Explainable prediction of preeclampsia using routinely collected clinical data for primary care in Honduras: Design and operational protocol of PRECLA-XAI-HN

Salvador Diaz [1]; Jorge Urmeneta[2]; Nelson Sánchez[3]; Selvin Reyes[4]; Gustavo Galo[5]; Leda Sánchez[6] and Marisol Maradiaga[7]

[2,3,4,5,6,7] Universidad Nacional Autónoma de Honduras, Honduras, *jorge.urmeneta@unah.edu.hn, nelson.sanchez@unah.edu.hn, selvin.reyes@unah.edu.hn, Gustavo.galo@unah.edu.hn, leda.sanche@unah.edu.hn, marisol.maradiaga@unah.edu.hn*

[1]*Secretaría de Salud, Honduras, sedicacdo@hotmail.com*

Abstract– ***Hypertensive disorders of pregnancy remain a major cause of maternal and perinatal morbidity, especially where first level facilities have limited access to angiogenic biomarkers, specialized ultrasound and continuous obstetric surveillance. This paper proposes PRECLA-XAI-HN, an explainable algorithm for preeclampsia risk stratification using simple variables routinely available in primary care. A design science approach was used to define the clinical variables, the modelling pipeline, the explanation layer and an operational protocol for triage and referral. Candidate models included logistic regression, pruned CART, random forest and XGBoost. A desk validation with 3,600 synthetic pregnancy records was used to test internal coherence, discrimination, calibration and decision utility. The operational model combined calibrated XGBoost predictions with SHAP based explanations and safety gates that override the algorithm when severe hypertension or alarm symptoms are present. In the simulated test set, the XGBoost model reached an AUC ROC of 0.786, AUC PR of 0.337, sensitivity of 0.871 at an 8% action threshold and negative predictive value of 0.981. The leading explanatory variables were mean arterial pressure, diastolic pressure, body mass index, maternal age and systolic pressure. The contribution of the work is an implementable algorithm and protocol for primary care, conceived as decision support and not as an autonomous diagnostic system.***

Keywords-- ***preeclampsia, explainable artificial intelligence, primary care, clinical decision support, Honduras.***

Predicción explicable de la preeclampsia utilizando datos clínicos recogidos rutinariamente para atención primaria en Honduras: Diseño y protocolo de PRECLA-XAI-HN

Salvador Diaz [1]; Jorge Urmeneta[2]; Nelson Sánchez[3]; Selvin Reyes[4]; Gustavo Galo[5]; Leda Sánchez[6] and Marisol Maradiaga[7]

[2,3,4,5,6,7] Universidad Nacional Autónoma de Honduras, Honduras, *jorge.urmeneta@unah.edu.hn, nelson.sanchez@unah.edu.hn, selvin.reyes@unah.edu.hn, Gustavo.galo@unah.edu.hn, leda.sanche@unah.edu.hn, marisol.maradiaga@unah.edu.hn*

[1]*Secretaría de Salud, Honduras, sedicacdo@hotmail.com*

Resumen– ***Las complicaciones hipertensivas del embarazo continúan produciendo una carga clínica importante en contextos donde el primer nivel de atención dispone de presión arterial, antecedentes maternos y pruebas básicas, pero no siempre de biomarcadores angiogénicos, ultrasonido especializado o vigilancia obstétrica continua. Este artículo propone PRECLA-XAI-HN, un algoritmo explicable para estratificar riesgo de preeclampsia con datos clínicos simples y un protocolo operativo para su uso en establecimientos de primer nivel. Se desarrolló un estudio de ciencia del diseño orientado a producto, basado en revisión de evidencia, definición de variables mínimas, modelado predictivo, explicabilidad local y validación de escritorio con datos sintéticos. Se compararon regresión logística, CART podado, Random Forest y XGBoost calibrable. El conjunto sintético incluyó 3,600 gestantes simuladas y una prevalencia de preeclampsia de 6.5%. En el conjunto de prueba, XGBoost alcanzó AUC ROC de 0.786, AUC PR de 0.337, sensibilidad de 0.871 al umbral operativo de 8% y valor predictivo negativo de 0.981. Las variables explicativas predominantes fueron presión arterial media, presión diastólica, índice de masa corporal, edad materna y presión sistólica. El protocolo clasifica a las gestantes en cinco estratos de acción, con derivación inmediata cuando existe presión arterial severa o síntomas de alarma. El resultado es un algoritmo y protocolo reproducibles para priorización clínica, no un sustituto del juicio profesional ni de la confirmación diagnóstica.***

Palabras clave-- preeclampsia, inteligencia artificial explicable, atención primaria, apoyo a la decisión clínica, Honduras.

I. Introducción

La preeclampsia es un trastorno hipertensivo del embarazo que suele aparecer después de las 20 semanas de gestación y se define clínicamente por hipertensión de novo, con o sin proteinuria, acompañada de criterios de daño orgánico cuando la presentación es grave [1], [2]. La Organización Mundial de la Salud estima que afecta entre 3% y 8% de las mujeres que dan a luz y que los trastornos hipertensivos contribuyeron a cerca de 16% de las muertes maternas mundiales en 2023 [3]. Su progresión hacia eclampsia, desprendimiento placentario, restricción del crecimiento fetal, parto prematuro o muerte materna mantiene a esta condición como una prioridad de salud pública.

