

Inteligencia artificial en el análisis predictivo de patologías frecuentes en medicina interna, ginecología y pediatría

Junior Andretti Melo Villamizar¹, Joyce Beatriz Solórzano Griego², Laura Natalia Cárdenas³, Lina Maryudi Rodríguez López⁴

¹ Junior Andretti Melo Villamizar, Universidad Antonio Nariño, juniormelo96@hotmail.com

² Joyce Beatriz Solórzano Griego, Corporación Universitaria Rafael Núñez, joycesolorzanogriego@gmail.com

³ Laura Natalia Cárdenas Castro, Universidad El Bosque, castronatic2001@gmail.com

⁴ Lina Maryudi Rodríguez López, Universidad del Tolima, linismrl@gmail.com

Historia del Artículo:

Recibido: Junio 2025

Aceptado: Diciembre 2025

Publicado: Enero de 2026

Palabras Clave:

Inteligencia Artificial, Aprendizaje Automático, Modelos Predictivos, Toma de Decisiones Clínicas, Medicina Interna, Ginecología y Obstetricia, Pediatría.

Keywords:

Artificial Intelligence, Machine Learning, Predictive Models, Clinical Decision-Making, Internal Medicine, Gynecology and Obstetrics, Pediatrics.

Resumen

La inteligencia artificial (IA) ha emergido como una herramienta relevante en el ámbito de la salud, especialmente en el desarrollo de modelos predictivos que apoyan la toma de decisiones clínicas. Este estudio tuvo como objetivo analizar, mediante una revisión sistemática, el papel de la inteligencia artificial en el desarrollo de modelos predictivos aplicados a patologías frecuentes en Medicina Interna, Ginecología/Obstetricia y Pediatría. A partir de una revisión sistemática de la literatura, se evalúa el rendimiento, la validación y la utilidad clínica de estos modelos, destacando los principales enfoques metodológicos utilizados y su aplicación en distintos contextos asistenciales. Asimismo, se examinan los avances, limitaciones y desafíos asociados a su implementación, evidenciando el potencial de la inteligencia artificial como herramienta para mejorar la estratificación del riesgo y apoyar la toma de decisiones clínicas.

Abstract

Artificial intelligence (AI) has emerged as a relevant tool in the field of healthcare, particularly in the development of predictive models that support clinical decision-making. This study aimed to analyze, through a systematic review, the role of artificial intelligence in the development of predictive models applied to common pathologies in Internal Medicine, Gynecology/Obstetrics, and Pediatrics. Based on a systematic review of the literature, the performance, validation, and clinical utility of these models are evaluated, highlighting the main methodological approaches used and their application across different healthcare settings. Additionally, the advances, limitations, and challenges associated with their implementation are examined, demonstrating the potential of artificial intelligence as a tool to improve risk stratification and support clinical decision-making.

* Autor para correspondencia:

Junior Andretti Melo Villamizar, Universidad Antonio Nariño, juniormelo96@hotmail.com

Cómo citar:

Melo et al. Inteligencia artificial en el análisis predictivo de patologías frecuentes en medicina interna, ginecología y pediatría. S&EMJ. Año 2026; Vol.6: 34-53.

Introducción

La inteligencia artificial (IA) ha adquirido un papel creciente en el ámbito de la salud, particularmente en el desarrollo de modelos predictivos orientados a apoyar la toma de decisiones clínicas. El aumento en la disponibilidad de datos clínicos y los avances en aprendizaje automático (ML) y aprendizaje profundo han impulsado su aplicación en diversas áreas médicas, aunque con una evidencia aún heterogénea.

Introduction

Artificial intelligence (AI) has gained an increasingly important role in the healthcare field, particularly in the development of predictive models aimed at supporting clinical decision-making. The growing availability of clinical data, along with advances in machine learning (ML) and deep learning, has driven its application across various medical fields, although the existing evidence remains heterogeneous.

Objetivo

Analizar, mediante una revisión sistemática, la evidencia disponible sobre el rendimiento, la validación y la utilidad clínica de los modelos predictivos basados en IA aplicados a patologías frecuentes en Medicina Interna, Ginecología/Obstetricia y Pediatría.

Objective

To analyze, through a systematic review, the available evidence on the performance, validation, and clinical utility of AI-based predictive models applied to common pathologies in Internal Medicine, Gynecology/Obstetrics, and Pediatrics.

Métodología

Se realizó una revisión sistemática de la literatura con el objetivo de analizar la evidencia sobre el rendimiento y la utilidad clínica de modelos predictivos basados en inteligencia artificial en patologías frecuentes de Medicina Interna, Ginecología/Obstetricia y Pediatría, siguiendo las recomendaciones de la guía PRISMA 2020. La búsqueda se llevó a cabo en PubMed/MEDLINE y ScienceDirect, utilizando términos MeSH y palabras clave relacionadas con inteligencia artificial, aprendizaje automático, predicción clínica y diagnóstico temprano, combinadas con operadores booleanos. Se incluyeron artículos de los últimos cinco años, en inglés o español, con texto completo disponible, excluyendo literatura gris.

Se consideraron estudios primarios originales que aplicaran modelos de inteligencia artificial con objetivos predictivos y reportaran métricas cuantitativas de desempeño. Se excluyeron revisiones, metaanálisis, editoriales, duplicados, estudios sin

aplicación real de IA, sin enfoque predictivo, sin validación adecuada o fuera de las áreas clínicas definidas. La selección se realizó mediante revisión de títulos, resúmenes y texto completo, documentada con diagrama PRISMA. Se extrajeron datos clínicos y metodológicos, evaluando la calidad con PROBAST y QUADAS-2. Debido a la heterogeneidad, los resultados se sintetizaron de forma descriptiva.

La búsqueda identificó 457 registros; tras eliminar 220 duplicados, quedaron 237 para cribado. Se excluyeron 80 por no cumplir criterios, principalmente por falta de enfoque en IA o predicción, quedando 157 estudios potencialmente relevantes, de los cuales 94 fueron evaluados a texto completo.

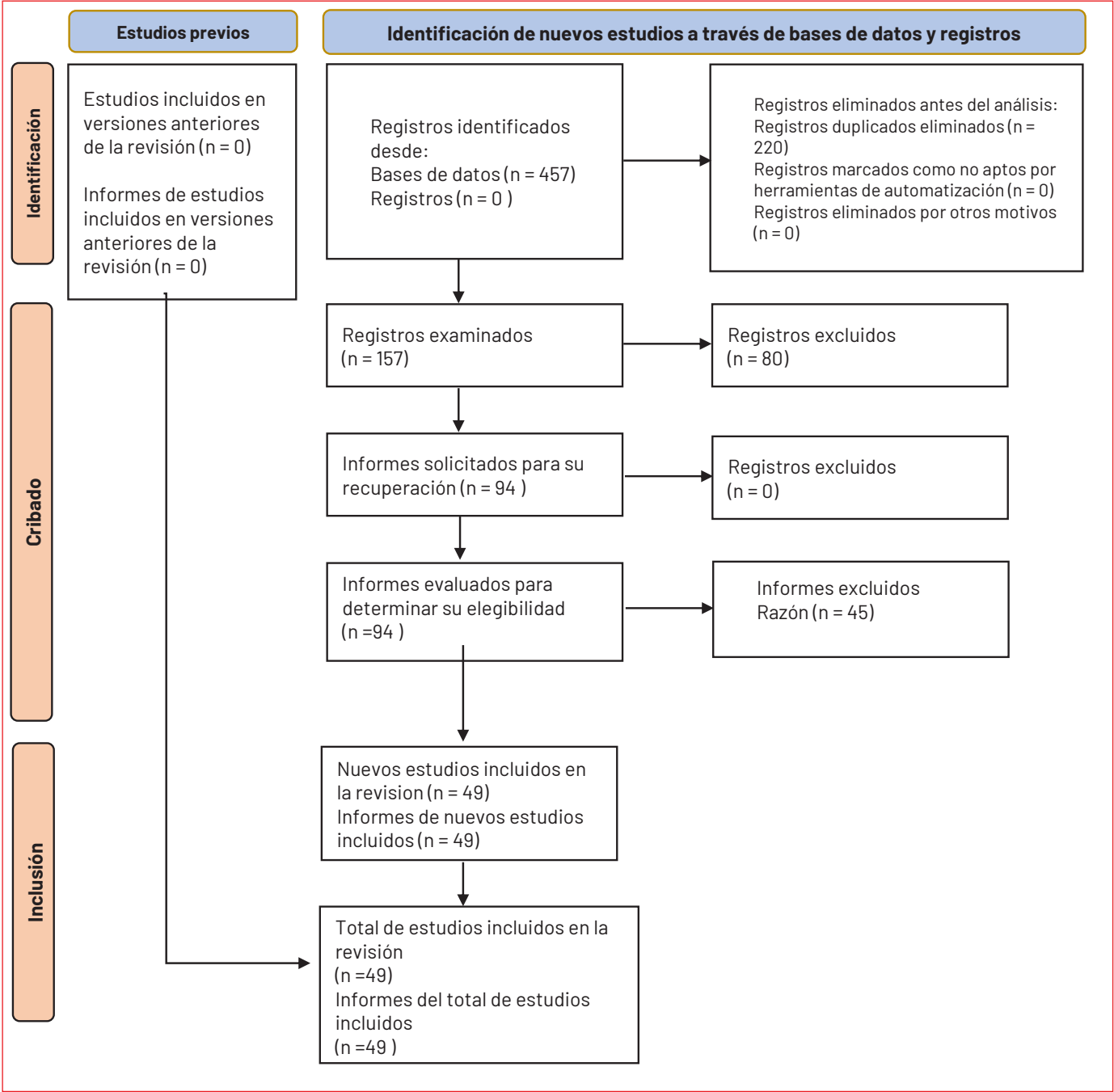
Tras la evaluación detallada de los 94 artículos, se excluyeron 45 estudios, principalmente por no reportar métricas cuantitativas de desempeño, presentar implementaciones de IA sin validación adecuada o no alinearse con los objetivos específicos de la revisión. Finalmente, 49 estudios primarios cumplieron con todos los criterios de inclusión y fueron incorporados en la síntesis cualitativa (1-49).

El proceso de selección de los estudios se resume en el diagrama de flujo PRISMA presentado en la **Figura 1**, el cual describe de manera transparente las etapas de identificación, cribado, elegibilidad e inclusión de los estudios analizados.

La **Figura 2** presenta un esquema metodológico de la revisión sistemática que sintetiza de manera visual las principales fases del estudio. En este esquema se resumen el diseño de la revisión conforme a PRISMA 2020, la identificación de estudios mediante la búsqueda sistemática en bases de datos especializadas, el proceso de selección y cribado de los registros, la recopilación de variables clínicas y metodológicas relevantes, la evaluación de la calidad metodológica de los estudios incluidos y el enfoque de síntesis de los resultados. Esta representación gráfica permite una comprensión integral y secuencial del proceso metodológico seguido en la revisión.

