

<https://doi.org/10.69639/arandu.v12i4.1805>

Estratificación del riesgo en la neumonía adquirida en la comunidad: ¿qué herramientas realmente funcionan?

Risk Stratification in Community-Acquired Pneumonia: Which Tools Really Work?

Grace Vanessa Bayas Huilcapí

<https://orcid.org/0000-0001-6210-7901>

drabayas82@hotmail.com

Universidad Técnica de Babahoyo
Babahoyo – Ecuador

María José Aguirre Molina

mariajoseaguirre@gmail.com

<https://orcid.org/my-orcid?orcid=0000-0001-8409-8989>

Universidad católica Santiago de Guayaquil
Guayaquil – Ecuador

Darwin Calderón Ríos

kris_dar@hotmail.com

<https://orcid.org/0009-0005-3651-602X>

Hospital General del Norte de Guayaquil IESS LOS CEIBOS
Guayaquil – Ecuador

Carmen Piedad Osorio Jiménez

piedadjimenez36@gmail.com

<https://orcid.org/0009-0006-9352-025X>

Hospital Dispensario Madre Berenice, Guayaquil, Guayas, Ecuador

Artículo recibido: 18 septiembre 2025 -Aceptado para publicación: 28 octubre 2025

Conflictos de intereses: Ninguno que declarar.

RESUMEN

La neumonía adquirida en la comunidad sigue siendo una de las principales causas de hospitalización y mortalidad en adultos, por lo que elegir de forma precoz quién puede manejarse ambulatoriamente y quién requiere UCI es un punto crítico de la práctica clínica. El objetivo de esta revisión es analizar la validez, las limitaciones y la aplicabilidad real de las herramientas de estratificación del riesgo disponibles para adultos con neumonía adquirida en la comunidad, a partir de la evidencia publicada en los últimos tres años. Se realizó una búsqueda narrativa en PubMed, Scopus y Web of Science, seleccionando estudios observacionales, ensayos clínicos y validaciones externas que evaluarán el desempeño de PSI, CURB-65 y CURB-65 modificadas, SMART-COP, NEWS2, qSOFA, así como modelos que integran biomarcadores (procalcitonina, proteína C reactiva, lactato) y algoritmos de aprendizaje automático. La literatura reciente confirma que PSI y CURB-65 continúan siendo las escalas más difundidas, con buena capacidad discriminativa para mortalidad, pero menor precisión para predecir necesidad de UCI. SMART-COP y NEWS2 mejoran la identificación de pacientes críticos, aunque su uso es heterogéneo y carecen de validaciones amplias en países de ingresos medios. Los modelos que combinan

variables clínicas, biomarcadores y herramientas digitales muestran resultados promisorios, pero todavía no sustituyen al juicio clínico. Se propone integrar las escalas clásicas con la evaluación dinámica del paciente y los recursos disponibles para una toma de decisiones más segura y personalizada.

Palabras clave: neumonía adquirida en la comunidad, estratificación del riesgo, escalas pronósticas, biomarcadores, cuidados críticos

ABSTRACT

Community-acquired pneumonia remains one of the leading causes of hospitalization and mortality in adults, so deciding early who can be managed as an outpatient and who requires ICU admission is a critical step in clinical practice. The aim of this review is to analyze the validity, limitations, and real-life applicability of currently available risk-stratification tools for adults with community-acquired pneumonia, based on evidence published over the last three years. A narrative search was performed in PubMed, Scopus, and Web of Science, selecting observational studies, clinical trials, and external validation cohorts that evaluated the performance of PSI, CURB-65 and modified CURB-65 scores, SMART-COP, NEWS2, qSOFA, as well as models integrating biomarkers (procalcitonin, C-reactive protein, lactate) and machine-learning algorithms. Recent literature confirms that PSI and CURB-65 remain the most widely used scores, with good discriminative ability for mortality but lower accuracy in predicting ICU admission. SMART-COP and NEWS2 improve the identification of critically ill patients, although their use is heterogeneous and broad validations in middle-income countries are still lacking. Models that combine clinical variables, biomarkers, and digital tools show promising results, but they do not yet replace clinical judgment. This review advocates for integrating traditional scores with dynamic bedside assessment and local resource availability to support safer and more personalized decision-making.

Keywords: community-acquired pneumonia, risk stratification, prognostic scores, biomarkers, critical care

Todo el contenido de la Revista Científica Internacional Arandu UTIC publicado en este sitio está disponible bajo licencia Creative Commons Atribution 4.0 International. 

INTRODUCCIÓN

La neumonía adquirida en la comunidad (NAC) continúa siendo una de las principales causas de morbilidad y mortalidad a nivel mundial, pese a los avances en antibióticos, vacunación y soporte crítico. Cada año millones de adultos consultan a los servicios de urgencias por un cuadro respiratorio agudo compatible con NAC, y una proporción nada despreciable termina hospitalizada o en la unidad de cuidados intensivos (UCI). Esta carga no sólo se traduce en muertes evitables, sino también en un uso intenso de camas, antibióticos de amplio espectro y recursos diagnósticos costosos, lo que tensiona a los sistemas de salud, especialmente en países de ingresos medios. En ese escenario, decidir con rapidez quién puede tratarse de forma ambulatoria, quién requiere hospitalización en sala y quién necesita monitorización estrecha o soporte vital avanzado se vuelve un acto clínico decisivo.