En Honduras, la mortalidad materna continúa siendo un indicador sensible de inequidad sanitaria. La OPS reporta una razón de mortalidad materna estimada de 71.8 por 100,000 nacidos vivos para 2020 y una cobertura de atención del parto por personal calificado de 52% en 2021 [4]. Un análisis reciente de la Revista Médica Hondureña documentó variación importante de la razón de mortalidad materna nacional, con 125 por 100,000 nacidos vivos en 2021 y 65 por 100,000 en 2023, y señaló que la hemorragia y la preeclampsia permanecen entre las principales causas obstétricas directas [5]. Esta carga es especialmente relevante para zonas rurales o periféricas, donde la detección tardía y la referencia demorada aumentan el riesgo de desenlaces adversos.

La aceptabilidad de herramientas de apoyo a la decisión clínica en contextos maternos ha sido evaluada recientemente en una revisión sistemática cualitativa que incluyó 16 estudios y 1,639 mujeres de países de ingresos bajos, medios y altos. Los hallazgos indican que estas herramientas son aceptables tanto para profesionales de salud como para gestantes, y que pueden mejorar la toma de decisiones clínicas cuando se integran adecuadamente al flujo de trabajo [6].

Los modelos predictivos de preeclampsia suelen depender de combinaciones de historia materna, presión arterial media, índices Doppler uterinos y biomarcadores como PAPP-A o PlGF [7]. Aunque estos enfoques son robustos en centros con alta capacidad diagnóstica, su transferibilidad al primer nivel de atención es limitada. En establecimientos rurales, el dato disponible con mayor regularidad es más simple: edad materna, paridad, presión arterial, antecedentes clínicos, presencia de síntomas y, cuando existe, proteinuria por tira reactiva. Por ello, la innovación no consiste en maximizar complejidad técnica, sino en convertir variables mínimas en una alerta comprensible, auditable y útil para la toma de decisiones.

La evidencia reciente confirma el crecimiento acelerado de modelos de aprendizaje automático para predicción de preeclampsia. Una revisión sistemática de 35 estudios publicados entre 2013 y 2023 identificó que las características predictivas más frecuentes en estos modelos son edad, número

de embarazos, índice de masa corporal, diabetes, hipertensión y presión arterial, mientras que datos genéticos e imagenología clínica fueron considerados con menor frecuencia [8]. Un meta-análisis publicada en 2026 evaluó modelos de ML para preeclampsia mediante PROBAST y calculó intervalos de predicción al 95%, concluyendo que la validación externa continúa siendo indispensable para determinar el rendimiento real en entornos clínicos futuros [9]. La evidencia reciente sintetizada en una revisión sistemática y metaanálisis de modelos de aprendizaje automático para predicción de preeclampsia respalda la necesidad de evaluar desempeño, calibración y aplicabilidad clínica antes de su implementación en atención prenatal. Es por ello que PRECLA-XAI-HN como un algoritmo diseñado para variables rutinarias, reconociendo que su validación empírica es la siguiente etapa obligatoria.

El objetivo de este artículo es proponer un algoritmo explicable de predicción de preeclampsia con datos clínicos simples, denominado PRECLA-XAI-HN, junto con un protocolo de uso para primer nivel de atención. El trabajo se orienta a una contribución tecnológica y operativa: un modelo de riesgo, una capa de explicación local y una ruta de acción clínica que pueda ser evaluada posteriormente con datos retrospectivos y prospectivos del sistema sanitario.

.

II. Materiales y métodos

A. *Diseño del estudio*

Se desarrolló un estudio de ciencia del diseño orientado a producto, con tres componentes: síntesis documental de evidencia clínica y metodológica, diseño de un pipeline predictivo explicable y validación de escritorio con datos sintéticos. El producto esperado fue un algoritmo con protocolo operativo, no una prueba diagnóstica autónoma. No se utilizaron datos personales ni historias clínicas reales. La simulación tuvo como finalidad verificar consistencia, flujo de cálculo, comportamiento de métricas y calidad de visualización.

B. *Variables clínicas simples*

La selección de variables priorizó disponibilidad en el primer nivel de atención, alineación con la Historia Clínica Perinatal y pertinencia clínica. El Sistema Informático Perinatal del CLAP/OPS organiza información desde la primera consulta prenatal hasta el alta, facilita la estandarización del registro y permite consolidación para análisis local y nacional [10]. Bajo esa lógica, PRECLA-XAI-HN utiliza variables que pueden registrarse sin equipamiento especializado: edad, índice de masa corporal, edad gestacional al tamizaje, paridad, antecedentes de preeclampsia, hipertensión crónica, diabetes, enfermedad renal, gestación múltiple, presión sistólica, presión diastólica, presión arterial media, proteinuria por tira reactiva, síntomas de alarma y variables operativas de acceso. Estas se describen en la Tabla I.

TABLA I
VARIABLES SIMPLES CANDIDATAS PARA PRECLA-XAI-HN

Dominio	Variable simple	Tipo	Justificación clínica y operativa
Identificación materna	Edad materna	Numérica	Riesgo incremental en extremos de edad y disponibilidad universal
Antropometría	IMC pregestacional o estimado	Numérica	Obesidad incrementa riesgo y orienta consejería preventiva
Historia obstétrica	Nuliparidad, preeclampsia previa, gestación múltiple	Binaria	Factores mayores incluidos en guías clínicas internacionales
Comorbilidad	Hipertensión crónica, diabetes, enfermedad renal	Binaria	Condiciones de alto riesgo que modifican vigilancia y referencia
Consulta actual	Presión sistólica, diastólica y presión arterial media	Numérica	Medición disponible en primer nivel y señal temprana del continuo hipertensivo
Consulta actual	Proteinuria por tira reactiva	Ordinal	Prueba de bajo costo cuando no hay proteinuria cuantitativa
Síntomas	Cefalea intensa, fosfenos, epigastralgia o alarma reportada	Binaria	Gatillo de seguridad para referencia inmediata
Acceso	Zona rural y tiempo de traslado mayor a 60 minutos	Binaria	Permite ajustar prioridad operativa sin sustituir criterio clínico

C. *Modelado predictivo y explicabilidad*

El riesgo individual se estimó mediante una función probabilística general en la Ecuación (1):

$$p_i = \frac{1}{\{1 + exp[-f\theta(x_i)]\}} \quad (1)$$

En Ec. (1), p_i representa la probabilidad estimada de preeclampsia para la gestante i, x_i es el vector de variables clínicas simples y $f\theta$ es la función aprendida por el modelo. Se compararon cuatro familias: regresión logística como referencia interpretable, árbol CART podado como regla clínica visual, Random Forest y XGBoost. Este último se incluyó por su buen rendimiento en datos tabulares, su tolerancia a relaciones no lineales y su disponibilidad de explicaciones TreeSHAP [11], [12].

