

EDITORIAL

Inteligencia artificial en urgencias: de la expectativa a la implementación

Artificial intelligence in emergency medicine: from expectation to implementation

Manuel Ramos-Casals^{1,2}

El trabajo en los servicios de urgencias suele caracterizarse por la necesidad de gestionar el caos y tomar decisiones críticas con información incompleta bajo una presión temporal implacable. En este escenario, la tecnología siempre se ha usado como un soporte pasivo (monitores, historias clínicas, sistemas de laboratorio). La rápida irrupción de la inteligencia artificial (IA) está empezando a transformar este entorno en un escenario de conversación activa y predictiva entre profesional y máquina.

La IA no es una única tecnología, sino un conjunto de enfoques y herramientas con usos distintos según el ámbito de aplicación. Las formas más “clásicas” de IA, como el aprendizaje automático (AA, *machine learning*) y las redes neuronales (RN)^{1,2}, son modelos predictivos/analíticos, a diferencia de la IA generativa, capaz de generar nuevo contenido. En medicina, los algoritmos creados por IA deben actuar como “exoesqueletos cognitivos” que amplifiquen las capacidades humanas para detectar patrones en escenarios complejos sin reemplazar el juicio clínico.

En este número de la Revista se publican dos caras de una misma moneda, IA para gestionar flujo y para refinar decisiones clínicas. En el primer artículo, Millán Soria *et al.*³ presentan un modelo de aprendizaje automático que predice en tiempo real el ingreso hospitalario integrado en un panel de control para la planificación de camas. El segundo trabajo de Gallardo Pizarro *et al.*⁴ analiza el sistema BOSS (Bacteremia Objective Score System) para estratificar el riesgo de bacteriemia en pacientes con neoplasias hematológicas atendidos en urgencias, comparando regresión logística (RL), medias K (*K-means*) y máquinas de vector soporte (MVS, *support vector machine*).

El primer trabajo³ tiene como objetivo predecir quién va a ingresar antes de que el médico firme el ingreso para preparar el recurso con antelación. El rendimiento de los algoritmos mostro que el que utilizó RL obtuvo un área bajo la curva (ABC) de 0,939, mientras que el que usó bosque aleatorio (*Random Forest*) alcanzó 0,958 (y hasta 0,97 en test), superando la mayoría

de modelos publicados (0,75-0,90)^{5,6}. El modelo RN obtuvo mayor sensibilidad (0,87), pero con menor precisión global (0,80). La solidez del trabajo reside no solo en el volumen de datos, sino en decisiones metodológicas acertadas: la agrupación temporal de variables predictoras en tres niveles (preepisodio, inicio-episodio e intraepisodio), la calibración y la ausencia de sobreajuste. La integración en un panel de control interactivo convierte el resultado estadístico en una herramienta de gestión dinámica, recalculando la predicción durante la evolución del episodio. Un hallazgo destacado del estudio es el peso diferencial de las variables predictoras, con “ubicación” y “medicación” como determinantes clave.

El estudio BOSS⁴ se sitúa en un terreno clínico de alto impacto, el paciente con neoplasia hematológica que acude a urgencias por fiebre y posible bacteriemia, donde el antibiótico empírico debe equilibrar rapidez, adecuación y prudencia. El riesgo es claramente asimétrico, tanto en el falso negativo (no tratar una bacteriemia en un paciente neutropénico puede llevar al *shock séptico*) como en el falso positivo (tratar con antibióticos de amplio espectro, alimentando la pandemia de resistencias antimicrobianas). El diseño compara tres modelados matemáticos: BOSS-1 (RL), BOSS-2 (K-medias), y BOSS-3 (MVS), una comparación directa infrecuente en la literatura^{2,7}. Los resultados revelan que mientras RL y K-medias ofrecen alta sensibilidad para identificar pacientes de bajo riesgo, MVS logra el mejor equilibrio entre sensibilidad y especificidad clasificando a los pacientes en dos grupos (bajo y alto riesgo), y obteniendo la mejor generalización en la validación externa al identificar un grupo de alto riesgo con probabilidad tres veces mayor de bacteriemia que concentró la mayoría de tratamiento empírico inadecuado.