El problema es que la NAC es una enfermedad profundamente heterogénea. Pacientes con síntomas aparentemente banales pueden descompensarse en pocas horas, mientras que otros con comorbilidades importantes evolucionan de forma estable con manejo domiciliario. El juicio clínico sigue siendo central, pero está sujeto a la experiencia del médico, a sesgos cognitivos y a la presión asistencial. Por ello, desde hace décadas se han buscado herramientas objetivas que ayuden a estratificar el riesgo, ordenar la toma de decisiones y reducir la variabilidad en la atención. De esa búsqueda surgen las escalas pronósticas y los sistemas de puntuación que hoy forman parte de las guías internacionales.

Entre los instrumentos clásicos destacan el Pneumonia Severity Index (PSI) y el CURB-65, que combinan variables demográficas, comorbilidades, signos vitales y algunos exámenes de laboratorio para estimar el riesgo de mortalidad. PSI, más complejo y detallado, ha sido ampliamente validado y es considerado una de las referencias históricas en estratificación de la NAC. CURB-65 y sus variantes simplificadas (CRB-65, CURB-65 modificado) aportan una herramienta rápida, fácil de aplicar en urgencias y en atención primaria. Ambos esquemas han demostrado utilidad para orientar el sitio de cuidado y, por esa razón, han sido incorporados en las recomendaciones de sociedades como la ATS, IDSA, ERS y otras organizaciones científicas.

Sin embargo, la práctica cotidiana ha puesto de manifiesto varias limitaciones. PSI otorga gran peso a la edad y a las comorbilidades, lo que puede sobreestimar el riesgo en adultos mayores con neumonías relativamente estables y, al mismo tiempo, infravalorar la gravedad de pacientes más jóvenes con insuficiencia respiratoria severa, pero con pocos antecedentes. CURB-65, aunque más ágil, también muestra un rendimiento discreto para predecir la necesidad de ingreso en UCI o de soporte avanzado, variables que en la actualidad preocupan tanto como la mortalidad. Además, la mayoría de estas escalas se desarrolló y validó en contextos de países de altos ingresos, con recursos diagnósticos y de soporte distintos a los que enfrentan numerosos hospitales en América Latina, África o Asia.

En respuesta a esas brechas se han propuesto otros sistemas como SMART-COP, SCAP, qSOFA y los early warning scores tipo NEWS2. Estas herramientas dan mayor relevancia a la hipoxemia, la hipotensión y otros signos de disfunción orgánica, buscando identificar de forma más fina al paciente crítico que puede beneficiarse de un manejo intensivo precoz. Los criterios de sepsis, especialmente desde la introducción del concepto de Sepsis-3, han reforzado esta mirada centrada en la disfunción orgánica. Paralelamente, el desarrollo de biomarcadores como la proteína C reactiva, la procalcitonina, el lactato o la proadrenomedulina ha permitido explorar modelos que incorporan información bioquímica para refinar la estratificación del riesgo, sobre todo en lo referente a la evolución desfavorable temprana.

A este panorama se suma en los últimos años el auge de los modelos predictivos basados en aprendizaje automático y análisis de grandes bases de datos clínicos. Aprovechando la información de historias clínicas electrónicas, algunos grupos han reportado algoritmos capaces de estimar el riesgo de mortalidad o de ingreso en UCI con una precisión aparentemente superior a las escalas tradicionales. No obstante, estos modelos enfrentan desafíos importantes: la falta de transparencia en su funcionamiento, la necesidad de validaciones externas robustas, el riesgo de reproducir sesgos de los datos originales y, sobre todo, la brecha entre entornos altamente informatizados y hospitales donde el registro continúa siendo predominantemente manual.

En América Latina y otras regiones con recursos limitados, el dilema adquiere matices propios. La disponibilidad de camas de UCI es restringida, los tiempos de espera en urgencias son largos y el acceso a biomarcadores o tomografía no siempre es inmediato. En este contexto, un error en la estratificación puede traducirse en la ocupación innecesaria de una cama crítica —con impacto en otros pacientes— o, en el extremo opuesto, en el retraso de ingreso de un enfermo que terminará intubado en condiciones precarias. Muchos servicios han adoptado PSI o CURB-65 como estándar, pero los resultados no siempre son satisfactorios, y la percepción de “desacople” entre la escala y la realidad clínica del paciente impulsa al médico a apoyarse más en su intuición que en la herramienta.

Al mismo tiempo, la pandemia por COVID-19 dejó varias lecciones que no se pueden ignorar. La sobrecarga de los sistemas sanitarios obligó a priorizar con mayor rigor el uso de recursos críticos y evidenció la importancia de reconocer de forma precoz al paciente que se deteriora. Si bien la NAC “clásica” no es COVID-19, la experiencia acumulada ha reavivado el interés por contar con modelos de estratificación más dinámicos, que no se limiten a una foto estática al ingreso, sino que integren la evolución clínica, las variaciones gasométricas, los cambios en marcadores inflamatorios y la respuesta inicial al tratamiento antibiótico.