Los estudios incluidos en esta revisión sistemática evidencian una amplia diversidad de contextos clínicos, poblaciones y enfoques metodológicos, lo que refleja el carácter transversal de la inteligencia artificial en la predicción clínica. Las investigaciones se desarrollaron en distintas regiones (Asia, Europa, África y Norteamérica) y correspondieron principalmente a estudios observacionales retrospectivos, con menor representación de diseños prospectivos. Asimismo, los tamaños muestrales fueron heterogéneos, abarcando desde cohortes pequeñas hasta grandes bases de datos poblacionales y registros electrónicos de salud, permitiendo evaluar el desempeño de los modelos en escenarios asistenciales diversos.

Figura 1. Diagrama de flujo PRISMA del proceso de selección de estudios



Fuente: elaboración propia a partir de Guías prisma 2020.

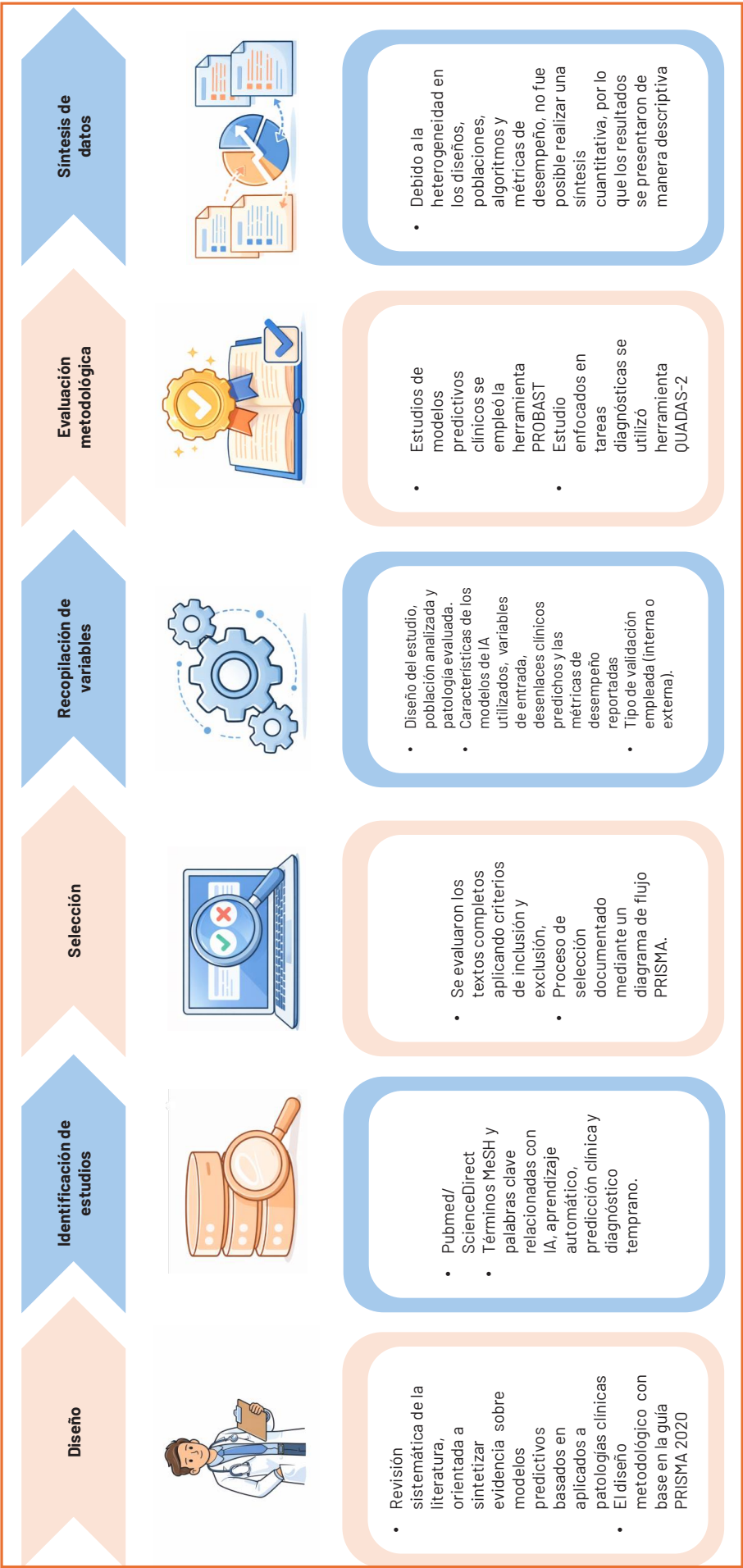
En cuanto a su distribución por áreas clínicas, predominó la medicina interna, seguida de ginecología y obstetricia, y la pediatría. Este patrón sugiere un mayor desarrollo de modelos de IA en patologías de alta carga asistencial y amplia disponibilidad de datos, como diabetes mellitus, hipertensión arterial, insuficiencia cardíaca aguda y EPOC. En Ginecología y Obstetricia se evidencia un interés creciente en la predicción temprana de complicaciones maternas y perinatales, como la preeclampsia y el parto pretérmino, destacando su potencial para mejorar la estratificación del riesgo y la toma de decisiones clínicas. (1-49).

En Pediatría, los modelos de IA se orientan

principalmente a la predicción temprana de alteraciones nutricionales y patologías agudas graves, como sepsis, infección urinaria, exacerbaciones asmáticas y apendicitis. Sin embargo, su aplicación enfrenta desafíos relacionados con las trayectorias de crecimiento, la heterogeneidad etaria, la menor disponibilidad de datos y consideraciones éticas y de privacidad, lo que podría explicar su menor desarrollo en comparación con otras áreas. (50) **Figura 3.**

La calidad metodológica y el riesgo de sesgo de los estudios incluidos fueron evaluados mediante herramientas validadas según el tipo de modelo y desenlace. En los estudios de modelos predictivos

Figura 2. Esquema metodológico de la revisión sistemática: resume las etapas de diseño del estudio, identificación y selección de los estudios, recopilación de variables, evaluación metodológica y síntesis de los datos incluidos.



Fuente: Elaboración propia a partir de estudio incluidos en la revisión sistemática.

clínicos se utilizó la herramienta PROBAST, mientras que en aquellos enfocados en tareas diagnósticas se empleó QUADAS-2 (51,52).

Las patologías abordadas incluyeron, entre otras, sepsis asociada a lesión renal (1,2), COVID-19 (3), ICA (4,40,41) diabetes mellitus (5-9,49), hipertensión arterial (10-13), EPOC (14-19), preeclampsia (24-27), cáncer cervicouterino (35-39), infección de tracto urinario (ITU) pediátrica (44), asma pediátrica (20,42,43), obesidad (46,47), desnutrición (45,48), aborto (28-34), sepsis neonatal (22,23) y apendicitis pediátrica (21). Los tamaños de la muestra variaron ampliamente entre los estudios, desde cohortes clínicas pequeñas hasta bases de datos poblacionales de gran escala, lo que permitió evaluar el desempeño de los modelos en distintos escenarios asistenciales.

Respecto a las técnicas de IA empleadas, predominó el uso de algoritmos de ML supervisado, tales como bosques aleatorios, *boosting* por gradiente, XGBoost y redes neuronales profundas, con una validación mayoritariamente interna mediante partición de los datos o validación cruzada. Las métricas de desempeño más frecuentemente reportadas fueron el AUC, la sensibilidad, la especificidad y la exactitud, reflejando un enfoque centrado en la evaluación del rendimiento predictivo de los modelos.

En coherencia con esta distribución por áreas clínicas y con las características de los datos disponibles en cada especialidad, la **Figura 4** presenta la distribución de los enfoques computacionales y de los algoritmos de inteligencia artificial utilizados en los estudios incluidos. Se observó un predominio de modelos de aprendizaje automático tradicional, entre los que destacaron los bosques aleatorios, XGBoost, LightGBM, las máquinas de vectores de soporte (SVM) y la regresión logística, seleccionados principalmente por su capacidad para manejar conjuntos de datos clínicos estructurados y por su relativa interpretabilidad. En contraste, los modelos de aprendizaje profundo, como las redes neuronales profundas, las redes neuronales convolucionales (CNN), las redes recurrentes (RNN) y las arquitecturas tipo *Transformers*, se emplearon con menor frecuencia y principalmente en escenarios que involucraron grandes volúmenes de datos o información de mayor complejidad, como imágenes médicas o señales fisiológicas. Esta menor adopción podría estar relacionada con mayores requerimientos computacionales, así como con limitaciones en la interpretabilidad y en la implementación clínica en entornos asistenciales (53).

Methodology

A systematic review of the literature was conducted to analyze the evidence on the performance and clinical utility of artificial intelligence-based predictive models in common pathologies in Internal Medicine,

Gynecology/Obstetrics, and Pediatrics, following the PRISMA 2020 guidelines. The search was carried out in PubMed/MEDLINE and ScienceDirect using MeSH terms and keywords related to artificial intelligence, machine learning, clinical prediction, and early diagnosis, combined with Boolean operators. Articles published in the last five years, in English or Spanish, with full text available were included, while gray literature was excluded. Original primary studies applying artificial intelligence models with predictive purposes and reporting quantitative performance metrics were considered. Reviews, meta-analyses, editorials, duplicates, studies without real AI application, without a predictive focus, lacking adequate validation, or outside the defined clinical areas were excluded. Study selection was performed through title, abstract, and full-text screening, documented using a PRISMA flow diagram. Clinical and methodological data were extracted, and study quality was assessed using PROBAST and QUADAS-2. Due to heterogeneity, results were synthesized descriptively.

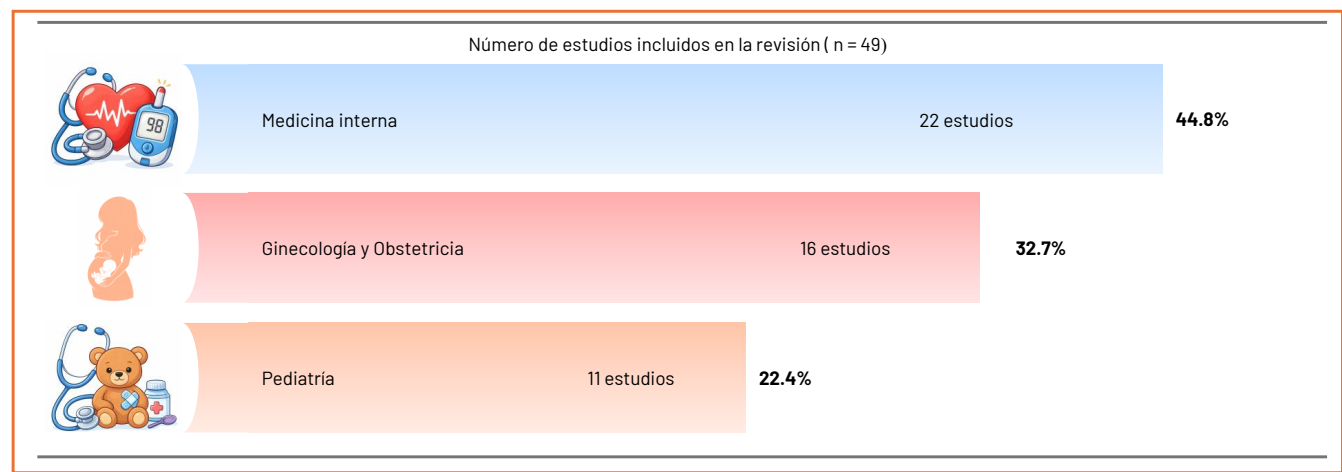
The search identified 457 records; after removing 220 duplicates, 237 remained for screening. Of these, 80 were excluded for not meeting inclusion criteria, mainly due to lack of focus on AI or prediction, leaving 157 potentially relevant studies, of which 94 were assessed in full text.