Para XGBoost se definieron los hiperparámetros mediante búsqueda en grilla con validación cruzada estratificada de cinco pliegues. Los parámetros explorados incluyeron profundidad máxima del árbol (3, 5, 7), tasa de aprendizaje (0.01, 0.05, 0.1), número de estimadores (100, 300, 500) y peso de la clase positiva (scale_pos_weight calculado como

ratio de la clase mayoritaria sobre la minoritaria). La calibración posterior se realizó mediante Platt scaling con regresión logística sobre las probabilidades predichas en el conjunto de validación. La implementación se realizó con scikit-learn 1.3 y XGBoost 2.0 en Python 3.11[13]-[16].

La explicabilidad se definió en dos niveles. En el nivel global, se calculó la importancia media absoluta de SHAP para ordenar variables. En el nivel local, cada predicción se traduce en una lista breve de factores que aumentan o reducen el riesgo de la gestante. LIME se consideró como método complementario para auditoría local, especialmente cuando se requiera comparar explicaciones entre modelos [17].

D. Validación de escritorio

Se generaron 3,600 registros sintéticos de gestantes, con distribución plausible de edad, índice de masa corporal, antecedentes, presión arterial, proteinuria y síntomas. El desenlace simulado fue preeclampsia diagnosticada durante el embarazo, con una prevalencia global de 6.5%. Los datos se dividieron en entrenamiento y prueba con proporción 70% y 30%, preservando la prevalencia del desenlace. Las métricas fueron AUC ROC, AUC PR, Brier score, sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo y curva de decisión. Esta última cuantifica beneficio neto para distintos umbrales de intervención y es apropiada cuando el objetivo no es sólo discriminar, sino decidir a quién intensificar seguimiento o referir [18].

El diseño de reporte se alineó con principios de TRIPOD+AI para estudios de modelos predictivos con aprendizaje automático y con PROBAST para evaluación de riesgo de sesgo [19]. En una fase empírica posterior, el protocolo deberá evaluarse con validación temporal, validación externa por región sanitaria, análisis de calibración, análisis de equidad y revisión de seguridad clínica.

La generación de datos sintéticos para validación de escritorio presenta limitaciones inherentes. La literatura reciente señala que los datos sintéticos pueden no capturar casos extremos clínicamente significativos y podrían amplificar sesgos existentes en los supuestos de la distribución original, representando patrones simulados sin incorporar la variabilidad completa del mundo real [20]. Por ello, los resultados reportados en este estudio deben interpretarse como verificación de coherencia interna del algoritmo, no como evidencia de rendimiento clínico generalizable. La técnica específica de simulación empleó distribuciones marginales basadas en estadísticas publicadas de la región centroamericana, con estructura de correlación entre variables derivada de la literatura clínica.

Dado que la prevalencia simulada de preeclampsia fue de 6.5%, se evaluaron tres estrategias para el manejo del desbalance: submuestreo aleatorio, SMOTE (Synthetic Minority Oversampling Technique) y ajuste del parámetro scale_pos_weight en XGBoost. Esta última estrategia fue seleccionada por no generar observaciones artificiales adicionales y por su integración nativa con el gradiente de la función de pérdida [21].

III. RESULTADOS

A. Comportamiento predictivo

La validación de escritorio mostró que los modelos lineales y de ensamble pueden operar con datos clínicos mínimos, aunque con perfiles distintos. La regresión logística obtuvo la mayor AUC ROC, mientras que XGBoost presentó mayor sensibilidad al umbral operativo de 8% y mejor utilidad para reducir falsos negativos en el escenario de triaje. Este patrón es consistente con el propósito del sistema: detectar gestantes que requieren reevaluación temprana o referencia, no confirmar diagnóstico; para más detalle ver Tabla II.

TABLA II
DESEMPEÑO COMPARATIVO EN CONJUNTO DE PRUEBA SINTÉTICO

	Regresión logística	**CART podado**	**Random Forest**	**XGBoost calibrable**	**Referencia clínica ACOG/SMFM (≥1 factor)**
AUC ROC (IC 95%)	0.800 (0.75–0.85)	0.752 (0.69–0.81)	0.780 (0.72–0.84)	0.786 (0.74–0.83)	0.653 (0.59–0.72)
AUC PR (IC 95%)	0.354 (0.28–0.44)	0.206 (0.14–0.29)	0.285 (0.21–0.38)	0.337 (0.26–0.42)	0.143 (0.09–0.22)
Br.	0.051	0.055	0.054	0.058	0.061
Sen. 8%	0.557	0.643	0.629	0.871	0.743
Esp.	0.846	0.735	0.798	0.47	0.561
VPP	0.2	0.144	0.177	0.102	0.105
VPN	0.965	0.967	0.969	0.981	0.969