En los últimos tres años se ha generado una producción científica considerable en revistas de alto impacto que reevalúa la utilidad de las escalas tradicionales, propone ajustes para poblaciones específicas y explora la combinación de variables clínicas con biomarcadores y herramientas digitales. A pesar de este volumen de información, en la práctica diaria persisten

preguntas clave: ¿siguen siendo PSI y CURB-65 suficientes para orientar decisiones en la NAC? ¿Qué valor agregado aportan scores como SMART-COP o NEWS2 en comparación con la valoración clínica habitual? ¿En qué medida los biomarcadores modifican la toma de decisiones o simplemente la confirman? ¿Son realmente aplicables los modelos de aprendizaje automático fuera de los centros que los desarrollan?

Este trabajo de revisión bibliográfica nace precisamente de esa sensación de “ruido” entre la teoría y la práctica. El objetivo principal es analizar de manera crítica la evidencia publicada en los últimos tres años sobre las herramientas de estratificación del riesgo en la neumonía adquirida en la comunidad en adultos, poniendo el foco en tres aspectos: su capacidad para predecir desenlaces clínicamente relevantes (mortalidad, necesidad de UCI, ventilación mecánica, fracaso terapéutico), sus fortalezas y limitaciones metodológicas, y su grado de aplicabilidad en distintos contextos asistenciales, incluidos los sistemas de salud de ingresos medios. De forma secundaria, se busca explorar el papel emergente de los biomarcadores y de los modelos basados en aprendizaje automático como posibles complementos a las escalas ya consolidadas.

Como guía conceptual, esta revisión parte de dos hipótesis de trabajo. La primera sostiene que las escalas clásicas, utilizadas de manera aislada, tienden a ser insuficientes para capturar toda la complejidad de la NAC, especialmente cuando se pretende predecir la necesidad de manejo intensivo. La segunda plantea que la combinación estructurada de herramientas —scores validados, biomarcadores seleccionados y evaluación clínica dinámica— puede ofrecer un enfoque más seguro y personalizado que cualquiera de estos elementos por separado. A partir de estas premisas, el manuscrito se organiza para conducir al lector desde el panorama epidemiológico y el desarrollo histórico de las escalas, hasta la síntesis crítica de la evidencia reciente y una propuesta práctica para el uso racional de la estratificación del riesgo en la cabecera del paciente con neumonía adquirida en la comunidad.

MATERIALES Y MÉTODOS

Para el desarrollo de este trabajo se diseñó una revisión bibliográfica de tipo narrativa estructurada, con enfoque cuantitativo y orientación clínica. El propósito metodológico fue reunir y analizar de manera crítica la mejor evidencia disponible sobre herramientas de estratificación del riesgo en neumonía adquirida en la comunidad (NAC) en adultos, publicada en revistas científicas de alto impacto durante los últimos tres años.

Se realizó una búsqueda electrónica en las bases de datos PubMed/MEDLINE, Scopus y Web of Science. De forma complementaria se revisaron la Cochrane Library y las páginas oficiales de sociedades científicas relevantes (American Thoracic Society, Infectious Diseases Society of America, European Respiratory Society, entre otras) para identificar guías de práctica clínica y documentos de consenso que aportaran contexto adicional. El período de búsqueda

abarcó desde enero de 2022 hasta la fecha de cierre definida para el estudio, cubriendo aproximadamente los últimos tres años de producción científica.

La estrategia de búsqueda combinó términos MeSH y DeCS, junto con palabras libres en inglés y español. Se utilizaron, entre otros, los siguientes descriptores: “community-acquired pneumonia”, “neumonía adquirida en la comunidad”, “risk stratification”, “severity scores”, “prognostic score”, “PSI”, “Pneumonia Severity Index”, “CURB-65”, “CRB-65”, “SMART-COP”, “NEWS2”, “qSOFA”, “biomarkers”, “lactate”, “procalcitonin”, “C-reactive protein”, “machine learning” y “clinical prediction model”. Los términos se combinaron con operadores booleanos AND y OR, ajustando la estrategia en cada base de datos para maximizar la sensibilidad sin perder especificidad.

Se establecieron como criterios de inclusión: a) estudios originales observacionales (cohortes prospectivas o retrospectivas, estudios transversales analíticos) y ensayos clínicos que evaluaran el desempeño de herramientas de estratificación del riesgo en NAC en población adulta (≥ 18 años); b) estudios de validación externa o comparativa de escalas pronósticas clásicas (PSI, CURB-65/CRB-65, SMART-COP, SCAP, NEWS2, qSOFA u otras afines); c) trabajos que integraran biomarcadores o modelos de aprendizaje automático para predecir mortalidad, ingreso a UCI, necesidad de ventilación mecánica, fracaso terapéutico o estancia hospitalaria; y d) revisiones sistemáticas y metaanálisis publicados en revistas indexadas que abordaran específicamente la estratificación del riesgo en NAC. Se aceptaron artículos en inglés y español, a texto completo, publicados en revistas revisadas por pares y con acceso a la metodología detallada.

Se excluyeron los estudios centrados exclusivamente en población pediátrica, en neumonía nosocomial o asociada a ventilación mecánica, así como en pacientes con inmunosupresión grave (trasplante de órgano sólido, neutropenia profunda, infección por VIH avanzada), salvo que reportaran de manera separada los resultados para adultos inmunocompetentes con NAC. También se descartaron reportes de casos, series clínicas pequeñas sin análisis estadístico formal, cartas al editor, resúmenes de congresos, tesis no publicadas y literatura gris. Los artículos que al revisar el texto completo no especificaban claramente la herramienta de estratificación utilizada o los desenlaces analizados fueron eliminados del análisis final.