Following detailed evaluation, 45 studies were excluded, primarily for not reporting quantitative performance metrics, lacking proper validation, or not aligning with the specific objectives of the review. Ultimately, 49 primary studies met all inclusion criteria and were included in the qualitative synthesis (1-49).

The study selection process is summarized in the PRISMA flow diagram presented in Figure 1, which transparently outlines the stages of identification, screening, eligibility, and inclusion of the analyzed studies. Figure 2 presents a methodological framework of the systematic review, visually summarizing the main phases of the study, including study design according to PRISMA 2020, identification of studies through systematic database searches, screening and selection processes, extraction of relevant clinical and methodological variables, quality assessment, and data synthesis. This graphical representation provides a comprehensive and sequential understanding of the methodological process followed in the review.

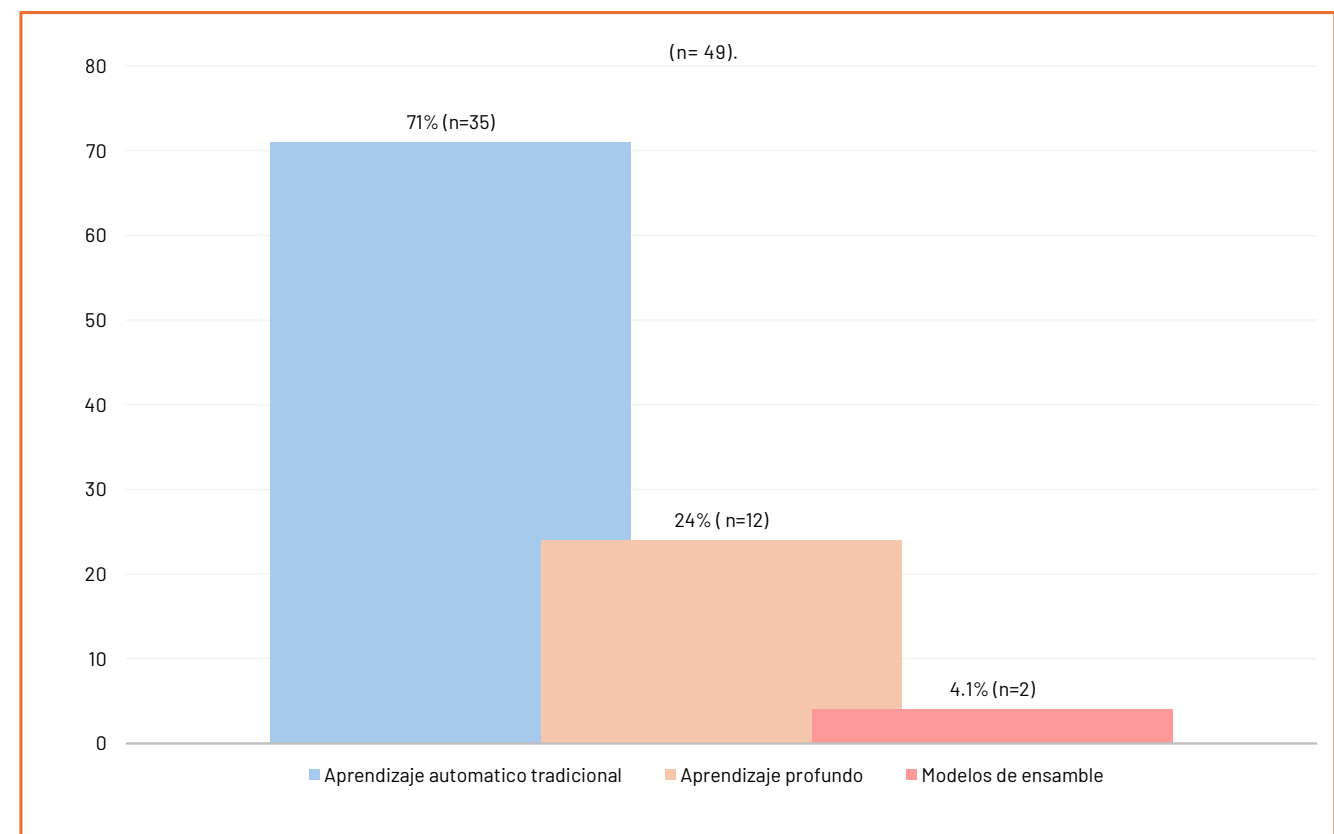
The studies included in this systematic review demonstrate a wide diversity of clinical contexts, populations, and methodological approaches, reflecting the cross-cutting nature of artificial intelligence in clinical prediction. The studies were conducted across different regions (Asia, Europe, Africa, and North America) and were predominantly retrospective observational studies, with fewer prospective designs. Sample sizes were heterogeneous, ranging from small

Figura 3. Distribución de los estudios incluidos según el área clínica: corresponde a estudios de Medicina interna, Ginecología y Obstetricia y Pediatría, expresados en frecuencia absoluta y porcentaje respecto al total de estudios incluidos en la revisión sistemática.



Fuente: Elaboración propia a partir de estudio incluidos en la revisión sistemática.

Figura 4: Distribución de los enfoques computacionales de inteligencia artificial en los estudios incluidos



Fuente: Elaboración propia a partir de estudio incluidos en la revisión sistemática.

clinical cohorts to large population-based datasets and electronic health records, allowing evaluation of model performance across diverse healthcare settings.

Regarding clinical distribution, Internal Medicine predominated, followed by Gynecology and Obstetrics, and to a lesser extent Pediatrics. This pattern suggests

greater development of AI models in high-burden conditions with broad data availability, such as diabetes mellitus, hypertension, acute heart failure, and chronic obstructive pulmonary disease (COPD). In Gynecology and Obstetrics, there is growing interest in the early prediction of maternal and perinatal complications, such as preeclampsia and preterm birth, highlighting

their potential to improve risk stratification and clinical decision-making (1-49).

In Pediatrics, AI models are primarily focused on the early prediction of nutritional disorders and severe acute conditions, such as sepsis, urinary tract infection, asthma exacerbations, and appendicitis. However, their application faces challenges related to growth trajectories, age heterogeneity, limited data availability, and ethical and privacy considerations, which may explain their relatively lower development compared to other fields (50).

The methodological quality and risk of bias of the included studies were assessed using validated tools according to the model type and outcomes. PROBAST was used for clinical prediction models, while QUADAS-2 was applied to diagnostic studies (51,52).

The pathologies addressed included, among others, sepsis associated with kidney injury (1,2), COVID-19 (3), acute heart failure (AHF) (4,40,41), diabetes mellitus (5-9,49), arterial hypertension (10-13), chronic obstructive pulmonary disease (COPD) (14-19), preeclampsia (24-27), cervical cancer (35-39), pediatric urinary tract infection (UTI) (44), pediatric asthma (20,42,43), obesity (46,47), malnutrition (45,48), abortion (28-34), neonatal sepsis (22,23), and pediatric appendicitis (21).

Sample sizes varied widely across studies, ranging from small clinical cohorts to large-scale population databases, allowing for the evaluation of model performance across different healthcare settings.

Regarding the artificial intelligence techniques employed, the use of supervised machine learning algorithms predominated, including random forests, gradient boosting, XGBoost, and deep neural networks, with validation conducted primarily through internal methods such as data splitting or cross-validation. The most frequently reported performance metrics were the area under the curve (AUC), sensitivity, specificity, and accuracy, reflecting a focus on evaluating the predictive performance of the models.

Medicina del siglo XXI: el papel emergente de la inteligencia artificial

Desde una perspectiva evolutiva, la capacidad de adaptación ha sido históricamente un factor clave para la supervivencia. En salud, esta idea puede trasladarse al desarrollo de herramientas capaces de responder a entornos clínicos complejos y cambiantes. En este contexto, los modelos predictivos basados en IA han cobrado relevancia porque pueden aprender de grandes volúmenes de información, integrar múltiples variables y ajustarse a nuevas condiciones clínicas, lo que los convierte en una alternativa prometedora para apoyar la toma de decisiones (54).

En los últimos años, la IA ha adquirido un papel creciente en el ámbito de la salud, particularmente en el desarrollo de modelos predictivos orientados a apoyar la toma de decisiones clínicas. El aumento en la sistematización y disponibilidad de los datos clínicos, junto con los avances en técnicas de aprendizaje automático y aprendizaje profundo, ha permitido el diseño de herramientas capaces de identificar patrones, estimar riesgos y anticipar desenlaces clínicos relevantes en distintos contextos asistenciales (53,55).

En áreas como Medicina Interna, Ginecología/Obstetricia y Pediatría, se han desarrollado múltiples modelos predictivos basados en IA con el objetivo de apoyar la identificación temprana del riesgo, la estratificación de pacientes y la predicción de desenlaces clínicos relevantes. Estas aplicaciones abarcan tanto enfermedades crónicas de alta carga asistencial como condiciones agudas y escenarios materno-infantiles, en los que la anticipación del riesgo puede tener un impacto directo en la oportunidad y calidad de la atención. En conjunto, estos enfoques reflejan el interés creciente por integrar herramientas de IA en la práctica clínica como apoyo a la toma de decisiones (55).

Sin embargo, el crecimiento acelerado de la literatura ha dado lugar a una evidencia heterogénea. Existen diferencias relevantes entre estudios en cuanto a diseños metodológicos, características poblacionales, algoritmos utilizados y métricas de desempeño reportadas. Esta variabilidad dificulta la comparación directa entre modelos y plantea interrogantes sobre su rendimiento real y su aplicabilidad clínica en escenarios asistenciales diversos (2).

Aplicaciones de la inteligencia artificial en la predicción clínica en Medicina Interna

Diabetes mellitus

Los estudios incluidos evaluaron modelos de aprendizaje automático y profundo aplicados a datos clínicos rutinarios para la detección temprana de diabetes, la predicción de complicaciones y el uso de servicios de salud (5-9,49).

Para la detección temprana, los modelos alcanzaron AUC entre 0,69 y 0,74 utilizando variables clínicas frecuentes, como hipertensión arterial, obesidad, hiperlipidemia y número de consultas médicas. **Algoritmo 1.** (9).

En relación con las complicaciones crónicas, los modelos orientados a neuropatía y nefropatía diabética reportaron AUC superiores a 0,70, mientras que los modelos basados en XGBoost lograron AUC de hasta 0,98 en la predicción de amputación por pie diabético (5-7). Asimismo, algunos estudios demostraron un

desempeño estable en la predicción del uso intensivo de servicios sanitarios en cohortes con validación externa (6).