Ambos estudios reconocen sus limitaciones. En el primero³, el ABC de 0,97 reportado en la fase de prueba es extraordinariamente alto. En predicción clínica, valores superiores a 0,90 invitan al escepticismo por posible fuga de datos (*data leakage*), aunque los auto-

Filiación de los autores: ¹Programa de Postgrado “IA generativa en Medicina y Ciencias de la Salud”, Facultad de Medicina, Universitat de Barcelona, Barcelona, España. ²Servicio de Enfermedades Autoinmunes, Institut Clínic de Medicina i Dermatologia (ICMiD), Hospital Clínic, Barcelona, España.

Contribución de los autores: El autor ha confirmado su autoría en el documento de responsabilidades del autor, acuerdo de publicación y cesión de derechos a EMERGENCIAS.

Autor para correspondencia: Manuel Ramos-Casals. Hospital Clínic. Villarroel 170. 08036 Barcelona, España.

Correo electrónico: mramos@clinic.cat

Información del artículo: Recibido: 16-2-2026. Aceptado: 17-2-2026. Online: 19-2-2026.

Editor responsable: Óscar Miró.

DOI: 10.55633/s3me/019.2026

res incluyeron validación temporal. El diseño unicéntrico y la ausencia de datos fisiológicos en tiempo real representan oportunidades de mejora, junto con la demostración del impacto sobre indicadores asistenciales como los tiempos de espera o la ocupación de camas, previsto para una segunda fase. Y es que la generalización de este tipo de modelos es compleja, ya que un algoritmo entrenado en un hospital puede fallar en otro donde los patrones de ubicación o prescripción sean diferentes. La portabilidad es baja y cada hospital necesitaría recalibrar los modelos⁵. Respecto al segundo trabajo⁴, incluye un tamaño de muestra modesto, la pérdida de datos en algunas variables y la necesidad, como los propios autores señalan, de validación prospectiva. En cambio, la validación externa multicéntrica internacional (Francia y España) es un punto fuerte, ya que, en el mundo de la IA médica, es común que los modelos sufran un “colapso de rendimiento” al salir de su hospital de origen (sobreajuste *–overfitting–* a datos locales).

Es esencial considerar el horizonte regulatorio de 2026. La entrada en vigor pleno de la Ley de Inteligencia Artificial de la Unión Europea (EU AI Act)⁸ transformará el panorama y herramientas que intervienen en el triaje o en la estratificación de riesgo vital serán clasificadas muy probablemente como sistemas de alto riesgo. Esto implicará requisitos sobre gobernanza de datos, transparencia (explicabilidad del algoritmo) y supervisión humana (*human-in-the-loop*). Además, la aceptación profesional requerirá formación, confianza y transparencia algorítmica. Es indispensable que este tipo de modelos se planifiquen desde un inicio para ser explicables, auditables y que su despliegue se acompañe de evaluaciones de impacto rigurosas.

La transformación digital de los servicios de urgencias está pasando de la fase de digitalización pasiva a la inteligencia predictiva⁹. Lo confirman los dos estudios analizados^{3,4}, que comparten fortalezas metodológicas notables como el uso de grandes cohortes de datos del mundo real, validación rigurosa y, lo que es más importante, un enfoque claro en la aplicabilidad clínica. Junto con otros estudios recientes¹⁰⁻¹³, se puede confirmar que la IA ha llegado a los servicios de urgencias no como un concepto abstracto, sino con herramientas concretas que ayudan a los profesionales a decidir mejor. El objetivo ya no va a ser cómo evaluar el uso de la IA¹⁴, sino cómo integrarla de forma segura, equitativa y eficiente, con validaciones multicéntricas y una evaluación prospectiva del impacto clínico, bajo un marco regulatorio que asegure que la IA añadirá valor a algo que a día de hoy es insustituible: el juicio clínico. Ya no cabe preguntarse si habrá IA en urgencias, sino qué problemas va a resolver, cómo se va a integrar y qué valor va a demostrar.