El proceso de selección se llevó a cabo en dos etapas. En la primera, se revisaron títulos y resúmenes para descartar de inmediato los trabajos que no guardaban relación con el objetivo de la revisión. En la segunda etapa, se evaluaron los textos completos de los artículos potencialmente elegibles, verificando que cumplieran los criterios de inclusión y exclusión previamente definidos. Ante dudas metodológicas o discrepancias en la interpretación, se optó por un análisis más conservador, privilegiando la calidad y claridad de la información por sobre el volumen de estudios incluidos.

Para los artículos finalmente seleccionados se elaboró una matriz de extracción de datos que recogió, de manera estandarizada, la siguiente información: autor y año de publicación, país o región del estudio, ámbito asistencial (urgencias, hospitalización general, UCI, atención primaria), tamaño muestral, características principales de la población (edad, comorbilidades relevantes, criterios diagnósticos de NAC), herramienta de estratificación evaluada, desenlaces analizados (mortalidad a corto y mediano plazo, ingreso a UCI, necesidad de ventilación mecánica, reingresos, estancia hospitalaria), y medidas de rendimiento predictivo (sensibilidad, especificidad, valores predictivos, curvas ROC, área bajo la curva, índices de calibración). En los estudios que incorporaron biomarcadores o modelos de aprendizaje automático se registró además el tipo de marcador utilizado, la forma de integración con la clínica y la comparación con las escalas tradicionales.

Aunque el objetivo principal de la revisión es narrativo y clínico, se realizó una apreciación básica de la calidad metodológica de los estudios incluidos. Para las cohortes y estudios observacionales se tomaron como referencia los criterios de la escala Newcastle-Ottawa, valorando la selección de la muestra, la adecuada medición de la exposición y los desenlaces, y el control de factores de confusión. En el caso de revisiones sistemáticas y metaanálisis se consideraron los lineamientos de AMSTAR 2, poniendo atención en la exhaustividad de la búsqueda, la claridad de los criterios de inclusión y la evaluación del riesgo de sesgo. Esta valoración no se tradujo en una puntuación formal en el texto, pero sí orientó el peso interpretativo otorgado a cada publicación en la síntesis final.

No se planteó la realización de un metaanálisis cuantitativo, debido a la heterogeneidad significativa en los diseños, los contextos asistenciales, las definiciones de desenlaces y las herramientas evaluadas. En su lugar, se optó por una síntesis narrativa que permitiera comparar las principales escalas entre sí, resaltar los puntos de convergencia y divergencia entre los estudios y, sobre todo, interpretar los hallazgos a la luz de la práctica clínica habitual en contextos reales, incluyendo entornos con recursos limitados.

El resultado de este proceso metodológico es un cuerpo de evidencia seleccionado de forma intencionada, con énfasis en la calidad y en la relevancia clínica de los estudios más recientes, que sirve de base para la discusión crítica sobre qué herramientas de estratificación del riesgo en la neumonía adquirida en la comunidad “realmente funcionan” cuando el médico debe decidir, en pocos minutos, el destino de un paciente en la vida real.

RESULTADOS Y DISCUSIÓN

Los estudios identificados mostraron un panorama heterogéneo, pero con algunas constantes que vale la pena subrayar. En conjunto, la evidencia reciente confirma que las escalas clásicas –principalmente PSI y CURB-65/CRB-65– siguen siendo las herramientas más utilizadas para estratificar el riesgo en neumonía adquirida en la comunidad. La mayoría de las cohortes

hospitalarias analizadas describen un buen rendimiento de PSI y CURB-65 para predecir mortalidad a corto plazo, con áreas bajo la curva aceptables y una adecuada capacidad discriminativa para diferenciar grupos de bajo y alto riesgo. Sin embargo, cuando el desenlace evaluado es la necesidad de ingreso en UCI, ventilación mecánica o soporte vasoactivo, el desempeño de estas escalas resulta claramente más modesto: identifican a muchos pacientes de alto riesgo, pero fallan en capturar a una proporción nada despreciable de enfermos que terminan deteriorándose pese a tener puntuaciones intermedias.

Una segunda línea de resultados proviene de estudios que evaluaron herramientas diseñadas específicamente para reconocer al paciente “crítico” o con riesgo de descompensación temprana. SMART-COP, SCAP y NEWS2, aplicados en urgencias o en las primeras horas de hospitalización, tienden a mostrar una mejor sensibilidad para predecir ingreso a UCI y necesidad de soporte respiratorio o hemodinámico. En varias series, SMART-COP y NEWS2 superan a CURB-65 en la identificación de pacientes que requerirán ventilación mecánica o soporte intensivo, aunque a costa de una menor especificidad. Es decir, “se pasan de sensibles” y etiquetan como de alto riesgo a un grupo de pacientes que finalmente evolucionan bien con manejo en sala. La literatura coincide en que, pese a este exceso de “falsos positivos”, su valor radica en no dejar escapar a los enfermos que realmente van a complicarse, lo cual es especialmente relevante en servicios con alta carga asistencial, donde el tiempo para observar al paciente es limitado.