En adultos mayores, modelos de *gradient boosting* mantuvieron AUC cercanos a 0,80 al integrar factores clínicos y conductuales modificables, reforzando su utilidad para la estratificación temprana del riesgo en atención primaria (49).

Hipertensión arterial

Los estudios sobre hipertensión arterial aplicaron modelos de aprendizaje automático y profundo para la predicción y el diagnóstico, empleando registros electrónicos de salud y variables fisiológicas (10-13), los modelos prediagnósticos mostraron valores de AUC cercanos a 0,75, identificando dislipidemia, obesidad y frecuencia de contacto con servicios de salud como predictores relevantes (10).

Otros estudios evaluaron la identificación de hipertensión enmascarada y no controlada mediante modelos basados en bosques aleatorios y *boosting* por gradiente, con elevada sensibilidad y valor predictivo negativo (12). En escenarios más complejos, los modelos de aprendizaje profundo integraron datos estructurados y no estructurados, alcanzando F1-scores superiores a 0,89 en el diagnóstico diferencial de hipertensión secundaria (13).

Enfermedad renal aguda

Dos estudios evaluaron modelos de ML para predecir mortalidad en pacientes con sepsis asociada a lesión renal aguda (SA-AKI) usando bases de datos MIMIC-IV y eICU. En el primero, el modelo bosques aleatorios mostró el mejor desempeño para mortalidad intrahospitalaria (AUC $\approx$ 0,80), identificando como predictores claves la edad, el puntaje GCS, el estadio de AKI, la creatinina, la diuresis y el SAPS II (1).

El segundo estudio comparó el inicio temprano versus tardío de la terapia de reemplazo renal continua (CRRT). Tras ajuste por puntaje de propensión, el inicio tardío se asoció con mayor supervivencia, aunque con estancias hospitalarias más prolongadas. El modelo GBM presentó un C-índice $\approx$ 0,69 y AUC $\approx$ 0,73 en validación externa, siendo el lactato máximo, la edad y la SpO₂ mínima los principales predictores de mortalidad (2).

Insuficiencia cardiaca aguda

Tres estudios evaluaron modelos de aprendizaje automático en pacientes hospitalizados por insuficiencia cardiaca aguda, principalmente en unidades de cuidados intensivos, utilizando cohortes retrospectivas de la base MIMIC-IV (4,40,41).

En todos los casos, los modelos basados en XGBoost

mostraron el mejor desempeño predictivo, con AUC entre 0,72 y 0,82 para la mortalidad intrahospitalaria, superando a escalas clínicas tradicionales como ADHERE, GWTG-HF, SOFA, SAPS II y LODS (40,41).

Para la predicción de readmisión a 30 días, XGBoost alcanzó un AUC de 0,76 y una sensibilidad de 0,66, permitiendo una estratificación del riesgo clínicamente relevante (4). Las variables más influyentes incluyeron edad, duración de la hospitalización, volumen urinario, presión arterial, función renal, parámetros respiratorios y uso de vasopresores, reflejando la gravedad multiorgánica de estos pacientes (4,40,41).

Enfermedad pulmonar obstructiva crónica

Los estudios incluidos evaluaron modelos aplicados a la predicción de exacerbaciones, severidad, hospitalización y mortalidad en EPOC (14-19). Los modelos de aprendizaje profundo mostraron valores de AUC superiores a los obtenidos mediante enfoques estadísticos tradicionales, especialmente en la predicción de exacerbaciones y mortalidad a largo plazo. La integración de datos clínicos, radiológicos y administrativos permitió una discriminación moderada a buena del riesgo clínico (14-19).

COVID-19

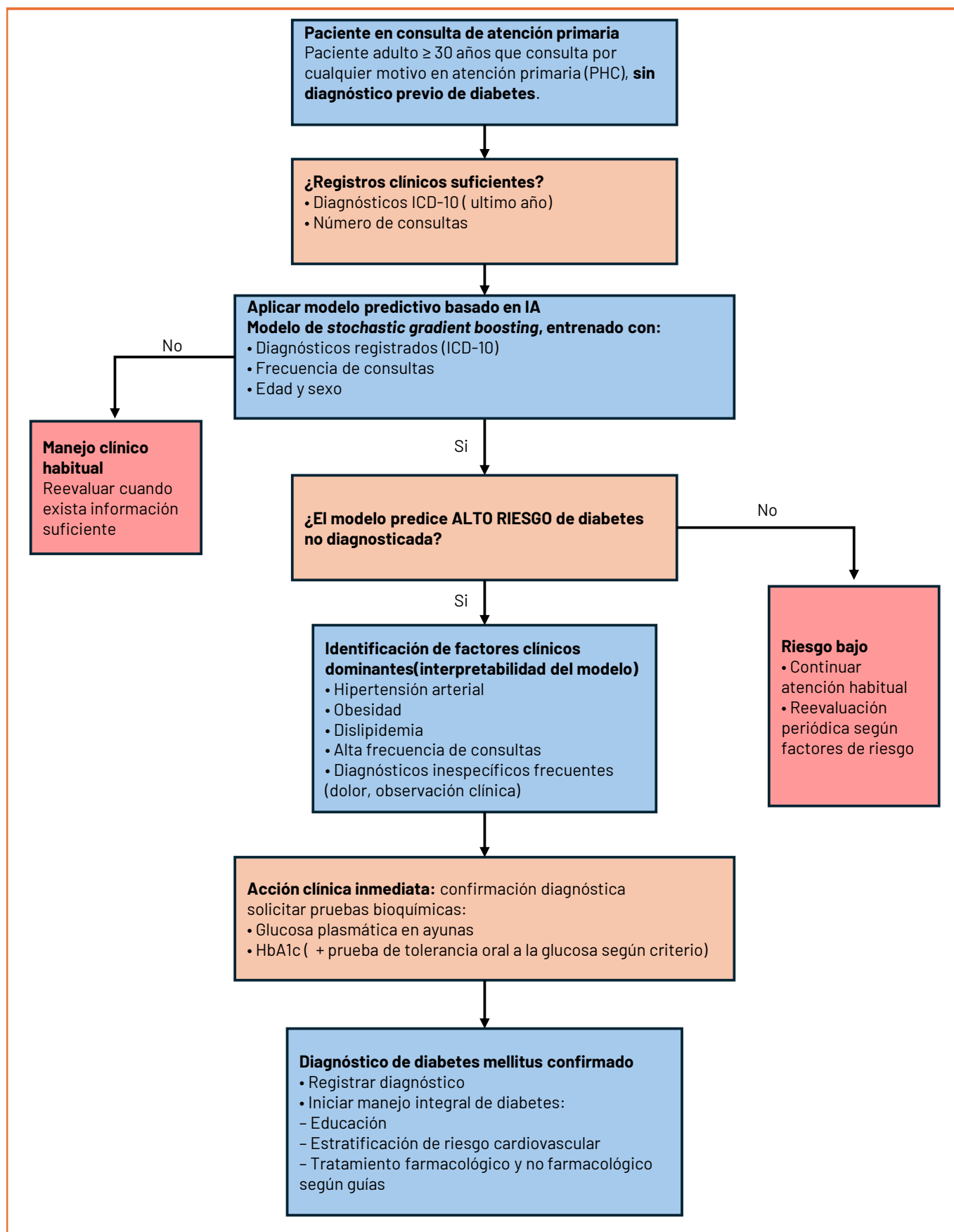
Un estudio evaluó modelos de aprendizaje automático para identificar pacientes hospitalizados con COVID-19 en alto riesgo de mortalidad utilizando exclusivamente información disponible al ingreso. En una cohorte de 824 pacientes, se analizaron 19 modelos, incluyendo enfoques supervisados y no supervisados, demostrando la viabilidad de estas herramientas para la estratificación temprana del riesgo clínico.

Aplicaciones de la inteligencia artificial en la predicción clínica en Pediatría

Desnutrición

Ambos estudios evaluaron el uso de modelos de ML para la predicción de desnutrición infantil, utilizando tanto bases de datos poblacionales como contextos clínicos reales (45,48). En conjunto, los modelos basados en aprendizaje en conjunto y algoritmos de árboles (bosques aleatorios y *boosting* por gradiente) mostraron el mejor desempeño predictivo.

En el análisis poblacional con datos DHS de Egipto, bosques aleatorios y *boosting* por gradiente alcanzaron exactitudes superiores al 90% y valores de AUC mayores a 0,95, superando consistentemente a modelos lineales y basados en distancia como KNN (48). De forma complementaria, en el contexto clínico de tratamiento ambulatorio de desnutrición aguda en Níger, un modelo de aprendizaje en conjunto logró un ajuste superior, enfrente a modelos individuales y una AUC de 0,78 para identificar niños con bajo aumento de

Algoritmo 1: Aplicación clínica de inteligencia artificial para la identificación de diabetes no diagnosticada en atención primaria.

Fuente: Elaboración propia de los autores para fines de este estudio a partir de (12).

peso, con una tasa de acierto del 84% en el grupo de mayor riesgo (45). En ambos estudios, los principales predictores incluyeron factores socioeconómicos, maternos, ambientales (agua y saneamiento) y características perinatales (45,48).

Obesidad

Dos estudios evaluaron el uso de modelos de ML para la predicción temprana de obesidad infantil a partir de datos longitudinales. En un estudio basado en un panel nacional, se desarrolló un modelo predictivo para obesidad a los 10 años mediante regresión logística con selección de variables por LASSO, alcanzando un AUC de 0,82 y una exactitud del 76%. Los principales predictores incluyeron el IMC previo del niño, el nivel de actividad física y variables maternas como el IMC y la autoestima (47).

En un estudio basado en registros electrónicos de salud, se compararon siete algoritmos de ML en una cohorte de gran tamaño. El modelo XGBoost mostró el mejor desempeño, con un AUC de 0,81 y una exactitud aproximada del 66% al fijar la sensibilidad en 80%. Entre las variables más influyentes se identificaron mediciones antropométricas tempranas y parámetros fisiológicos como la circunferencia cefálica y signos vitales registrados en los primeros dos años de vida (46).

Asma en la infancia

Los estudios incluidos evaluaron modelos de IA y ML para predecir desenlaces clínicos relevantes en asma pediátrica, utilizando datos de registros clínicos electrónicos y bases administrativas (20,42,43). Los modelos aplicados en servicios de urgencias para predecir hospitalización mostraron un alto rendimiento, con AUC entre 0,91 y 0,94, superando a modelos tradicionales y permitiendo una identificación temprana de pacientes de alto riesgo (43).

En modelos poblacionales y longitudinales, la predicción de exacerbaciones graves y del pronóstico del asma (remisión vs persistencia) alcanzó un desempeño moderado a alto, con AUC de 0,72–0,93, dependiendo del desenlace y del horizonte temporal analizado. Los predictores más relevantes fueron los antecedentes de exacerbaciones, el uso previo de medicación y la historia clínica longitudinal, mientras que variables sociales y demográficas aportaron un beneficio limitado en la capacidad predictiva global, con abordaje clínico como es descrito en el **Algoritmo 2** (20,42).