Conflicto de intereses: El autor declara no tener conflictos de interés en relación con el presente artículo.

Financiación: El autor declara la no existencia de financiación en relación con el presente artículo.

Responsabilidades éticas: El autor ha confirmado el mantenimiento de la confidencialidad y respeto de los derechos de los pacientes en el documento de responsabilidades del autor, acuerdo de publicación y cesión de derechos a EMERGENCIAS.

Artículo encargado y con revisión interna por el Comité Editorial.

Nota sobre el uso de la IA en este artículo: Durante la redacción de este trabajo, el autor utilizó Claude Opus 4.6 con el fin de asistir en lenguaje, gramática, numeración de las citas bibliográficas y flujo narrativo. Tras el uso de esta herramienta, el autor revisó y editó según fue necesario, y asume la plena responsabilidad del contenido de la publicación final.

Bibliografía

- Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378.
- García-Vidal C, Sanjuan G, Puerta-Alcalde P, Moreno-García E, Soriano A. Artificial intelligence to support clinical decision-making processes. *EBioMedicine*. 2019;46:27-9.
- Millán Soria J, Medina Álvarez J, Sánchez Suárez L, Cano Cerviño C, Serrano Lázaro L, García-Minguillán Castillo MC, et al. Predicción en tiempo real de los ingresos del servicio de urgencias mediante un modelo de inteligencia artificial y su integración en la planificación de las camas de hospitalización. *Emergencias*. 2026;38:98-106.
- Gallardo-Pizarro A, Peyrony O, Teijón-Lumbreras C, Castiella-Aranzasti A, Martínez-Urrea A, Terrones-Campos C, et al. Evaluación del riesgo de bacteriemia en pacientes con neoplasias hematológicas en el servicio de urgencias. Estudio comparativo entre regresión logística y algoritmos de aprendizaje automático. *Emergencias*. 2026;38:89-97.
- Larburu N, Azkue L, Kerexeta J. Predicting hospital ward admission from the emergency department: a systematic review. *J Pers Med*. 2023;13:849.
- Hong WS, Haimovich AD, Taylor RA. Predicting hospital admission at emergency department triage using machine learning. *PLoS One*. 2018;13:e0201016.
- Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med*. 2022;28:31-8.
- Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial). *Diario Oficial de la Unión Europea*. L 2024/1689. (Consultado 10 Febrero 2026). Disponible en: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Castro-Delgado R, Pardo Ríos M. La inteligencia artificial y los servicios de urgencias y emergencias: debemos dar un paso adelante. *Emergencias*. 2024;36:145-7.
- Tahir A, Hassan M, Farooq A, Ashfaq S, Ayub M, Hameed I, et al. In-hospital mortality prognostication for cancer patients with febrile neutropenia. *Commun Med*. 2025;5:460.
- Long H, Huang G, Yin X, Zheng X, Cao S, Wang N, et al. ChatGPT-4o assists emergency physicians in enhancing diagnostic accuracy for fever of unknown origin. *BMC Emerg Med*. 2026;26:45.
- Gallardo-Pizarro A, Teijón-Lumbreras C, Monzo-Gallo P, Aiello TF, Chumbita M, Peyrony O, et al. Towards personalised empirical antibiotic therapy in febrile neutropenia: a theoretical model based on machine learning and prior colonisation with MDR gram-negative bacilli. *Ther Adv Infect Dis*. 2026;13:20499361251411061.
- Agius S, Cassar V, Magri C, Khan W, Obe DA, Caruana G, et al. Predicting Emergency Severity Index (ESI) level, hospital admission, and admitting ward in an emergency department using data-driven machine learning. *BMC Med Inform Decis Mak*. 2025;25:281.
- Sáenz-Abad D, Sachi Martínez-Mihara M, Lahoza-Pérez MC. La inteligencia artificial como herramienta de apoyo diagnóstico en urgencias. *Emergencias*. 2025;37:78-9.