Nota: Br. = Brier score. Sen. = Sensibilidad. Esp. = Especificidad

Como referencia clínica se incluyó un comparador basado en la suma simple de factores binarios reconocidos por ACOG/SMFM para estratificación de riesgo de preeclampsia, incluyendo nuliparidad, antecedente de preeclampsia, hipertensión crónica, diabetes pregestacional, gestación múltiple, edad materna ≥35 años e IMC ≥30 kg/m². Para mantener aplicabilidad en el primer nivel de atención, se utilizó un umbral operativo de ≥1 factor presente. En la validación sintética, este comparador presentó una AUC ROC de 0.653 (IC 95%: 0.59–0.72), AUC PR de 0.143 (IC 95%: 0.09–0.22), sensibilidad de 0.743 y especificidad de 0.561. En comparación, el modelo XGBoost calibrable alcanzó una AUC ROC de 0.786 (IC 95%: 0.74–0.83) y una AUC PR de 0.337 (IC 95%: 0.26–0.42), lo que representa un incremento absoluto de 0.133 en AUC ROC y de 0.194 en AUC PR. Estos resultados sugieren que el modelado no lineal aporta ganancia discriminativa sobre una regla clínica simple, sin reemplazar el criterio clínico ni los criterios de referencia inmediata.

B. Calibración y utilidad clínica

La calibración por deciles permitió revisar si el riesgo predicho coincide con la frecuencia observada en grupos ordenados por riesgo. La calibración es crítica en preeclampsia porque un modelo con buena discriminación, pero probabilidades exageradas puede saturar la red de referencia. La curva de decisión mostró que PRECLA-XAI-HN ofrece beneficio neto en umbrales compatibles con vigilancia

intensificada, especialmente entre 8% y 18%. Este comportamiento se ha descrito en las Fig. 1 y 2.

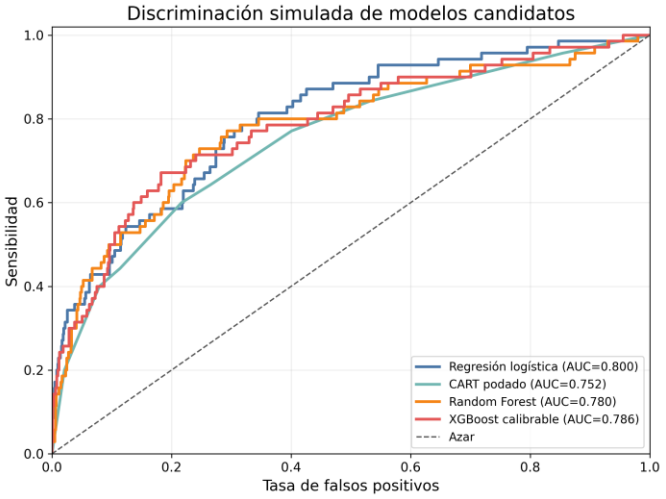


Fig. 1 Curvas de característica operativa del receptor (ROC) para los cuatro modelos candidatos evaluados en el conjunto de prueba sintético (n = 1,080). Se comparan regresión logística (AUC = 0.800), CART podado (AUC = 0.752), Random Forest (AUC = 0.780) y XGBoost calibrado (AUC = 0.786). La línea discontinua diagonal representa la clasificación aleatoria (AUC = 0.500). Los intervalos de confianza al 95% fueron calculados mediante 1,000 réplicas bootstrap. La regresión logística obtuvo la mayor AUC global, mientras que XGBoost ofreció mayor sensibilidad al umbral operativo de 8%.

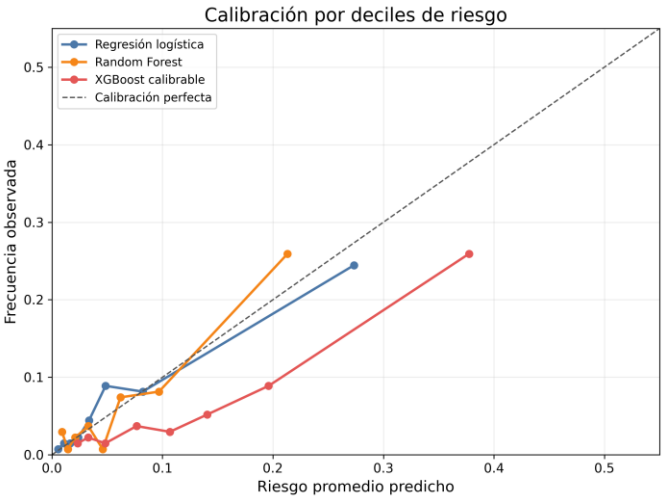


Fig. 2 Diagrama de calibración del modelo XGBoost calibrado, construido por deciles de riesgo predicho en el conjunto de prueba sintético. El eje horizontal muestra la probabilidad predicha media por decil y el eje vertical la proporción observada de preeclampsia en cada grupo. La línea diagonal (pendiente = 1, intercepto = 0) representa calibración perfecta.

C. Explicabilidad clínica

La explicación global identificó presión arterial media, presión diastólica, índice de masa corporal, edad materna y presión sistólica como los contribuyentes principales. Este resultado es clínicamente coherente porque la presión arterial media integra información sistólica y diastólica, mientras que los antecedentes y la antropometría capturan susceptibilidad basal. La utilidad práctica de SHAP no está en mostrar una gráfica al paciente, sino en entregar al trabajador de salud una explicación breve: el riesgo sube por presión elevada, antecedente obstétrico, comorbilidad o síntomas, y baja cuando la presión y los antecedentes son favorables. Ver Fig. 3, conteniendo el análisis de la curva de decisión.

