission and death in the Emergency Department: A comparison of six early warning scores. *Resuscitation*. 2023;190:109876. DOI: 10.1016/j.resuscitation.2023.109876
 15. Romero-Brufau S, Huddleston JM, Naessens JM, Johnson MG, Hickman J, Morlan BW, et al. Widely used track and trigger scores: are they ready for automation in practice? *Resuscitation*. 2014;85(4):549–52. DOI: 10.1016/j.resuscitation.2013.12.017
 16. Vergara P, Forero D, Bastidas A, Garcia J-C, Blanco J, Azocar J, et al. Validation of the National Early Warning Score (NEWS)-2 for adults in the emergency department in a tertiary-level clinic in Colombia. *Medicine (Baltimore)*. 2021;100(40):e27325. DOI: 10.1097/MD.00000000000027325
 17. Lee SB, Kim DH, Kim T, Kang C, Lee SH, Jeong JH, et al. Triage in Emergency Department Early Warning Score (TREWS) is predicting in-hospital mortality in the emergency department. *Am J Emerg Med*. Elsevier Inc.; 2019;38(2):203–10. DOI: 10.1016/j.ajem.2019.02.004
 18. Aygun H, Eraybar S. The role of emergency department triage early warning score (TREWS) and modified early warning score (MEWS) to predict in-hospital mortality in COVID-19 patients. *Ir J Med Sci*. Springer International Publishing; 2021;(0123456789). DOI: 10.1007/s11845-021-02696-y
 19. Alam N, Hobbelenk EL, van Tienhoven AJ, van de Ven PM, Jansma EP, Nanayakkara PWB. The impact of the use of the Early Warning Score (EWS) on patient outcomes: A systematic review. *Resuscitation*. European Resuscitation Council, American Heart Association, Inc., and International Liaison Committee on Resuscitation.~Published by Elsevier Ireland Ltd; 2014;85(5):587–94. DOI: 10.1016/j.resuscitation.2014.01.013
 20. Emergency Nurses Association. *Emergency Severity Index Handbook*. 5th ed. Woodfield Road; 2023.
 21. Ministerio de salud y protección social. Resolución 5596 del 2015: "Por la cual se definen los criterios técnicos para el Sistema de Selección y Clasificación de pacientes en los servicios de urgencias "Triage." [Internet]. 2015 Dec 24 p. 5.
 22. Gao H, McDonnell A, Harrison DA, Moore T, Adam S, Daly K, et al. Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. *Intensive Care Med*. 2007;33(4):667–79. DOI: 10.1007/s00134-007-0532-3
 23. Vergara P, Forero D, Bastidas A, Garcia J-C, Blanco J, Azocar J, et al. Validation of the National Early Warning Score (NEWS)-2 for adults in the emergency department in a tertiary-level clinic in Colombia. *Medicine (Baltimore)*. 2021;100(40):e27325. DOI: 10.1097/MD.00000000000027325
 24. Churpek MM, Yuen TC, Edelson DP. Risk stratification of hospitalized patients on the wards. *Chest*. 2013;143(6):1758–65. DOI: 10.1378/chest.12-1605
 25. Hillman K. Introduction of the medical emergency team (MET) system: A cluster-randomised controlled trial. *Lancet*. 2005;365(9477):2091–7. DOI: 10.1016/S0140-6736(05)66733-5

26. Bleyer AJ, Vidya S, Russell GB, Jones CM, Sujata L, Daeihagh P, et al. Longitudinal analysis of one million vital signs in patients in an academic medical center. *Resuscitation*. European Resuscitation Council, American Heart Association, Inc., and International Liaison Committee on Resuscitation.~Published by Elsevier Ireland Ltd; 2011;82(11):1387–92. DOI: 10.1016/j.resuscitation.2011.06.033
27. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. *Sci Transl Med*. 2015;7(299):299ra122. DOI: 10.1126/scitranslmed.aab3719
28. Niemantsverdriet MSA, Varkila MRJ, Vromen-Wijnsman JLP, Hoefler IE, Bellomo D, van Vliet MH, et al. Transportability and Implementation Challenges of Early Warning Scores for Septic Shock in the ICU: A Perspective on the TREWScore. *Front Med*. 2021;8:793815. DOI: 10.3389/fmed.2021.793815
29. Adams R, Henry KE, Sridharan A, Soleimani H, Zhan A, Rawat N, et al. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. *Nat Med*. Nature Research; 2022;28(7):1455–60. DOI: 10.1038/s41591-022-01894-0
30. Chang H, Park JE, Lee D, Lee K, Jekal SY, Moon KT, et al. Development and validation of a transformer model-based early warning score for real-time prediction of adverse outcomes in the emergency department. *Sci Rep*. 2025;15(1):23021. DOI: 10.1038/s41598-025-07511-7
31. Arévalo-Buitrago P, Morales-Cané I, Olivares Luque E, Guler I, Rodríguez-Borrego MA, López-Soto PJ. Predictive power of early-warning scores used in hospital emergency departments: a systematic review and meta-analysis. *Emerg Rev Soc Espanola Med Emerg*. 2021;33(5):374–81.