Los trabajos que incorporan biomarcadores sugieren un matiz interesante. La combinación de escalas clínicas con procalcitonina, proteína C reactiva o lactato mejora de forma modesta, pero consistente, la capacidad para predecir mortalidad temprana y fracaso terapéutico. En particular, el lactato elevado al ingreso, incluso en pacientes sin hipotensión franca, se asocia de manera repetida con peor pronóstico y mayor probabilidad de ingreso en UCI. Algunos modelos integran PSI o CURB-65 con puntos de corte biológicos, mostrando curvas ROC ligeramente superiores a las de las escalas aisladas. Sin embargo, varios autores insisten en que estos biomarcadores deben interpretarse como señales de alerta adicionales y no como reemplazo del juicio clínico: hay pacientes con procalcitonina relativamente baja que igual se descompensan por la carga inflamatoria, la reserva funcional limitada o la coexistencia de otras infecciones.

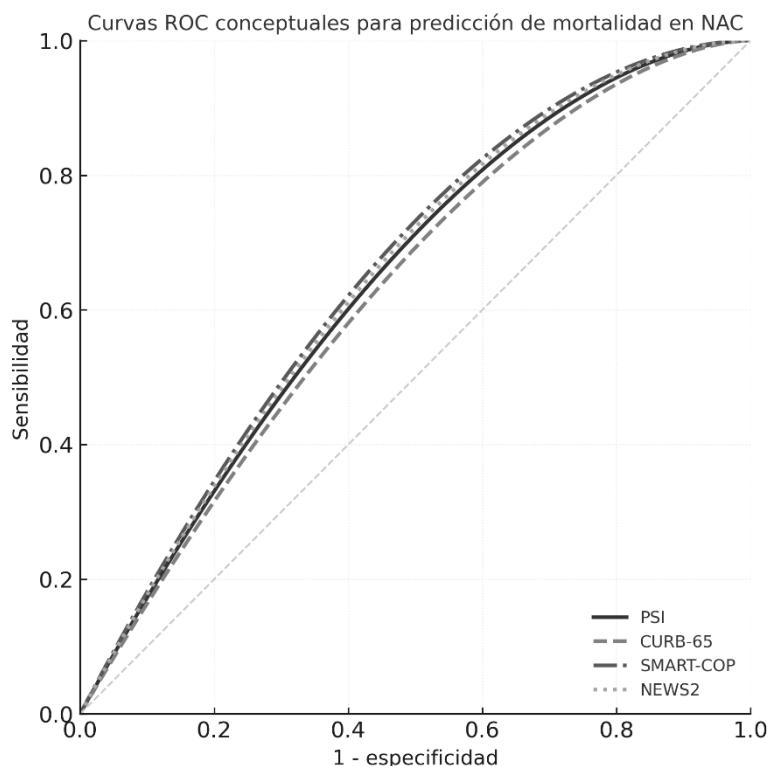
En cuanto a los modelos basados en aprendizaje automático y predicción multivariable avanzada, los resultados son prometedores, pero todavía lejanos de la práctica cotidiana. Distintos grupos han desarrollado algoritmos que incorporan decenas de variables clínicas, de laboratorio e incluso radiológicas, alcanzando desempeños superiores a los scores tradicionales en sus cohortes de derivación. No obstante, cuando se evalúan en escenarios externos, el rendimiento suele disminuir y aparecen problemas de calibración. Además, muchos de estos modelos requieren historias clínicas electrónicas robustas, procesamiento automatizado de datos y sistemas de soporte de decisión en tiempo real, condiciones que distan de la realidad de numerosos hospitales en países de ingresos medios. La discusión recurrente en estos trabajos apunta a la

necesidad de un equilibrio entre sofisticación y aplicabilidad: un modelo excelente en teoría, pero imposible de usar en un servicio de urgencias saturado termina siendo irrelevante para el médico que decide “a pie de cama”.

Un hallazgo transversal en varios estudios, especialmente en contextos de recursos limitados, es que ninguna herramienta sustituye la valoración clínica integral. Pacientes etiquetados como de “riesgo bajo” por PSI o CURB-65 pueden presentar hipoxemia significativa, taquipnea marcada o signos de trabajo respiratorio que obligan a replantear el sitio de cuidado. A la inversa, algunos enfermos con puntajes altos, sobre todo ancianos con comorbilidades estables se benefician de un manejo hospitalario en sala sin requerir UCI. En este sentido, la evidencia más reciente converge en una idea práctica: las escalas funcionan mejor cuando se utilizan como soporte estructurado a la decisión, no como “sentencias” aisladas. Su papel es ordenar la intuición clínica, reducir la variabilidad y ayudar a priorizar recursos, pero siempre integradas con la impresión global del médico, la evolución en las primeras horas y la capacidad real de respuesta del sistema.