Infección de tracto urinario

El análisis del estudio incluido muestra que los modelos de ML permiten una predicción precisa del riesgo de ITU en niños febriles menores de 3 años. A

partir de 1.556 pacientes, la selección de variables mediante regresión LASSO identificó siete predictores relevantes: edad, recuento de leucocitos, antecedentes de ITU, plaquetas, pico de fiebre, proteína C reactiva y anomalías renales prenatales, con ello entre los algoritmos evaluados, el modelo bosques aleatorios presentó el mejor rendimiento diagnóstico, con un AUC de 0,88 y un AUPRC de 0,82, superando a otros modelos y a la regresión logística. El análisis explicativo mediante SHAP permitió identificar la contribución individual de cada predictor en la estimación del riesgo de ITU (44).

Apendicitis aguda

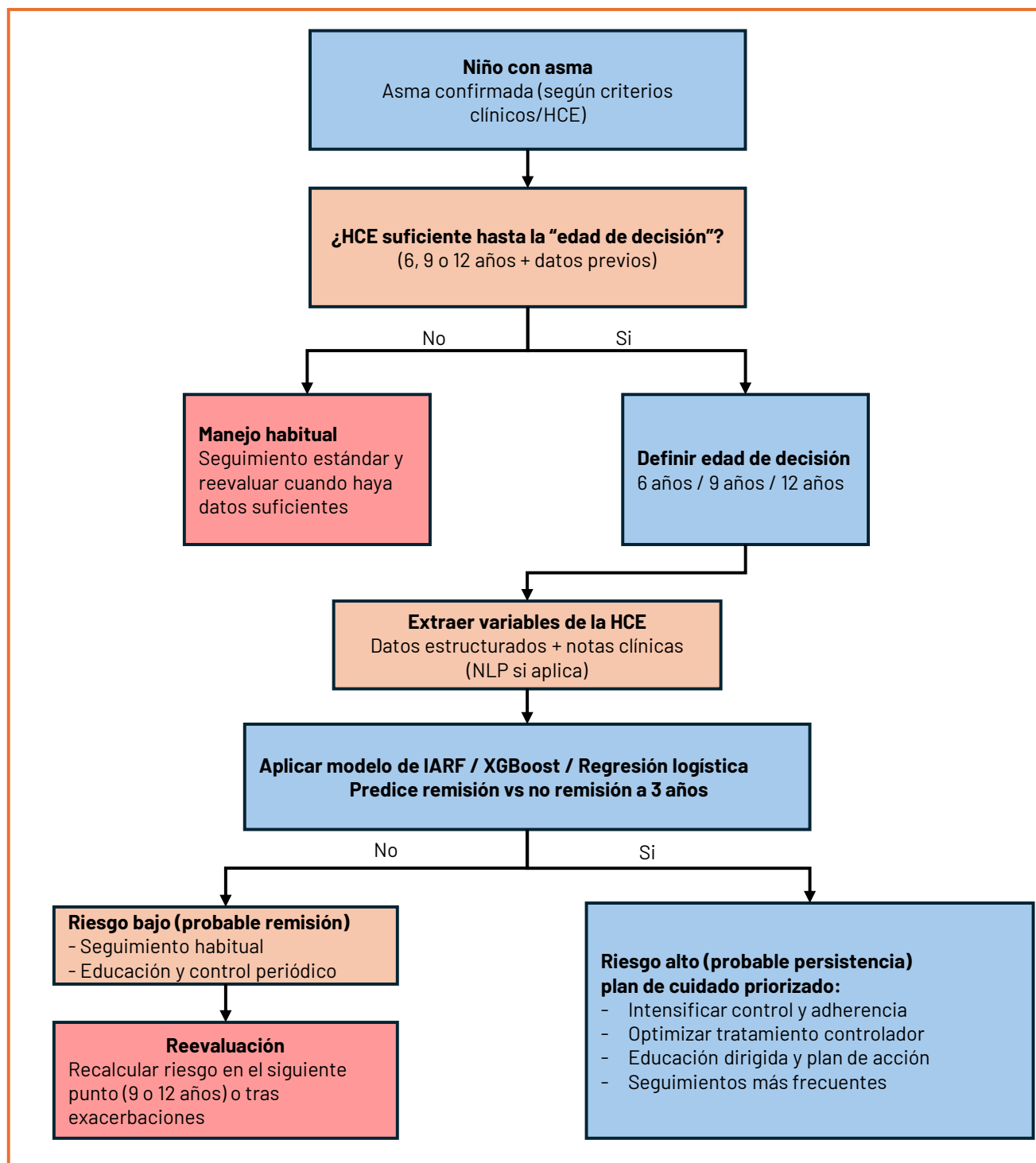
Un estudio multicéntrico internacional validó el uso de una plataforma de IA para el diagnóstico de apendicitis aguda utilizando datos clínicos, analíticos y ecográficos, con confirmación histopatológica como estándar de referencia (21).

La plataforma mostró alto desempeño diagnóstico, con sensibilidad del 92,2 %, especificidad del 97,2 % y valor predictivo negativo del 98,7 %. El rendimiento global fue consistente, con un AUC de 0,97, *F1-score* de 0,89 y coeficiente de Matthews de 0,88, indicando una elevada capacidad discriminativa, evidenciaron un beneficio en reducción sustancial de apendicectomías innecesarias, lo que sugiere su utilidad como herramienta de apoyo diagnóstico, especialmente en contextos con acceso limitado a Tomografía (21).

Sepsis neonatal

Dos estudios evaluaron la aplicación de modelos de IA en neonatos con sepsis atendidos en unidades de cuidados intensivos neonatales, abordando tanto la detección temprana de la enfermedad como la predicción de mortalidad hospitalaria (22,23). En uno de los estudios, se desarrollaron modelos de ML utilizando datos clínicos y de laboratorio de una cohorte de 388 neonatos, en los cuales los algoritmos *Bagging* y bosques aleatorios mostraron el mejor desempeño, con exactitudes cercanas al 98,4 % y valores de AUC superiores a 0,98. Entre las variables con mayor relevancia predictiva se identificaron el bajo peso al nacer, la prematuridad y la proteína C reactiva elevada (22).

En el otro estudio, realizado en una cohorte de 1.095 neonatos con sepsis clínicamente sospechada, se evaluaron distintos modelos de ML para la predicción de mortalidad intrahospitalaria. La red neuronal profunda (DNN) alcanzó el mejor rendimiento predictivo, con un AUC de 0,923 y una exactitud del 95,6 %, destacándose la necesidad de ventilación mecánica y las condiciones de alimentación como variables predictivas relevantes (23).

Algoritmo 2: Inteligencia artificial para la predicción del pronóstico del asma en población pediátrica.

Fuente: Elaboración propia a partir de (23).

Aplicaciones de la inteligencia artificial en la predicción clínica en Ginecología y Obstetricia Preeclampsia

Los estudios incluidos evaluaron modelos de IA para la predicción y detección temprana de la preeclampsia utilizando variables clínicas, bioquímicas y fisiológicas (24-27). En una cohorte multicéntrica

de 17.520 gestantes, modelos de regresión logística y redes neuronales basados en características maternas y biomarcadores de ADN libre circulante (cfDNA) mostraron AUC cercanos a 0,80 para la predicción de preeclampsia pretérmino, con un beneficio limitado del cfDNA, excepto en la preeclampsia de inicio temprano, donde aumentó la sensibilidad en aproximadamente un 7% (24).

Otros estudios reportaron un desempeño superior

con modelos basados en variables clínicas de rutina, como LightGBM, que alcanzó un AUC de 0,83 y una especificidad superior al 90%, así como árboles de decisión aplicados a impedanciocardiografía no invasiva, con AUC de hasta 0,98 para preeclampsia tardía, destacando predictores cardiovasculares relevantes, extrapolados en la práctica clínica Algoritmo 3 (27).

Aborto

Siete estudios evaluaron modelos de aprendizaje automático para la predicción de aborto espontáneo, aborto recurrente y otros desenlaces adversos del embarazo, empleando datos clínicos, biológicos, embrionarios y sociodemográficos (28–34). En el aborto espontáneo temprano, los modelos que integraron imágenes embrionarias y variables clínicas lograron sensibilidades superiores al 70%, asociando una mejor calidad embrionaria con menor riesgo de pérdida gestacional (28).

En el aborto recurrente, los modelos que incorporaron biomarcadores endocrinos e inmunológicos alcanzaron desempeños elevados, con valores de exactitud, sensibilidad y especificidad cercanos o superiores al 93% (29).

En poblaciones de alto riesgo y en reproducción asistida, algoritmos como CHAID y modelos morfocinéticos como PREFER demostraron una adecuada capacidad predictiva antes de la concepción o transferencia embrionaria (32–34). Estudios poblacionales en África Oriental mostraron que los bosques aleatorios alcanzaron altos valores de exactitud y AUC para predecir la terminación del embarazo, resaltando el papel de los determinantes sociales (30,31).

Cáncer cervicouterino

Cinco estudios analizaron la aplicación de IA y ML en el cáncer cervicouterino para diagnóstico y predicción pronóstica, utilizando datos de imagen y variables clínicas (35–39). En el análisis de imágenes, modelos basados en redes neuronales convolucionales identificaron lesiones cervicales de alto grado con exactitudes cercanas al 90%, mientras que enfoques profundos híbridos aplicados a citología cervical alcanzaron exactitudes y AUC superiores al 99% en diferentes conjuntos de datos (35,38).

En estudios pronósticos, los modelos de ML superaron de forma consistente a la regresión logística y a los modelos de Cox en la predicción de supervivencia y recurrencia, con mayores valores de AUC y C-index. En cohortes multicéntricas y bases poblacionales como SEER, estos modelos lograron AUC entre 0,78 y 0,95 y C-index de hasta 0,85, confirmando su potencial para la estratificación del riesgo y la predicción de recurrencia

sitio-específica (36,37,39).