32. Wei S, Xiong D, Wang J, Liang X, Wang J, Chen Y. The accuracy of the National Early Warning Score 2 in predicting early death in prehospital and emergency department settings: a systematic review and meta-analysis. *Ann Transl Med*. 2023;11(2):95. DOI: 10.21037/atm-22-6587
33. Patel MS, Jones MA, Jiggins M, Williams SC. Does the use of a “track and trigger” warning system reduce mortality in trauma patients? *Injury*. 2011;42(12):1455–9. DOI: 10.1016/j.injury.2011.05.030
34. De Meester K, Haegdorens F, Monsieurs KG, Verpooten GA, Holvoet A, Van Bogaert P. Six-day postoperative impact of a standardized nurse observation and escalation protocol: a preintervention and postintervention study. *J Crit Care*. 2013;28(6):1068–74. DOI: 10.1016/j.jcrc.2013.07.061
35. Vergara P, Forero D, Bastidas A, Garcia JC, Blanco J, Azocar J, et al. Validation of the National Early Warning Score (NEWS)-2 for adults in the emergency department in a tertiary-level clinic in Colombia: Cohort study. *Medicine (Baltimore)*. 2021;100(40):e27325. DOI: 10.1097/MD.00000000000027325
36. Zachariasse JM, van der Hagen V, Seiger N, Mackway-Jones K, van Veen M, Moll HA. Performance of triage systems in emergency care: a systematic review and meta-analysis. *BMJ Open*. 2019;9(5):e026471. DOI: 10.1136/bmjopen-2018-026471

37. Saberian P, Abdollahi A, Hasani-Sharamin P, Modaber M, Karimialavijeh E. Comparing the prehospital NEWS with in-hospital ESI in predicting 30-day severe outcomes in emergency patients. *BMC Emerg Med.* 2022;22(1):42. DOI: 10.1186/s12873-022-00598-5
38. Fortaleza BrasilG 64^a A. Declaración de helsinki de la AMM – Principios éticos para las investigaciones médicas en seres humanos. 2013.
39. Ministerio de Salud, Colombia. Resolución 8430 de 1993: Por la cual se establecen las normas científicas, técnicas y administrativas para la investigación en salud. 1993.
40. Euser AM, Zoccali C, Jager KJ, Dekker FW. Cohort Studies: Prospective versus Retrospective. *Nephron Clin Pract.* 2009;113(3):c214–7. DOI: 10.1159/000235241
41. Obermeyer Z, Cohn B, Wilson M, Jena AB, Cutler DM. Early death after discharge from emergency departments: analysis of national US insurance claims data. *BMJ. British Medical Journal Publishing Group;* 2017;356:j239. DOI: 10.1136/bmj.j239
42. Pimentel MAF, Redfern OC, Gerry S, Collins GS, Malycha J, Prytherch D, et al. A comparison of the ability of the National Early Warning Score and the National Early Warning Score 2 to identify patients at risk of in-hospital mortality: A multi-centre database study. *Resuscitation.* 2019;134:147–56. DOI: 10.1016/j.resuscitation.2018.09.026
43. Alam N, Vegting IL, Houben E, van Berkel B, Vaughan L, Kramer MHH, et al. Exploring the performance of the National Early Warning Score (NEWS) in a European emergency department. *Resuscitation.* 2015;90:111–5. DOI: 10.1016/j.resuscitation.2015.02.011
44. Abbott TEF, Vaid N, Ip D, Cron N, Wells M, Torrance HDT, et al. A single-centre observational cohort study of admission National Early Warning Score (NEWS). *Resuscitation.* 2015;92:89–93. DOI: 10.1016/j.resuscitation.2015.04.020
45. [Internet]. National Early Warning Score (NEWS) 2 [cited 2025 Aug 12]. Available from: <https://www.rcp.ac.uk/improving-care/resources/national-early-warning-score-news-2/>
46. Holland M, Kellett J. A systematic review of the discrimination and absolute mortality predicted by the National Early Warning Scores according to different cut-off values and prediction windows. *Eur J Intern Med.* 2022;98:15–26. DOI: 10.1016/j.ejim.2021.12.024
47. Zhang K, Zhang X, Ding W, Xuan N, Tian B, Huang T, et al. National Early Warning Score Does Not Accurately Predict Mortality for Patients With Infection Outside the Intensive Care Unit: A Systematic Review and Meta-Analysis. *Front Med.* 2021;8:704358. DOI: 10.3389/fmed.2021.704358
48. Aygun H, Eraybar S. The role of emergency department triage early warning score (TREWS) and modified early warning score (MEWS) to predict in-hospital mortality in COVID-19 patients. *Ir J Med Sci.* 2022;191(3):997–1003. DOI: 10.1007/s11845-021-02696-y