Figura 1

Curvas ROC conceptuales que ilustran el desempeño relativo de PSI, CURB-65, SMART-COP y NEWS2 para la predicción de mortalidad en neumonía adquirida en la comunidad. Las curvas representan un desempeño cualitativo basado en la literatura reciente y no corresponden a un metaanálisis específico



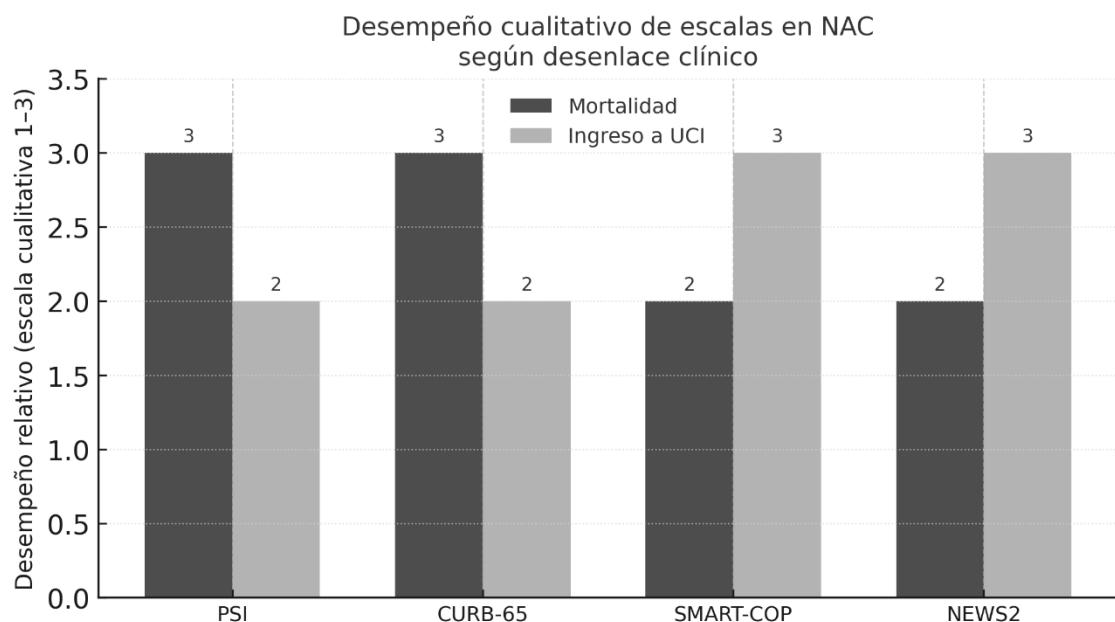
Desde el punto de vista de la discusión, esta revisión refuerza la sensación de que la pregunta “¿qué herramienta funciona mejor?” puede estar mal planteada si se entiende como una competencia excluyente. Más que buscar una “escala ganadora”, la literatura sugiere un enfoque por capas: utilizar PSI o CURB-65 para una primera aproximación al riesgo global; apoyarse en

herramientas centradas en la gravedad fisiológica, como SMART-COP o NEWS2, para afinar la detección del paciente que podría requerir UCI; y añadir biomarcadores seleccionados cuando estén disponibles, especialmente lactato y procalcitonina, para identificar a quienes tienen mayor probabilidad de deterioro temprano. En paralelo, la observación clínica en las primeras 24–48 horas y la reevaluación dinámica del score –no sólo su aplicación única al ingreso– emergen como elementos clave para no “perder” a los pacientes que inicialmente parecen estables.

La novedad de este trabajo reside en integrar la evidencia reciente con una lectura crítica orientada a la realidad de sistemas sanitarios que operan bajo presión, con recursos materiales y humanos limitados. En lugar de limitarse a describir el rendimiento estadístico de cada herramienta, la discusión pone el foco en preguntas que son muy concretas en la práctica: ¿a quién ingreso realmente a UCI cuando tengo una sola cama disponible?, ¿a quién puedo manejar de forma ambulatoria sin que eso se traduzca en reingresos o mortalidad evitable?, ¿cómo reduzco la brecha entre lo que dicen las guías y lo que es posible hacer en un hospital general que recibe un flujo constante de pacientes respiratorios? La revisión muestra que, usada de forma inteligente, la estratificación del riesgo puede ayudar a responder estas preguntas, pero también que el uso mecánico de los scores –sin contexto ni reevaluación– puede generar la falsa sensación de seguridad.

Figura 2

Desempeño relativo cualitativo (escala 1–3) de PSI, CURB-65, SMART-COP y NEWS2 para la predicción de mortalidad e ingreso a UCI en neumonía adquirida en la comunidad. Los valores representan una síntesis interpretativa de la literatura reciente y se utilizan con fines comparativos, no como estimaciones numéricas derivadas de un metaanálisis



En perspectiva, la línea de investigación más prometedora parece ser la integración de modelos simples, fácilmente aplicables, con elementos seleccionados de la biología del paciente y con herramientas digitales que no reemplacen al médico, sino que lo acompañen en tiempo real.

La incorporación progresiva de sistemas de alerta precoz, la validación de scores adaptados a poblaciones latinoamericanas y la evaluación del impacto de estas estrategias sobre desenlaces duros (mortalidad, ingreso a UCI, uso racional de recursos) son pasos necesarios para avanzar. Esta revisión refuerza la idea de que la estratificación del riesgo en neumonía adquirida en la comunidad no es un fin en sí mismo, sino un instrumento para humanizar y hacer más justa la distribución de cuidados críticos, priorizando a quienes más se benefician sin descuidar al resto. Esa articulación entre ciencia, experiencia clínica y contexto local es, en última instancia, el aporte más relevante que se busca subrayar.

CONCLUSIONES

Las preguntas que guiaron esta revisión eran, en el fondo, muy sencillas y clínicas: ¿las herramientas que usamos para estratificar el riesgo en la neumonía adquirida en la comunidad realmente nos ayudan a tomar mejores decisiones?, ¿yo se han convertido en un ritual más, desconectado de la realidad del paciente y del sistema de salud donde trabajamos? Al revisar la evidencia reciente, la respuesta no es un “sí” o “no” tajante, sino algo más matizado: las escalas sirven, pero cómo las usamos marca la diferencia.