Resultados

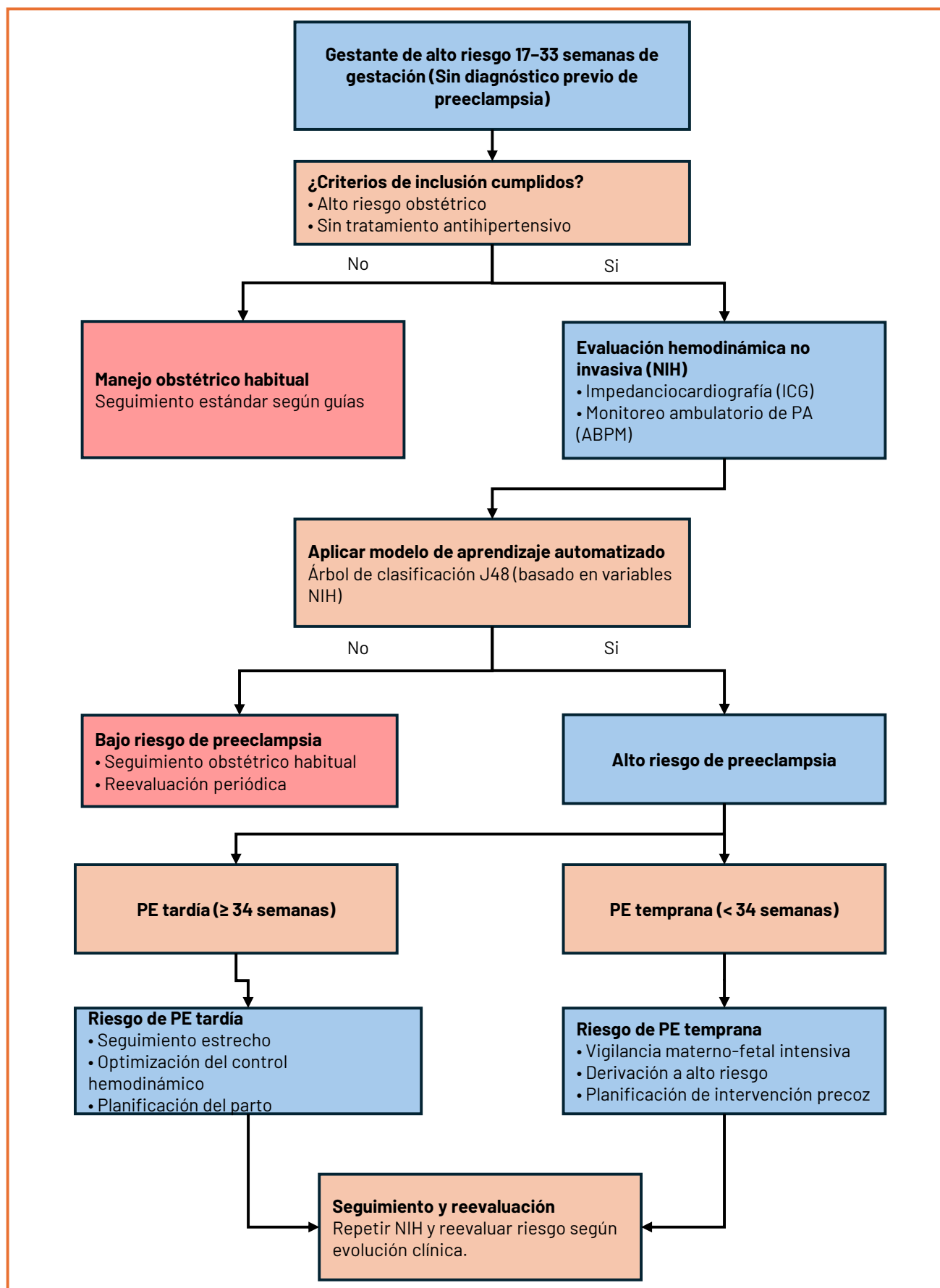
Se incluyeron 49 estudios que abordaron una amplia variedad de patologías clínicas en las tres áreas de interés. Predominaron los modelos de ML tradicional, particularmente aquellos basados en árboles de decisión y métodos de *boosting*. La mayoría de los estudios reportó validación interna, mientras que la validación externa fue menos frecuente.

Discusión

La presente revisión sistemática integra la evidencia reciente sobre la aplicación de modelos de IA y ML en múltiples patologías clínicas, incluyendo enfermedades crónicas no transmisibles, condiciones agudas, patologías materno-infantiles y escenarios críticos. En conjunto, los estudios analizados muestran que los enfoques basados en IA ofrecen un desempeño predictivo consistentemente superior o comparable al de los modelos estadísticos tradicionales, particularmente cuando se aplican a desenlaces clínicamente relevantes y se apoyan en grandes volúmenes de datos clínicos reales (1–49).

Un hallazgo transversal es que los modelos de ML muestran mayor utilidad en contextos donde el desenlace está bien definido y existe disponibilidad de variables clínicas longitudinales. En enfermedades crónicas como diabetes, hipertensión, insuficiencia cardíaca, EPOC y asma, los algoritmos lograron valores de AUC generalmente entre 0,75 y 0,90, destacándose los modelos de aprendizaje en conjunto basados en árboles (bosques aleatorios, *boosting* por gradiente, XGBoost) y, en escenarios longitudinales complejos, las redes neuronales profundas tipo LSTM (5–20,40–43). Estos resultados sugieren que el principal valor de la IA no radica en identificar factores de riesgo completamente nuevos, sino en integrar múltiples variables simultáneamente, capturar relaciones no lineales y generar predicciones individualizadas de riesgo clínico.

En patologías metabólicas como diabetes y obesidad infantil, los estudios evidencian que la inclusión de antecedentes clínicos tempranos, patrones longitudinales y variables familiares o maternas mejora de forma significativa el desempeño predictivo (5–9,46,47,49). En diabetes, los modelos mostraron un rendimiento particularmente alto para la predicción de complicaciones microvasculares, amputación y hospitalización por cetoacidosis diabética, con AUC que en algunos casos superaron 0,85 (5–7). De manera similar, en obesidad infantil, los modelos basados en cohortes longitudinales y registros electrónicos de salud permitieron identificar riesgo desde edades tempranas, reforzando el potencial de estas herramientas para estrategias preventivas

Algoritmo 3: Predicción con IA de preeclampsia temprana y tardía, mediante medición hemodinámica no invasiva

Fuente: Elaboración propia a partir de (30).

personalizadas.

En el ámbito de las enfermedades cardiovasculares y renales críticas, como la insuficiencia cardíaca aguda y la sepsis asociada a lesión renal aguda, los modelos de ML superaron de forma consistente a escalas clínicas tradicionales en la predicción de mortalidad y readmisión temprana (1,2,4,40,41). Un aspecto relevante es que los predictores más importantes identificados no se limitaron a parámetros específicos de órgano, sino que reflejaron la gravedad sistémica del paciente, incluyendo variables hemodinámicas, respiratorias, neurológicas y metabólicas. Esto respalda la hipótesis de que la IA es especialmente útil para modelar síndromes complejos y multiorgánicos donde la interacción entre variables es difícil de capturar mediante modelos lineales.

En patologías respiratorias como EPOC y asma, los modelos demostraron utilidad para la predicción de exacerbaciones, hospitalización y mortalidad, con desempeños moderados, pero clínicamente relevantes (14–20,42,43). La incorporación de datos heterogéneos —como espirometría, tomografía computarizada, señales de voz, medicación y registros administrativos— resalta la capacidad del ML para integrar fuentes diversas de información. No obstante, las mejoras en desempeño frente a modelos convencionales fueron en general modestas, lo que sugiere que el valor añadido en estos contextos podría residir más en la automatización, la escalabilidad y la estratificación del riesgo que en un incremento sustancial de la precisión predictiva.

En el campo materno-infantil, la IA mostró resultados especialmente prometedores. En preeclampsia, los modelos lograron altos valores de AUC, particularmente para la predicción de preeclampsia de inicio temprano, utilizando datos clínicos, biomarcadores, electrocardiogramas y técnicas no invasivas de hemodinámica (24–27). En aborto espontáneo y terminación del embarazo, los modelos permitieron integrar factores clínicos, endocrinos, inmunológicos y sociales, mostrando un alto potencial tanto para la medicina personalizada como para el análisis poblacional y la formulación de políticas públicas (28–34). En sepsis neonatal, los modelos de ML y aprendizaje profundo superaron a los enfoques tradicionales para la predicción de mortalidad, destacando su capacidad para integrar múltiples variables clínicas y de laboratorio en escenarios de alta complejidad (46,47).

En patologías oncológicas como el cáncer cervicouterino, los modelos de IA demostraron un rendimiento elevado tanto en el diagnóstico como en la predicción pronóstica y de recurrencia, superando a los modelos de regresión tradicionales, especialmente cuando se incorporaron múltiples variables clínico-patológicas o datos de imagen (35–39). La posibilidad

de predecir recurrencia sitio-específica representa un avance relevante hacia la medicina de precisión, aunque su implementación clínica aún requiere validación prospectiva.

Un tema recurrente en todos los ejes clínicos es la tensión entre desempeño predictivo e interpretabilidad. Mientras que los modelos más complejos tienden a ofrecer mejores métricas, su adopción clínica puede verse limitada si no se acompañan de estrategias de explicación claras. En este sentido, el uso de técnicas de IA explicable, como SHAP, reglas basadas en árboles o modelos híbridos, representa un avance clave para facilitar la aceptación clínica y la toma de decisiones informada (1,2,4,7,8,17–19,37). La interpretabilidad emerge como un requisito esencial para la traslación de estos modelos desde la investigación hacia la práctica asistencial.

A pesar de los resultados prometedores, esta revisión identifica limitaciones comunes en la literatura actual. La mayoría de los estudios presentan diseños retrospectivos, validación interna y dependencia de un solo sistema de salud o base de datos, lo que limita la generalización de los hallazgos. Asimismo, la calidad y completitud de los registros clínicos electrónicos sigue siendo un factor crítico, con riesgos de sesgo por datos faltantes, codificación heterogénea y cambios en las prácticas clínicas a lo largo del tiempo (1–49). Finalmente, pocos estudios evalúan de manera prospectiva el impacto clínico real de la implementación de estos modelos en la toma de decisiones y en los desenlaces en salud.

La evidencia disponible sugiere que los modelos de IA y ML tienen un alto potencial para apoyar la predicción de desenlaces clínicos relevantes en múltiples áreas de la medicina. Su mayor fortaleza radica en la integración de grandes volúmenes de datos clínicos y longitudinales, permitiendo una estratificación del riesgo más precisa. No obstante, la generalización de estos modelos se ve limitada por diseños retrospectivos, validación interna predominante y heterogeneidad metodológica. Futuras investigaciones deberán centrarse en validaciones externas, estudios prospectivos y evaluación del impacto clínico real para facilitar su implementación en la práctica asistencial.

Conclusiones

La evidencia sintetizada en esta revisión sistemática indica que los modelos predictivos basados en IA tienen un alto potencial para apoyar la predicción de desenlaces clínicos relevantes en Medicina Interna, Ginecología/Obstetricia y Pediatría. En los distintos contextos analizados, estos modelos mostraron un desempeño generalmente favorable, en particular aquellos basados en ML tradicional y enfoques de aprendizaje en conjunto, lo que respalda su utilidad para la estratificación del riesgo y la identificación

temprana de pacientes con mayor probabilidad de complicaciones clínicas.

No obstante, la aplicabilidad clínica de estas herramientas continúa limitada por la predominancia de diseños retrospectivos, la validación principalmente interna y la heterogeneidad metodológica entre los estudios, factores que restringen la generalización de los resultados y su implementación rutinaria en distintos entornos asistenciales.

En conjunto, los hallazgos sugieren que la IA constituye una herramienta prometedora para el análisis predictivo en salud; sin embargo, su integración clínica efectiva requerirá estudios prospectivos, validaciones externas independientes y evaluaciones del impacto clínico real, con el fin de garantizar un uso seguro, eficaz y éticamente responsable en los sistemas de salud.