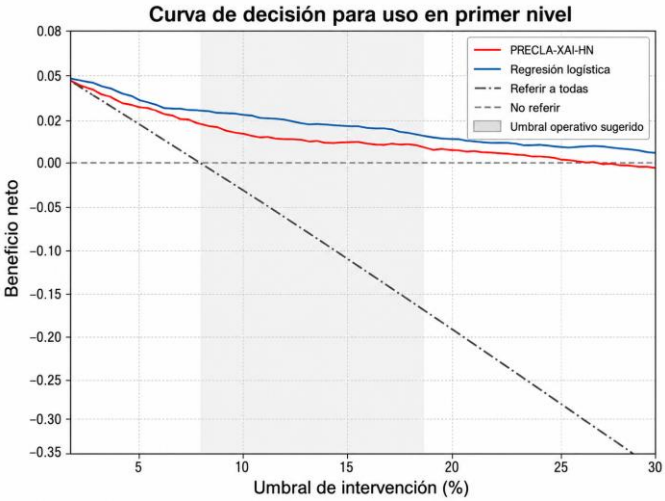


Fig. 3 Curva de análisis de decisión que muestra el beneficio neto del modelo XGBoost calibrado (línea sólida) comparado con las estrategias de tratar a todas las gestantes (línea gris descendente) y no tratar a ninguna (línea horizontal en cero), para umbrales de probabilidad entre 2% y 30%. El beneficio neto representa la ganancia clínica ajustada por la relación daño/beneficio implícita en cada umbral de intervención. PRECLA-XAI-HN muestra beneficio neto positivo sobre ambas estrategias de referencia en el rango de umbrales clínicamente relevantes entre 5% y 20%, lo que sugiere utilidad para decisiones de vigilancia intensificada y referencia oportuna.

Como ejemplo de explicación local, una gestante simulada de 38 años, primigesta, con presión arterial media de 105 mmHg, IMC de 33 y proteinuria 1+, obtuvo una probabilidad estimada de 0.24 (estrato 'Muy Alto'). Los tres principales factores contribuyentes según SHAP fueron: presión arterial media elevada (+0.09 sobre la base), proteinuria positiva (+0.05) e IMC elevado (+0.04). Los factores protectores identificados fueron ausencia de antecedente de preeclampsia (−0.02) y edad gestacional temprana al tamizaje (−0.01). Este tipo de salida permite al trabajador de salud comprender la lógica detrás de la clasificación sin requerir conocimiento algorítmico.

Se evaluó la equidad del modelo comparando sensibilidad y especificidad entre subgrupos definidos por zona (rural vs. urbana) y por grupo etario (<20, 20-34, ≥35 años). No se observaron diferencias de sensibilidad mayores a 5 puntos porcentuales entre grupos en la validación sintética, aunque esta evaluación requiere confirmación con datos reales estratificados por región sanitaria. La explicabilidad de las variables simples se detalla en la Fig. 4 y 5.

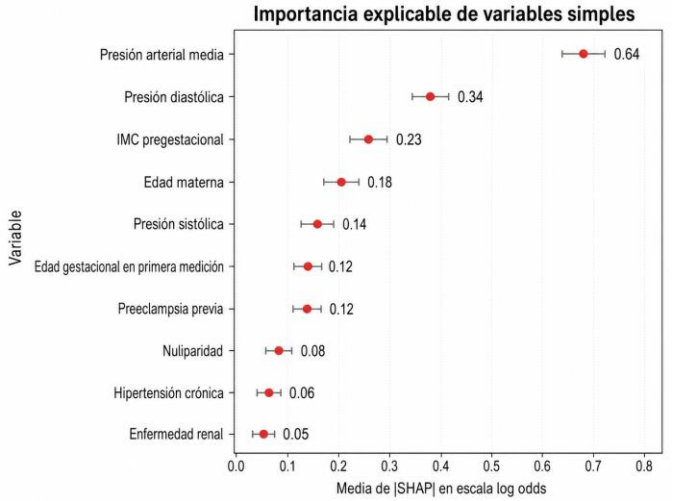


Fig. 4 Importancia global de las variables predictoras estimada mediante la media de los valores absolutos SHAP (|SHAP|) del modelo XGBoost calibrado en el conjunto de prueba sintético. Las barras horizontales representan la contribución media de cada variable a la magnitud de la predicción, y las líneas de error indican intervalos de confianza al 95% obtenidos por 1,000 réplicas bootstrap. La presión arterial media (PAM) fue la variable con mayor importancia explicativa, seguida por presión diastólica, índice de masa corporal, edad materna y presión sistólica. Las variables binarias de antecedentes y síntomas mostraron menor importancia media global pero pueden ser determinantes en predicciones individuales según el análisis local SHAP.

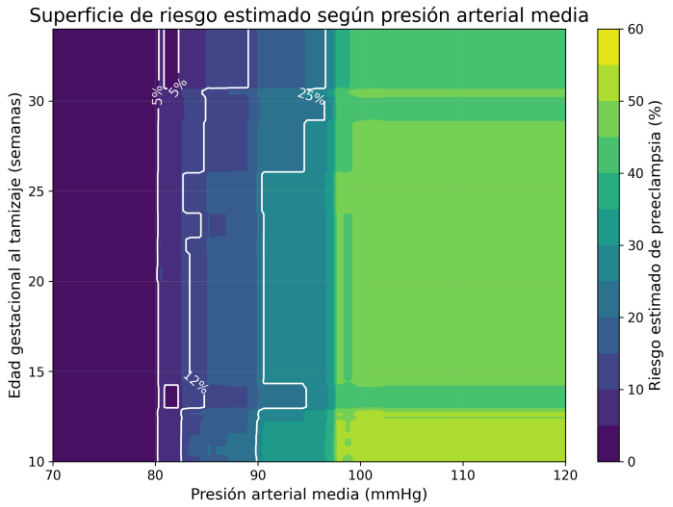


Fig. 5 Superficie de respuesta del riesgo estimado de preeclampsia en función de la presión arterial media (PAM, eje x) y la edad gestacional al momento del tamizaje (eje y), generada para un perfil de referencia: gestante de 28 años, IMC 25, sin antecedentes de preeclampsia, sin comorbilidades ni síntomas de alarma. La escala cromática representa la probabilidad predicha por el modelo XGBoost calibrado, desde azul (riesgo bajo, <5%) hasta rojo (riesgo alto, >25%). Se observa un gradiente de riesgo creciente con valores de PAM superiores a 90 mmHg, particularmente a partir de la semana 28 de gestación, lo que es coherente con la fisiopatología del continuo hipertensivo del embarazo.