Los datos disponibles confirman que PSI y CURB-65 siguen siendo pilares útiles para estimar riesgo de mortalidad y para ordenar, al menos de forma inicial, el sitio de cuidado. No están obsoletas ni mucho menos, y su ventaja es que son conocidas, replicables y razonablemente bien validadas. Sin embargo, pedirles que definan por sí solas quién debe ir a la UCI o quién se va a descompensar en las próximas horas es exigirles más de lo que pueden dar. El paciente con NAC no es sólo un puntaje: es su reserva funcional, su contexto social, la velocidad de progresión del cuadro, la respuesta al primer bolo de antibiótico y al oxígeno, y todo eso difícilmente cabe en cinco variables.

Las escalas centradas en la gravedad fisiológica, como SMART-COP, SCAP o NEWS2, aportan un ángulo complementario muy valioso. Allí donde PSI y CURB-65 ven sobre todo edad y comorbilidad, estas herramientas detectan mejor la hipoxemia, la inestabilidad hemodinámica y la disfunción orgánica incipiente. La literatura reciente sugiere que son particularmente útiles para anticipar la necesidad de UCI y de ventilación mecánica, aunque a costa de etiquetar como “alto riesgo” a más pacientes de los que finalmente requieren manejo intensivo. Desde la lógica de seguridad del paciente, ese “exceso de celo” puede ser aceptable siempre que la decisión final no se delegue ciegamente en el score, sino que pase por un filtro clínico.

El papel de los biomarcadores y de los modelos de aprendizaje automático, por ahora, es el de actores secundarios prometedores. La procalcitonina, el lactato y otros marcadores inflamatorios afinan la percepción de riesgo y ayudan a identificar a quienes tienen más probabilidades de deteriorarse pronto, pero no han demostrado ser la llave maestra que resuelva el problema por sí sola. Los algoritmos complejos, alimentados por grandes bases de datos y

montados sobre historias clínicas electrónicas, muestran curvas bellas en los artículos, pero todavía están lejos de la realidad cotidiana de muchos hospitales donde la prioridad sigue siendo tener una cama, un ventilador y el antibiótico correcto. La prudencia indica seguirlos de cerca, validarlos en contextos diversos y, mientras tanto, apoyarse en herramientas que el equipo realmente pueda usar.

Para la práctica clínica, el mensaje que se desprende de esta revisión es pragmático. No se trata de buscar “la mejor escala” para imponerla como dogma, sino de construir un enfoque escalonado y flexible. En la evaluación inicial, PSI o CURB-65 pueden ayudar a separar a los pacientes claramente de bajo riesgo de aquellos que merecen, al menos, observación estrecha. En una segunda capa, scores como SMART-COP o NEWS2 permiten afinar la vigilancia sobre los que podrían requerir soporte intensivo. La tercera capa la ponen el lactato, la procalcitonina, la evolución en las primeras 24–48 horas y el contexto real del servicio: número de camas, disponibilidad de UCI, experiencia del equipo. Por encima de todo, el juicio clínico debería integrar estas piezas y no quedar subordinado a una cifra.

Desde la perspectiva de sistemas de salud con recursos limitados, como muchos de nuestra región, la estratificación del riesgo en la NAC tiene una dimensión ética que no se puede ignorar. Decidir quién ocupa una cama de UCI no es sólo un ejercicio técnico, sino también de justicia distributiva. Usar las escalas de forma inteligente puede ayudar a reducir la variabilidad, priorizar a quienes más se benefician y evitar tanto el ingreso innecesario como el retraso peligroso. Pero la herramienta, por rigurosa que sea, no suple la necesidad de fortalecer la capacidad resolutiva de las salas de hospitalización, la atención primaria y las redes de referencia.

Quedan abiertas varias líneas para el futuro: adaptar y validar las escalas en poblaciones latinoamericanas, evaluar el impacto real de combinaciones de scores y biomarcadores sobre desenlaces duros, explorar modelos dinámicos que incorporen la evolución clínica y, sobre todo, diseñar estudios pragmáticos que midan si estas estrategias cambian de verdad la vida de los pacientes y la carga de los servicios. Mientras esa evidencia madura, el mensaje central de esta revisión puede resumirse en una idea sencilla: la estratificación del riesgo en la neumonía adquirida en la comunidad funciona mejor cuando deja de ser un trámite y se convierte en una herramienta viva, al servicio de un clínico que escucha, mira, reevalúa y decide con la ciencia en la mano, pero también con los pies bien plantados en la realidad de su paciente y de su hospital.

