

Catalina Peralta-Cortázar<sup>1a</sup>, Zulem Santiago-Loya<sup>2b</sup>, Talia Minerva Rivera-Villanueva<sup>3c</sup>, Daniel Omar Pérez-Godínez<sup>3d</sup>, Oscar Abraham José Padilla-Solís<sup>3e</sup>, Agustín Ramiro Urzúa-González<sup>3f</sup>, Alma Patricia González<sup>2g</sup>, Carlos Paque-Bautista<sup>2h</sup>, José Luis Felipe Luna-Anguiano<sup>4i</sup>, Gloria Patricia Sosa-Bustamante<sup>2j</sup>

### Resumen

**Introducción:** la estimación de la edad ósea (EO) con métodos automatizados puede eliminar la variación interindividual.

**Objetivo:** evaluar la EO mediante la creación de un algoritmo de inteligencia artificial (IA) en población pediátrica de la región del Bajío, México.

**Material y métodos:** estudio observacional, transversal, retrospectivo, analítico. Se incluyeron radiografías de la mano izquierda de menores de 18 años, obtenidas de la base de datos del Servicio de Radiología para crear el algoritmo de IA para estimar la EO. Se comparó el resultado de EO por IA con el obtenido por 2 observadores expertos con el método de Greulich y Pyle.

**Resultados:** se analizaron 271 radiografías para evaluar la EO y esta fue similar entre los observadores 1, 2 e IA, al considerar todas las imágenes ( $p = 0.68$ ). El tiempo de duración de la estimación de la EO fue mayor con la IA ( $p < 0.001$ ). La medición con IA no mostró diferencias entre edad cronológica (EC) y EO, al considerar tanto la totalidad de imágenes ( $p = 0.12$ ), como cuando se distribuyeron por grupos etarios  $< 6$  años,  $6 a < 10$  años y  $\geq 10$  años ( $p = 0.60$ ,  $p = 0.06$ ,  $p = 0.67$ , respectivamente). Los coeficientes de correlación de concordancia (CCC) de EO más altos se registraron cuando se evaluaron todas las imágenes (observador 1 y 2, observador 1 e IA, y observador 2 e IA [ $p < 0.001$ , en los 3 escenarios]).

**Conclusiones:** el algoritmo de IA permite la estimación objetiva de la EO de niños y adolescentes como primer entrenamiento; su perfeccionamiento optimizará su uso y utilidad en la práctica clínica.

### Abstract

**Background:** Bone age (BA) estimation with automated methods can eliminate interindividual variation.

**Objective:** To evaluate BA by creating an artificial intelligence (AI) algorithm in a pediatric population from the Bajío region of Mexico.

**Material and methods:** Observational, cross-sectional, retrospective, analytical study. Left-hand radiographs of children under 18 years of age, obtained from the Radiology Department database were included to create the AI algorithm for estimating BA. The BA result obtained by AI was compared with that obtained by 2 expert observers using the Greulich and Pyle method.

**Results:** 271 radiographs were analyzed to assess BA and this was similar between observers 1, 2, and AI when considering all images ( $p = 0.68$ ). The time taken to estimate BA was longer with AI ( $p < 0.001$ ). AI measurement showed no differences between chronological age (CA) and BA when considering both the total number of images ( $p = 0.12$ ) and when they were distributed by age group:  $< 6$  years,  $6 to < 10$  years, and  $\geq 10$  years ( $p = 0.60$ ,  $p = 0.06$ ,  $p = 0.67$ , respectively). The highest EO concordance correlation coefficients (CCCs) were recorded when all images were evaluated (observer 1 and 2, observer 1 and AI, and observer 2 and AI [ $p < 0.001$ , in all 3 scenarios]).

**Conclusions:** The AI algorithm allows for objective estimation of BA in children and adolescents as a first training approach; its refinement will optimize its use and utility in clinical practice.

<sup>1</sup>Instituto Mexicano del Seguro Social, Centro Médico Nacional del Bajío, Hospital de Gineco Pediatría No. 48, División de Pediatría. León, Guanajuato, México

<sup>2</sup>Instituto Mexicano del Seguro Social, Centro Médico Nacional del Bajío, Hospital de Gineco Pediatría No 48, Dirección de Educación e Investigación en Salud. León, Guanajuato, México

<sup>3</sup>Grupo Multidisciplinario de Investigación en Inteligencia Artificial en Salud. León, Guanajuato, México

<sup>4</sup>Instituto Mexicano del Seguro Social, Centro Médico Nacional del Bajío, Hospital de Gineco Pediatría No 48, Dirección General. León, Guanajuato, México

ORCID: [0000-0002-8905-9863](https://orcid.org/0000-0002-8905-9863)<sup>a</sup>, [0009-0009-0811-556X](https://orcid.org/0009-0009-0811-556X)<sup>b</sup>, [0000-0003-0186-9720](https://orcid.org/0000-0003-0186-9720)<sup>c</sup>, [0009-0006-0237-3347](https://orcid.org/0009-0006-0237-3347)<sup>d</sup>, [0009-0000-0043-5512](https://orcid.org/0009-0000-0043-5512)<sup>e</sup>, [0000-0002-9403-0686](https://orcid.org/0000-0002-9403-0686)<sup>f</sup>, [0000-0002-3401-7519](https://orcid.org/0000-0002-3401-7519)<sup>g</sup>, [0000-0002-2658-0491](https://orcid.org/0000-0002-2658-0491)<sup>h</sup>, [0000-0003-3739-8334](https://orcid.org/0000-0003-3739-8334)<sup>i</sup>, [0000-0002-8460-4965](https://orcid.org/0000-0002-8460-4965)<sup>j</sup>