D. Protocolo operativo propuesto

El protocolo transforma el puntaje probabilístico en una acción clínica. Ningún umbral reemplaza el criterio profesional. La regla de seguridad es explícita: presión arterial igual o mayor de 160/110 mmHg, síntomas neurológicos, epigastralgia severa, convulsión, alteración visual persistente o sospecha clínica de deterioro materno activan referencia inmediata, independientemente del puntaje calculado. La aspirina preventiva, cuando corresponda, debe seguir guías vigentes y evaluación clínica individual; la recomendación de la USPSTF indica 81 mg diarios después de las 12 semanas en personas con alto riesgo [22]. Los protocolos se detallan en la Tabla III y los indicadores en la Tabla IV.

TABLA III
PROTOCOLO OPERATIVO DE USO EN PRIMER NIVEL

Estrato	Criterio operativo	Acción en primer nivel	Tiempo máximo sugerido
Bajo	Riesgo <5% y sin alarma	Control prenatal habitual, repetir tamizaje en próxima visita	Según agenda ANC
Moderado	Riesgo 5% a <12% y sin alarma	Repetir presión, verificar técnica, consejería y reevaluación próxima	7 a 14 días
Alto	Riesgo 12% a <25% o incremento sostenido del riesgo	Consulta médica, proteinuria, plan de vigilancia, considerar referencia según red	72 horas
Muy alto	Riesgo ≥25% o combinación de HTA, proteinuria y comorbilidad	Referencia prioritaria a nivel con capacidad obstétrica	24 horas
Emergencia	PA ≥160/110 mmHg, síntomas neurológicos, epigastralgia severa o sospecha de eclampsia	Referencia inmediata, estabilización conforme protocolo local	Inmediato

TABLA IV
INDICADORES DE DESEMPEÑO EN PRIMER NIVEL

Estrato	Criterio operativo	Acción en primer nivel	Tiempo máximo sugerido
Cobertura de tamizaje	Gestantes con cálculo de riesgo / gestantes en control	≥85%	Detectar establecimientos con baja adopción
Tiempo a acción	Horas entre alerta alta y decisión documentada	<72 h	Medir respuesta clínica real
Tasa de referencia pertinente	Referencias con criterio clínico confirmado / referencias por alerta	≥70%	Evitar sobrecarga innecesaria
Eventos sin alerta previa	Preeclampsia diagnosticada sin alerta en 30 días previos	Tendencia descendente	Monitorear falsos negativos
Aceptabilidad	Trabajadores que consideran útil la explicación local	≥80%	Validar transferencia al flujo de trabajo
Equidad operativa	Diferencia de sensibilidad rural-urbana	<10 p.p.	Vigilar sesgo territorial

IV. Discusión

PRECLA-XAI-HN aborda una brecha concreta: la necesidad de anticipar riesgo obstétrico con datos disponibles en establecimientos de baja complejidad. La mayor parte de la literatura internacional sobre predicción de preeclampsia incorpora biomarcadores y ultrasonido especializado [7], [23]. Esos enfoques deben seguir siendo el estándar en centros con capacidad diagnóstica. Sin embargo, la innovación propuesta se ubica en otra capa del sistema: priorizar vigilancia, ordenar demanda y acelerar referencia cuando el primer contacto ocurre en un punto periférico.

Desde la perspectiva de implementación, la evidencia reciente subraya que la inteligencia artificial puede ser aprovechada para estratificación de riesgo y apoyo a la decisión clínica especialmente en países de ingresos bajos y medios, donde las disparidades en mortalidad materna reflejan inequidades en el acceso a atención de calidad [24]. Un estudio de validación con 14 profesionales de salud en Bangladesh demostró que las explicaciones híbridas que combinan reglas interpretables con importancia de variables tipo SHAP generan mayor confianza clínica que cualquiera de los dos métodos por separado, con un 71.4% de preferencia por explicaciones combinadas [25]. PRECLA-XAI-HN se alinea con este enfoque al integrar probabilidades calibradas, factores SHAP y una regla de seguridad explícita.

El modelo de competencia de riesgos de la Fetal Medicine Foundation (FMF), ampliamente utilizado para cribado de preeclampsia en el primer trimestre, combina factores maternos con presión arterial media, índice de pulsatilidad de arterias uterinas, PAPP-A y PlGF. Aunque su rendimiento reportado supera consistentemente AUC de 0.85 para preeclampsia pretérmino, su aplicabilidad depende de ultrasonido Doppler y biomarcadores séricos que no están disponibles de forma rutinaria en establecimientos de primer nivel en Honduras. PRECLA-XAI-HN no pretende reemplazar al modelo FMF, sino ofrecer una capa previa de priorización en entornos donde esos insumos diagnósticos no existen.