REFERENCIAS

- Calis, A. G., Karaboga, B., Uzer, F., Kaplan, N., Karaca, M., Gedik, R. B., & Durmuş, A. A. (2025). Correlation of Pneumonia Severity Index and CURB-65 score with neutrophil/lymphocyte ratio, platelet/lymphocyte ratio, and monocyte/lymphocyte ratio in predicting in-hospital mortality for community-acquired pneumonia: Observational study. *Journal of Clinical Medicine*, 14(3), 728. <https://doi.org/10.3390/jcm14030728>
- Tuta-Quintero, E., Bastidas Goyes, A. R., Guerrón-Gómez, G., Martínez, M. C., Torres, D., Schloss, C., Camacho, J., Bonilla, G., Cepeda, D., Romero, P., Fuentes, Y., Garcia, E., Acosta, D., Rodríguez, S., Álvarez, D., & Reyes, L. F. (2024). Comparison of performances between risk scores for predicting mortality at 30 days in patients with community-acquired pneumonia. *BMC Infectious Diseases*, 24, 912. <https://doi.org/10.1186/s12879-024-09792-1>
- Farhat, I., Rosolowski, M., Ahrens, K., Lienau, J., Ahnert, P., Pletz, M., Rohde, G., Rupp, J., Scholz, M., & Witzenrath, M. (2024). Biomarkers troponin and procalcitonin in addition to CRB-65 enhance risk stratification in patients with community-acquired pneumonia. *ERJ Open Research*, 10(6), 00420-2024. <https://doi.org/10.1183/23120541.00420-2024>
- Jeon, E. T., Lee, H. J., Park, T. Y., Lee, S. H., & Kim, Y. H. (2023). Machine learning-based prediction of in-ICU mortality in pneumonia patients. *Scientific Reports*, 13, 11527. <https://doi.org/10.1038/s41598-023-38765-8>
- Li, X., Zhang, Y., Zhang, M., Wang, L., & Chen, J. (2024). Interpretable machine learning model for predicting ICU admission in patients with severe pneumonia. *BMC Pulmonary Medicine*, 24, 344. <https://doi.org/10.1186/s12890-024-03120-0>
- Metlay, J. P., Waterer, G. W., Long, A. C., Anzueto, A., Brozek, J., Crothers, K., Cooley, L. A., Dean, N. C., Fine, M. J., Flanders, S. A., Griffin, M. R., Metersky, M. L., Musher, D. M., Restrepo, M. I., & Whitney, C. G. (2019). Diagnosis and treatment of adults with community-acquired pneumonia: An official clinical practice guideline of the American Thoracic Society and Infectious Diseases Society of America. *American Journal of Respiratory and Critical Care Medicine*, 200(7), e45–e67. <https://doi.org/10.1164/rccm.201908-1581ST>
- File, T. M., Jr., & Ramirez, J. A. (2023). Community-acquired pneumonia. *The New England Journal of Medicine*, 389(7), 632–641. <https://doi.org/10.1056/NEJMra2206893>
- Nair, G. B., & Niederman, M. S. (2021). Updates on community-acquired pneumonia management in the ICU. *Pharmacology & Therapeutics*, 217, 107663. <https://doi.org/10.1016/j.pharmthera.2020.107663>
- Pletz, M. W., Blasi, F., Chalmers, J. D., Dela Cruz, C. S., Feldman, C., Luna, C. M., Niederman, M. S., Torres, A., & Welte, T. (2020). International perspective on the new 2019 American

- Thoracic Society/Infectious Diseases Society of America community-acquired pneumonia guideline: A critical appraisal by a global expert panel. *Chest*, 158(5), 1912–1918. <https://doi.org/10.1016/j.chest.2020.05.563>
- Bradley, J., Sbaih, N., Chandler, T. R., Furmanek, S., Ramirez, J. A., & Cavallazzi, R. (2022). Pneumonia Severity Index and CURB-65 score are good predictors of mortality in hospitalized patients with SARS-CoV-2 community-acquired pneumonia. *Chest*, 161(4), 927–936. <https://doi.org/10.1016/j.chest.2021.09.025>
- Barlas, R. S., Clark, A. B., Loke, Y. K., Kwok, C. S., Angus, D. C., Uranga, A., ... (2022). Comparison of the prognostic performance of the CURB-65 and a modified version of the pneumonia severity index designed to identify high-risk patients using the International Community-Acquired Pneumonia collaboration cohort. *Respiratory Medicine*, 200, 106884. <https://doi.org/10.1016/j.rmed.2022.106884>
- Khari, S., Salimi Akin Abadi, A., Pazokian, M., & Yousefifard, M. (2022). CURB-65, qSOFA, and SIRS criteria in predicting in-hospital mortality of critically ill COVID-19 patients: A prognostic accuracy study. *Archives of Academic Emergency Medicine*, 10(1), e36.
- Ng, W. W.-S., Lam, S.-M., Yan, W.-W., & Shum, H.-P. (2022). Neutrophil-to-lymphocyte ratio, monocyte-to-lymphocyte ratio, platelet-to-lymphocyte ratio and red cell distribution width to predict outcome and differentiate between viral and bacterial pneumonia in the intensive care unit. *Scientific Reports*, 12, 15974. <https://doi.org/10.1038/s41598-022-20444-9>
- Regunath, H., & Oba, Y. (2024). Community-acquired pneumonia. In StatPearls. StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK430749/>
- Bender, M. T., & Niederman, M. S. (2016). Improving outcomes in community-acquired pneumonia. *Current Opinion in Pulmonary Medicine*, 22(3), 235–242.
- Kolditz, M., Scherag, A., Rohde, G., Ewig, S., Welte, T., & Pletz, M. W. (2016). Comparison of the qSOFA and CRB-65 for risk prediction in patients with community-acquired pneumonia. *Intensive Care Medicine*, 42(12), 2108–2110. <https://doi.org/10.1007/s00134-016-4517-y>