Reflexiones del autor

La historia ha demostrado que cada generación está marcada por procesos de cambio que transforman la manera en que las sociedades comprenden y enfrentan su realidad. Frecuentemente, las generaciones anteriores perciben a las nuevas como carentes de ciertas habilidades o valores en comparación con su propio contexto; sin embargo, esta percepción responde a las múltiples variables que acompañan la evolución del tiempo y a las distintas realidades sociales en las que se desarrolla cada grupo humano.

En la actualidad, la inteligencia artificial representa uno de los cambios más significativos en la forma en que vivimos, pensamos y actuamos, integrándose progresivamente en la vida cotidiana. Este fenómeno no es ajeno a la medicina, una disciplina históricamente caracterizada por su componente humano y su interacción directa con el paciente. Por el contrario, los cambios tecnológicos transforman también la manera en que los profesionales de la salud abordan los problemas clínicos y toman decisiones.

Más allá del reconocimiento individual o del desarrollo tecnológico en sí mismo, la medicina mantiene un propósito fundamental: brindar el mejor resultado posible para el paciente. En este sentido, como se evidenció en la presente revisión sistemática, múltiples modelos de inteligencia artificial han demostrado un desempeño predictivo favorable en distintos desenlaces clínicos, lo que respalda su potencial como herramientas de apoyo en la práctica médica.

No obstante, la evidencia actual aún es limitada para avalar su implementación rutinaria en medicina interna, ginecología y pediatría (las especialidades estudiadas en esta revisión). Como toda herramienta en desarrollo, la inteligencia artificial requiere un proceso continuo de perfeccionamiento técnico, validación y evaluación

de su impacto real. En el ámbito médico, donde las decisiones tienen consecuencias directas sobre la vida y la salud de las personas, el margen de error debe ser mínimo, ya que sus efectos pueden ser graves e incluso irreversibles.

En consecuencia, la integración de la inteligencia artificial en la medicina no debe entenderse como un reemplazo del juicio clínico, sino como una oportunidad para fortalecerlo. Su desarrollo y aplicación deben estar guiados por principios éticos, rigor científico y un enfoque centrado en el paciente, asegurando que la innovación tecnológica se traduzca en una mejora real en la calidad de la atención en salud.

Consideraciones éticas, sesgos y conflictos de interés

Consideraciones éticas

La presente revisión sistemática se basó exclusivamente en el análisis de estudios previamente publicados, por lo que no implicó la intervención directa sobre seres humanos ni el uso de datos personales identificables. En consecuencia, no fue necesaria la aprobación por un comité de ética en investigación. No obstante, el desarrollo del estudio se llevó a cabo siguiendo principios éticos de rigor científico, transparencia metodológica y reporte responsable de los resultados.

Sesgos de publicación

Un sesgo potencial relevante en esta revisión corresponde al sesgo de publicación, dado que los estudios con resultados negativos o con bajo desempeño de los modelos predictivos basados en inteligencia artificial pueden no ser publicados con la misma frecuencia que aquellos con resultados positivos. Esta situación podría conducir a una sobreestimación del rendimiento real de los modelos reportados en la literatura disponible.

Sesgo de selección

El sesgo de selección constituye otra limitación importante, ya que muchos de los estudios incluidos emplearon diseños observacionales retrospectivos y poblaciones provenientes de bases de datos específicas o de un único sistema de salud. Estas diferencias sistemáticas entre las poblaciones estudiadas pueden afectar la comparabilidad entre los estudios y limitar la generalización de los hallazgos a otros contextos clínicos. Asimismo, la ausencia de validación externa independiente en varios trabajos refuerza este riesgo.

Sesgo del observador

Aunque el sesgo del observador es menos frecuente en revisiones sistemáticas, no puede excluirse

completamente su presencia durante las etapas de selección y evaluación de los estudios incluidos. En este trabajo, se procuró reducir este riesgo mediante la definición previa de criterios de inclusión y exclusión y la aplicación de un proceso de cribado sistemático. Sin embargo, es posible que factores como la familiaridad con determinados autores, instituciones o líneas de investigación hayan ejercido una influencia no intencionada en la valoración de algunos estudios.

Conflictos de interés

Los autores declaran no tener conflictos de interés financieros, académicos ni personales que hayan influido en la realización, el análisis o la interpretación de los resultados de esta revisión sistemática.

Referencias

- Gao T, Nong Z, Luo Y, Mo M, Chen Z, Yang Z, et al. Machine learning-based prediction of in-hospital mortality for critically ill patients with sepsis-associated acute kidney injury. *Ren Fail.* 2024;46(1):2316267. doi:10.1080/0886022X.2024.2316267 PubMed PMID: 38369749.
- Zhuang C, Hu R, Li K, Liu Z, Bai S, Zhang S, et al. Machine learning prediction models for mortality risk in sepsis-associated acute kidney injury: evaluating early versus late CRRT initiation. *Front Med (Lausanne).* 2024;11:1483710. doi:10.3389/FMED.2024.1483710/FULL
- Bottrighi A, Pennisi M, Roveta A, Massarino C, Cassinari A, Betti M, et al. A machine learning approach for predicting high risk hospitalized patients with COVID-19 SARS-Cov-2. *BMC Med Inform Decis Mak.* 2022 Dec 1;22(1). doi:10.1186/S12911-022-02076-1 PubMed PMID: 36578017.
- Zhang Y, Xiang T, Wang Y, Shu T, Yin C, Li H, et al. Explainable machine learning for predicting 30-day readmission in acute heart failure patients. *iScience.* 2024 Jul 19;27(7):110281. doi:10.1016/J.ISCI.2024.110281
- Lai YR, Chiu WC, Huang CC, Cheng BC, Kung C Te, Lin TY, et al. Longitudinal artificial intelligence-based deep learning models for diagnosis and prediction of the future occurrence of polyneuropathy in diabetes and prediabetes. *Neurophysiologie Clinique.* 2024 Jul 1;54(4):102982. doi:10.1016/J.NEUCLI.2024.102982 PubMed PMID: 38761793.
- Tan JK, Quan L, Salim NNM, Tan JH, Goh SY, Thumboo J, et al. Machine Learning-Based Prediction for High Health Care Utilizers by Using a Multi-Institutional Diabetes Registry: Model Training and Evaluation. *JMIR AI.* 2024;3:e58463. doi:10.2196/58463 PubMed PMID: 39418089.
- Wen Y, Wan Z, Ren H, Wang X, Wang W. Interpretable Machine Learning Model for Predicting and Assessing the Risk of Diabetic Nephropathy: Prediction Model Study. *JMIR Med Inform.* 2025 Oct 22;13:e64979. doi:10.2196/64979 PubMed PMID: 41124652.
- Gao L, Liu ZX, Wang JN. Predictive model and risk analysis for outcomes in diabetic foot ulcer using eXtreme Gradient Boosting algorithm and SHapley Additive exPlanation. *World J Diabetes.* 2025 Jul 15;16(7):104789. doi:10.4239/WJD.V16.I7.104789 PubMed PMID: 40697589.
- Wändell P, Carlsson AC, Wierzbicka M, Sigurdsson K, Årnlöv J, Eriksson J, et al. A machine learning tool for identifying patients with newly diagnosed diabetes in primary care. *Prim Care Diabetes.* 2024 Oct 1;18(5):501-5. doi:10.1016/j.pcd.2024.06.010 PubMed PMID: 38944562.
- Norrman A, Hasselström J, Ljunggren G, Wachtler C, Eriksson J, Kahan T, et al. Predicting new cases of hypertension in Swedish primary care with a machine learning tool. *Prev Med Rep.* 2024 Aug 1;44:102806. doi:10.1016/J.PMEDR.2024.102806 PubMed PMID: 39091569.
- Li W, Huang G, Tang N, Lu P, Jiang L, Lv J, et al. Effects of heavy metal exposure on hypertension: A machine learning modeling approach. *Chemosphere.* 2023 Oct 1;337:139435. doi:10.1016/J.CHEMOSPHERE.2023.139435 PubMed PMID: 37422210.
- Hung MH, Shih LC, Wang YC, Leu HB, Huang PH, Wu TC, et al. Prediction of Masked Hypertension and Masked Uncontrolled Hypertension Using Machine Learning. *Front Cardiovasc Med.* 2021;8:778306. doi:10.3389/FCVM.2021.778306/FULL
- Wu L, Huang L, Li M, Xiong Z, Liu D, Liu Y, et al. Differential diagnosis of secondary hypertension based on deep learning. *Artif Intell Med.* 2023 Jul 1;141:102554. doi:10.1016/J.ARTMED.2023.102554 PubMed PMID: 37295898.
- Jeon ET, Park H, Lee JK, Heo EY, Lee CH, Kim DK, et al. Deep Learning-Based Chronic Obstructive Pulmonary Disease Exacerbation Prediction Using Flow-Volume and Volume-Time Curve Imaging: Retrospective Cohort Study. *J Med Internet Res.* 2025;27(1):e69785. doi:10.2196/69785 PubMed PMID: 40373296.
- Panettieri RA, Roy J, Gontarczyk Uczkowski N, Tyler A, Attanucci J, O'Riordan TG, et al. Leveraging Machine Learning and Real-World Data to Predict Chronic Obstructive Pulmonary Disease Exacerbations. *Int J Chron Obstruct Pulmon Dis.* 2025;20:3451. doi:10.2147/COPD.S536395 PubMed PMID: 41195056.
- Zhao X, Qiu X. Computed tomography radiomics-based prognosis prediction of patients with chronic obstructive pulmonary disease. *J Radiat Res Appl Sci.* 2024 Dec 1;17(4):101168. doi:10.1016/J.JRRAS.2024.101168
- Idrisoglu A, Dallora AL, Cheddad A, Anderberg P, Jakobsson A, Sanmartin Berglund J. COPDVID: Automated classification of chronic obstructive pulmonary disease on a new collected and evaluated voice dataset. *Artif Intell Med.* 2024 Oct 1;156:102953. doi:10.1016/J.ARTMED.2024.102953 PubMed PMID: 39222579.
- Alhur AA, Sani J, Ahmed MM. Predicting COPD admissions using machine learning and SHAP: An exploratory multi-hospital study in Riyadh, Saudi Arabia. *Intell Based Med.* 2025 Jan 1;12:100312. doi:10.1016/J.IBMED.2025.100312
- Pena-Gralle APB, Forget A, Chiu YM, Legault MA, Beauchesne MF, Blais L. Medication-based mortality prediction in COPD using machine learning and conventional statistical methods. *Int J Med Inform.* 2025 Feb 1;206:106177. doi:10.1016/J.IJMEDINF.2025.106177 PubMed PMID: 41202398.
- Sagheb E, Wi C II, King KS, Agnikula Kshatriya BS, Ryu E, Liu H, et al. AI model for predicting asthma prognosis in children. *The Journal of Allergy and Clinical Immunology: Global.* 2025 May 1;4(2):100429. doi:10.1016/J.JACIG.2025.100429 PubMed PMID: 40091884.
- Ghareeb WM, Draz E, Chen X, Zhang J, Tu P, Madbouly K, et al. Multicenter validation of an artificial intelligence (AI)-based platform for the diagnosis of acute appendicitis. *Surgery.* 2024 Sep 1;176(3):569-76. doi:10.1016/J.SURG.2024.05.007 PubMed PMID: 38910047.
- Iqbal F, Chandra P, Lewis LES, Acharya D, Purkayastha J, Shenoy PA, et al. Application of artificial intelligence to predict the sepsis in neonates admitted in neonatal