Los resultados simulados no deben interpretarse como evidencia clínica de efectividad. Su valor radica en mostrar que el algoritmo puede estructurarse con variables mínimas, producir explicaciones comprensibles y conectarse con un protocolo de acción. La literatura reciente confirma el potencial de los modelos de aprendizaje automático para predecir preeclampsia a partir de datos rutinarios, aunque también advierte que la validación externa sigue siendo indispensable [13], [14]. En ese sentido, PRECLA-XAI-HN debe avanzar hacia un estudio retrospectivo con registros perinatales reales, seguido de evaluación prospectiva pragmática en establecimientos seleccionados.

No obstante, la ausencia de validación con datos clínicos reales impide establecer su desempeño y utilidad efectiva en la toma de decisiones. Además, los datos sintéticos pueden subrepresentar casos extremos y reproducir sesgos derivados de los supuestos de simulación. La implementación real deberá evaluar integración al flujo asistencial, carga operativa, aceptabilidad del personal, seguridad y posibles efectos sobre la equidad entre establecimientos y poblaciones.

Desde la perspectiva tecnológica, la propuesta favorece interoperabilidad y bajo costo. Puede implementarse como módulo web, formulario móvil sin conexión permanente o componente del expediente perinatal. El cálculo del riesgo requiere pocos campos, puede ejecutarse localmente y puede sincronizar sólo resultados agregados para monitoreo. Esta arquitectura reduce dependencia de conectividad continua y evita enviar datos nominales innecesarios. Desde la perspectiva de negocios e innovación pública, el producto es transferible como paquete institucional: algoritmo versionado, protocolo de uso, tablero de indicadores, material de capacitación y auditoría de desempeño.

El componente explicable es central para adopción. Un trabajador de salud no necesita conocer la matemática de XGBoost para confiar en el sistema, pero sí necesita saber por qué una gestante fue clasificada como alta prioridad. La salida propuesta debe mostrar riesgo estimado, nivel de acción, principales factores contribuyentes y siguiente paso. Esta combinación responde a una exigencia ética y operativa: la IA en salud materna debe aumentar claridad clínica, no desplazar responsabilidad profesional.

Las limitaciones son relevantes. Primero, el conjunto de validación es sintético y no reproduce sesgos, errores de medición ni heterogeneidad real de establecimientos. Segundo, los umbrales propuestos son operativos y deben calibrarse con desenlaces locales, capacidad de referencia y costo de oportunidad. Tercero, variables como proteinuria por tira reactiva pueden tener calidad variable, especialmente si no se estandariza la técnica. Cuarto, la equidad debe evaluarse por región, ruralidad, edad, pertenencia étnica, idioma y acceso geográfico. El algoritmo debe auditarse periódicamente para evitar que la digitalización reproduzca desigualdades preexistentes. Quinto, existe el riesgo de sesgo de automatización (“automation bias”), en el cual el personal de salud podría confiar excesivamente en la alerta algorítmica y reducir su vigilancia clínica independiente. Para mitigar este riesgo, el protocolo enfatiza que PRECLA-XAI-HN es un complemento y no un sustituto del juicio profesional, y que la capacitación del personal debe incluir escenarios donde el algoritmo produce falsos negativos.

V. Conclusiones

La predicción explicable de preeclampsia con datos clínicos simples es una estrategia factible para fortalecer el primer nivel de atención, siempre que se implemente como apoyo a la decisión y no como diagnóstico autónomo. PRECLA-XAI-HN integra variables disponibles en consulta prenatal, modelos tabulares, explicación SHAP y un protocolo

de acción por estratos. Su contribución principal es operativa: convertir datos mínimos en una alerta trazable, comprensible y útil para priorizar vigilancia y referencia.

Adicionalmente, se recomienda que la validación empírica incluya análisis de calibración por subgrupo (edad, paridad, ruralidad), medición del Net Reclassification Improvement (NRI) comparado con el tamizaje clínico habitual, y evaluación de la latencia entre la alerta algorítmica y la acción clínica documentada como indicador de impacto operativo.

La siguiente etapa debe consistir en validación retrospectiva con registros perinatales reales, calibración por región sanitaria, evaluación prospectiva de seguridad, medición de aceptabilidad del personal y análisis de impacto en tiempos de referencia. Sólo después de ese proceso podría considerarse su incorporación gradual como módulo institucional de apoyo al control prenatal.

AGRADECIMIENTOS

Agradecemos el apoyo recibido del Instituto de Investigación en Ciencias Médicas y Derecho a la Salud (ICIMEDES) de la Facultad de Ciencias Médicas (FCM) y la Dirección de Investigaciones Científicas, Humanísticas y Tecnológicas (DICIHT) todas ellas de la Universidad Nacional Autónoma de Honduras (UNAH).

REFERENCIAS

[1] American College of Obstetricians and Gynecologists, "Gestational hypertension and preeclampsia: ACOG Practice Bulletin No. 222," *Obstet. Gynecol.*, vol. 135, no. 6, pp. e237–e260, Jun. 2020, doi: 10.1097/AOG.0000000000003891.

[2] L. A. Magee et al., "The 2021 International Society for the Study of Hypertension in Pregnancy classification, diagnosis and management recommendations for international practice," *Pregnancy Hypertens.*, vol. 27, pp. 148–169, Mar. 2022, doi: 10.1016/j.preghy.2021.09.008.

[3] World Health Organization, "Pre-eclampsia," *WHO Fact Sheet*, Dec. 10, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/pre-eclampsia. Accessed: May 18, 2026.

[4] Pan American Health Organization, "Honduras country profile," *Health in the Americas+*, Oct. 25, 2024. [Online]. Available: https://hia.paho.org/en/country-profiles/honduras. Accessed: May 18, 2026.