- intensive care unit. *Journal of Neonatal Nursing*. 2024 Apr 1;30(2):141-7. doi:10.1016/J.JNN.2023.07.016
23. Hsu JF, Chang YF, Cheng HJ, Yang C, Lin CY, Chu SM, et al. Machine Learning Approaches to Predict In-Hospital Mortality among Neonates with Clinically Suspected Sepsis in the Neonatal Intensive Care Unit. *J Pers Med*. 2021 Aug 1;11(8):695. doi:10.3390/JPM11080695 PubMed PMID: 34442338.
 24. Khalil A, Bellesia G, Norton ME, Jacobsson B, Haeri S, Egbert M, et al. The role of cell-free DNA biomarkers and patient data in the early prediction of preeclampsia: an artificial intelligence model. *Am J Obstet Gynecol*. 2024 Nov 1;231(5):554.e1-554.e18. doi:10.1016/j.ajog.2024.02.299 PubMed PMID: 38432413.
 25. Bülüz A, Hansu K, Çağan ES, Şahin AR, Dokumaci H. Artificial Intelligence in Early Diagnosis of Preeclampsia. *Niger J Clin Pract*. 2024 Mar 1;27(3):383-8. doi:10.4103/NJCP.NJCP_222_23 PubMed PMID: 38528360.
 26. Butler L, Gunturkun F, Chinthala L, Karabayir I, Tootooni MS, Bakir-Batu B, et al. AI-based preeclampsia detection and prediction with electrocardiogram data. *Front Cardiovasc Med*. 2024;11:1360238. doi:10.3389/FCVM.2024.1360238/FULL
 27. Olano D, Espeche W, Minetto J, Leiva Sisniegues BC, Cerri G, Martinez C, et al. Prediction for the development of preeclampsia through non-invasive hemodynamics using machine learning, distinguishing early from late. *Pregnancy Hypertens*. 2025 Sep 1;41:101242. doi:10.1016/J.PREGHY.2025.101242 PubMed PMID: 40737756.
 28. Chavez-Badiola A, Farias AFS, Mendizabal-Ruiz G, Silvestri G, Griffin DK, Valencia-Murillo R, et al. Use of artificial intelligence embryo selection based on static images to predict first-trimester pregnancy loss. *Reprod Biomed Online*. 2024 Aug 1;49(2):103934. doi:10.1016/j.rbmo.2024.103934 PubMed PMID: 38824762.
 29. Shi B, Chen J, Chen H, Lin W, Yang J, Chen Y, et al. Prediction of recurrent spontaneous abortion using evolutionary machine learning with joint self-adaptive sime mould algorithm. *Comput Biol Med*. 2022 Sep 1;148:105885. doi:10.1016/J.COMPBIOMED.2022.105885 PubMed PMID: 35930957.
 30. Setegn GM, Dejene BE. Explainable artificial intelligence models for predicting pregnancy termination among reproductive-aged women in six east African countries: machine learning approach. *BMC Pregnancy Childbirth*. 2024 Dec 1;24(1). doi:10.1186/S12884-024-06773-9 PubMed PMID: 39285277.
 31. Tesfa GA, Demeke AD, Seboka BT, Tebeje TM, Kasaye MD, Gebremeskele BT, et al. Employing machine learning models to predict pregnancy termination among adolescent and young women aged 15-24 years in East Africa. *Sci Rep*. 2024 Dec 1;14(1):30047. doi:10.1038/S41598-024-81197-1 PubMed PMID: 39627430.
 32. Liu Y, Geng Y, Yang L, Xiang S, Wang Q, Hu L, et al. Traditional Chinese Medicine Constitution and Clinical Data Association with Machine Learning for Prediction of Spontaneous Abortion. *Clinical Complementary Medicine and Pharmacology*. 2022 Jun 1;2(2):100016. doi:10.1016/J.CCMP.2021.100016
 33. Paydar K, Sheikhtaheri A. Prediction of pregnancy outcomes in women with systemic lupus erythematosus before pregnancy: Application of machine learning models. *Heliyon*. 2025 Feb 28;11(4):e42679. doi:10.1016/J.HELİYON.2025.E42679 PubMed PMID: 40040993.
 34. Bamford T, Smith R, Easter C, Dhillon-Smith R, Barrie A, Montgomery S, et al. Association between a morphokinetic ploidy prediction model risk score and miscarriage and live birth: a multicentre cohort study. *Fertil Steril*. 2023 Oct 1;120(4):834-43. doi:10.1016/J.FERTNSTERT.2023.06.006 PubMed PMID: 37307891.
 35. Takahashi T, Matsuoka H, Sakurai R, Akatsuka J, Kobayashi Y, Nakamura M, et al. Development of a prognostic prediction support system for cervical intraepithelial neoplasia using artificial intelligence-based diagnosis. *J Gynecol Oncol*. 2022 Sep 1;33(5):e57. doi:10.3802/JGO.2022.33.E57 PubMed PMID: 35712970.
 36. Nikookar E, Naderi E, Rahnavard A. Cervical Cancer Prediction by Merging Features of Different Colposcopic Images and Using Ensemble Classifier. *J Med Signals Sens*. 2021 Apr 1;11(2):67. doi:10.4103/JMSS.JMSS_16_20 PubMed PMID: 34268095.
 37. Roy P, Hasan M, Islam MR, Uddin MP. Interpretable artificial intelligence (AI) for cervical cancer risk analysis leveraging stacking ensemble and expert knowledge. *Digit Health*. 2025 Jan 1;11:20552076251327944. doi:10.1177/20552076251327945 PubMed PMID: 40144051.
 38. Ameta GK, Ramkumar MS, Jenifa G, Sivaramkrishnan M, Thangavelsamy N, Ghouse M, et al. Cervical cancer prediction using deformable kernel darknet-53 and depth wise separable convolutional neural networks. *Sci Rep*. 2025 Dec 1;15(1):37236. doi:10.1038/S41598-025-21119-X PubMed PMID: 41136553.
 39. Guo C, Wang J, Wang Y, Qu X, Shi Z, Meng Y, et al. Novel artificial intelligence machine learning approaches to precisely predict survival and site-specific recurrence in cervical cancer: A multi-institutional study. *Transl Oncol*. 2021 May 1;14(5):101032. doi:10.1016/J.TRANON.2021.101032 PubMed PMID: 33618238.
 40. Palmer DQ, Gismondi RA, Gemal P, Lomba GS, Seixas FL. Mortality Prediction in Heart Failure Patients: Machine Learning Versus Get With The Guidelines-Heart Failure (GWTG-HF) and Acute Decompensated Heart Failure National Registry (ADHERE). *Cureus*. 2025 Aug 1;17(8):e89226. doi:10.7759/CUREUS.89226 PubMed PMID: 40895687.
 41. Li J, Sun Y, Ren J, Wu Y, He Z. Machine Learning for In-hospital Mortality Prediction in Critically Ill Patients With Acute Heart Failure: A Retrospective Analysis Based on the MIMIC-IV Database. *J Cardiothorac Vasc Anesth*. 2025 Mar 1;39(3):666-74. doi:10.1053/J.JVCA.2024.12.016 PubMed PMID: 39779429.
 42. Rezaeiahari M, Brown CC, Eyimina A, Perry TT, Goudie A, Boyd M, et al. Predicting Pediatric Severe Asthma Exacerbations: An Administrative Claims-Based Predictive Model. *J Asthma*. 2024;61(3):203. doi:10.1080/02770903.2023.2260881 PubMed PMID: 37725084.
 43. Sills MR, Ozkaynak M, Jang H. Predicting hospitalization of pediatric asthma patients in emergency departments using machine learning. *Int J Med Inform*. 2021 Jul 1;151:104468. doi:10.1016/J.IJMEDINF.2021.104468 PubMed PMID: 33940479.
 44. Ye LZ, Sun JX, Chen J, Cen KK, Bi Y, Lu YC. Machine learning model for predicting urinary tract infection risk in febrile children under 3 years of age. *Front Pediatr*. 2025;13:1677292. doi:10.3389/FPED.2025.1677292 PubMed PMID: 41438455.
 45. Sánchez-Martínez LJ, Charle-Cuéllar P, Gado AA, Ousmane N, Hernández CL, López-Ejeda N. Using Machine Learning to Fight Child Acute Malnutrition and Predict Weight Gain During Outpatient Treatment with a Simplified Combined Protocol. *Nutrients*. 2024 Dec 1;16(23):4213. doi:10.3390/NU16234213 PubMed PMID: 39683605.
 46. Pang X, Forrest CB, Lê-Scherban F, Masino AJ. Prediction of early childhood obesity with machine learning and

- electronic health record data. *Int J Med Inform.* 2021 Jun 1;150:104454. doi:10.1016/J.IJMEDINF.2021.104454 PubMed PMID: 33866231.
47. Lim H, Lee H, Kim J. A prediction model for childhood obesity risk using the machine learning method: a panel study on Korean children. *Sci Rep.* 2023 Dec 1;13(1):10122. doi:10.1038/S41598-023-37171-4 PubMed PMID: 37344518.
48. Hendy A, Ibrahim RK, Abdelaliem SMF, Zaher A, Alkubati SA, El-kader RGA, et al. Supervised machine learning for classification and prediction of stunting among under-five Egyptian children. *BMC Pediatr.* 2025 Dec 1;25(1):681. doi:10.1186/S12887-025-06138-X PubMed PMID: 40963124.
49. Lee H, Park MB, Won YJ. AI Machine Learning-Based Diabetes Prediction in Older Adults in South Korea: Cross-Sectional Analysis. *JMIR Form Res.* 2025 Jan 1;9. doi:10.2196/57874
50. Richter F, Holmes E, Richter F, Guttman K, Duong SQ, Gangadharan S, et al. Toward governance of artificial intelligence in pediatric healthcare. *npj Digital Medicine* 2025 8:1. 2025 Oct 30;8(1):636-. doi:10.1038/s41746-025-02000-7
51. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-36. doi:10.7326/0003-4819-155-8-201110180-00009 PubMed PMID: 22007046.
52. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. *Ann Intern Med.* 2019 Jan 1;170(1):W1-33. doi:10.7326/M18-1377 PubMed PMID: 30596876.
53. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol.* 2017 Dec 1;2(4):230-43. doi:10.1136/SVN-2017-000101 PubMed PMID: 29507784.
54. Charles Darwin (1859). London: John Murray [Internet]. 1859 [cited 2025 Dec 15]. Charles Darwin (1859): El origen de las especies. Available from: <https://www.marxists.org/espanol/darwin/1859/origenespecies/index.htm>
55. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine* 2019 25:1. 2019 Jan 7;25(1):44-56. doi:10.1038/s41591-018-0300-7 PubMed PMID: 30617339.