[5] C. E. Claudino Fajardo, "Mortalidad materna en Honduras, ¿qué podemos hacer en los servicios de salud?," *Rev. Méd. Hondur.*, vol. 92, no. Supl. 2, pp. S49–S51, 2024, doi: 10.5377/rmh.v92iSupl.2.19504.

[6] V. Kabeya, S. Tariq, A. Delicate, H. Chong, J. Aunger, H. Naidu, C. Dunlop, D. Yates, A. Muthirulandi, A. Muthirulandi, R. Lilford, S. Sobhy, and S. Thangaratinam, "Acceptability and use of clinical decision support tools in maternity settings: Systematic review of qualitative studies," *Eur. J. Obstet. Gynecol. Reprod. Biol*., vol. 314, p. 114718, Nov. 2025, doi: 10.1016/j.ejogrb.2025.114718. PMID: 40946542.

[7] L. C. Poon et al., "The International Federation of Gynecology and Obstetrics initiative on pre-eclampsia: A pragmatic guide for first-trimester screening and prevention," *Int. J. Gynecol. Obstet.*, vol. 145, Suppl. 1, pp. 1–33, May 2019, doi: 10.1002/ijgo.12802.

[8] V. Malik, N. Agrawal, S. Prasad *et al.*, "Prediction of preeclampsia using machine learning: A systematic review," *Cureus*, vol. 16, Art. no. e76095, 2024, doi: 10.7759/cureus.76095

[9] L. Liu, Q. Zhu, Y. Zong, X. Chen, W. Zhang, and J. Wang, "Machine learning prediction models for preeclampsia: Systematic review and meta-analysis," *J. Med. Internet Res*., vol. 28, p. e78714, Jan. 2026, doi: 10.2196/78714. PMID: 41554533

[10] Pan American Health Organization, "Perinatal Information System," *Centro Latinoamericano de Perinatología, Salud de la Mujer y Reproductiva*, PAHO/WHO. [Online]. Available: https://www.paho.org/en/clap/perinatal-information-system. Accessed: May 18, 2026.

[11] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Adv. Neural Inf. Process. Syst.*, vol. 30, Long Beach, CA, USA, 2017, pp. 4765–4774.

[12] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?': Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. (KDD)*, San Francisco, CA, USA, 2016, pp. 1135–1144, doi: 10.1145/2939672.2939778.

[13] A. Ranjbar et al., "Machine learning models for predicting preeclampsia: A systematic review," *BMC Pregnancy Childbirth*, vol. 24, Art. no. 6, Jan. 2024, doi: 10.1186/s12884-023-06220-1.

[14] S. A. Tiruneh, D. L. Rolnik, H. J. Teede, and J. Enticott, "Prediction of pre-eclampsia with machine learning approaches: Leveraging important information from routinely collected data," *Int. J. Med. Inform.*, vol. 192, Art. no. 105645, Dec. 2024, doi: 10.1016/j.ijmedinf.2024.105645.

[15] S. Li et al., "Improving preeclampsia risk prediction by modeling pregnancy trajectories from routinely collected electronic medical record data," *npj Digit. Med.*, vol. 5, Art. no. 68, Jun. 2022, doi: 10.1038/s41746-022-00612-x.

[16] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. (KDD)*, San Francisco, CA, USA, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[17] A. J. Vickers and E. B. Elkin, "Decision curve analysis: A novel method for evaluating prediction models," *Med. Decis. Making*, vol. 26, no. 6, pp. 565–574, Nov. 2006, doi: 10.1177/0272989X06295361.

[18] G. S. Collins et al., "TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods," *BMJ*, vol. 385, Art. no. e078378, Apr. 2024, doi: 10.1136/bmj-2023-078378.

[19] R. F. Wolff et al., "PROBAST: A tool to assess the risk of bias and applicability of prediction model studies," *Ann. Intern. Med.*, vol. 170, no. 1, pp. 51–58, Jan. 2019, doi: 10.7326/M18-1376.

[20] J. C. Ahn *et al.*, "Use of artificial intelligence-generated synthetic data to augment and enhance the performance of clinical prediction models in patients with alcohol-associated hepatitis and acute cholangitis," *Gastro Hep Adv.*, Art. no. 100643, Feb. 2025, doi: 10.1016/j.gastha.2025.100643

[21] Y. Ma, H. Lv, Y. Ma, X. Wang, L. Lv, X. Liang, and L. Wang, "Advancing preeclampsia prediction: A tailored machine learning pipeline integrating resampling and ensemble models for handling imbalanced medical data," *BioData Min.*, vol. 18, Art. no. 25, Mar. 2025, doi: 10.1186/s13040-025-00440-1. PMID: 40128863.

[22] K. W. Davidson et al., "Aspirin use to prevent preeclampsia and related morbidity and mortality: US Preventive Services Task Force recommendation statement," *JAMA*, vol. 326, no. 12, pp. 1186–1191, Sep. 2021, doi: 10.1001/jama.2021.14781.

[23] T. Li et al., "Prediction model of preeclampsia using machine learning based methods: A population based cohort study in China," *Front. Endocrinol.*, vol. 15, Art. no. 1345573, 2024, doi: 10.3389/fendo.2024.1345573.

[24] Z. Hoodbhoy, S. Martinez-Sanchez, M. I. Nisar, and E. Smith, "Editorial: Use of artificial intelligence to improve maternal and neonatal health in low-resource settings," *Front. Glob. Womens Health*, vol. 6, Art. no. 1675578, Sep. 2025, doi: 10.3389/fgwh.2025.1675578.

[25] F. Yesmin, N. Shirmin, and S. S. Bristy, "Bridging the trust gap: Clinician-validated hybrid explainable AI for maternal health risk assessment in Bangladesh," *arXiv preprint*, arXiv:2601.07866, Jan. 10, 2026.