### Palabras clave

Inteligencia Artificial  
Determinación de la Edad por el Esqueleto  
Radiografía  
Niño  
Adolescentes



### Keywords

Artificial Intelligence  
Age Determination by Skeleton  
Radiography  
Child  
Adolescents

**Fecha de recibido:** 19/07/2025

**Fecha de aceptado:** 08/09/2025

### Comunicación con:

Gloria Patricia Sosa Bustamante  
 [patriciasosab@hotmail.com](mailto:patriciasosab@hotmail.com)  
 477 717 4800, extensión 31804

**Cómo citar este artículo:** Peralta-Cortázar C, Santiago-Loya Z, Rivera-Villanueva TM, *et al.* Algoritmo de inteligencia artificial para la estimación de la edad ósea. Rev Med Inst Mex Seguro Soc. 2026;64(3):e6753. doi: 10.5281/zenodo.18715240

## Introducción

La evaluación de la edad ósea (EO) es una herramienta importante en el diagnóstico de enfermedades endocrinas y metabólicas relacionadas con el crecimiento y desarrollo. La EO es un indicador de madurez biológica independiente de la talla.<sup>1</sup>

Existen métodos tradicionales para evaluar la EO y el más utilizado es Greulich-Pyle, el cual se basa en un atlas de 57 radiografías de referencia (31 para hombres, 26 para mujeres).<sup>2,3</sup> Otro es el método de Tanner-Whitehouse, el cual se basa en un sistema de puntuación que evalúa 20 huesos, tanto radio, cúbito, huesos cortos, como, los huesos del carpo. Con este método se calcula la puntuación final y se transforma en EO. El método de Tanner-Whitehouse fue estandarizado por medio de una muestra de niños británicos y se han actualizado 3 versiones, la última es del 2001 y fue nombrada TW3.<sup>4</sup>

El método de Bayley y Pinneau se basa en el porcentaje de talla adulta alcanzada a una edad cronológica determinada. Si el pronóstico de talla adulta se encuentra entre  $\pm 5$  cm de la talla diana, orienta a que existe concordancia de la talla del niño con la de sus padres, lo que indica que lo más probable es que no exista alguna patología del crecimiento. Este método utiliza la lectura del atlas de EO por Greulich-Pyle.<sup>5</sup>

Hay métodos automatizados que destacan para la estimación de la EO, como el *Computer Aided Skeletal Age Scores* y *BoneExpert* (Visian, Dinamarca), cuya ventaja principal es la eliminación de la variación interindividual, la cual oscila entre 0.45 a 0.82 años.<sup>6</sup> Estos métodos automáticos realizan la evaluación cuantitativa de los cambios morfológicos de varios huesos, lo cual resulta útil, pues el ojo humano no es óptimo para graduar dichos cambios, aunque sí puede detectar alteraciones cualitativas.<sup>7</sup>

La determinación de la EO convencional está asociada a la variabilidad dependiente de la interpretación humana, por lo que ha surgido el auge de nuevos métodos de determinación automática.<sup>8</sup>

Dadas las desventajas y dificultades que presenta la estimación de la EO por el humano, se requiere de métodos objetivos y eficientes en la rutina diaria y es así que la inteligencia artificial (IA) tiene potencial para reducir las variaciones intra- e interobservador, por lo que está obteniendo popularidad en el campo médico, especialmente en radiología.<sup>9,10</sup>

En años recientes, con el rápido desarrollo de algoritmos de aprendizaje profundo y la mejora del *hardware* informá-

tico, el *software* de diagnóstico asistido por IA ha comenzado a aplicarse en los hospitales, entre ellos, el *software* asistido por IA para la EO, ya que ha logrado buen rendimiento debido a que sus resultados han mostrado tanta precisión como la de los expertos.<sup>11,12</sup>

La necesidad de precisión y exactitud en la estimación de la EO en la edad pediátrica, así como la alta demanda de valoraciones por especialistas (que conlleva a retraso en la atención) justifica la necesidad de la creación de un algoritmo de IA en población mexicana; por lo tanto, el objetivo de la presente investigación fue estimar la EO mediante la creación e implementación de un algoritmo de IA en población pediátrica de la región del Bajío, México.

## Material y métodos

Estudio observacional, transversal, retrospectivo, analítico, realizado en la Unidad Médica de Alta Especialidad Hospital de Gineco Pediatría No. 48, del Centro Médico Nacional del Bajío, Instituto Mexicano del Seguro Social (IMSS), ubicada en León, Guanajuato, México. Se incluyeron radiografías de mano izquierda realizadas por alguna indicación médica, de pacientes con edad cronológica entre 2 y 17 años 11 meses, almacenadas en la base de datos del Sistema Infosalud Versión 0.18.17.0. Se excluyeron radiografías de pacientes en tratamiento por deficiencia de hormona de crecimiento, pubertad precoz, enfermedades del metabolismo del calcio, con insuficiencia renal, enfermedad reumatológica o que se encontraran con tratamiento esteroideo. Se consideraron como criterio de eliminación las radiografías de carpo ilegibles, incompletas o con artefactos.

La presente investigación se consideró sin riesgo de acuerdo con el artículo 17 del Reglamento de la Ley General de Salud en Materia de Investigación para la Salud. No se requirió de consentimiento informado, dada la naturaleza retrospectiva del estudio, por lo que fue autorizada la carta de excepción de consentimiento informado con el debido manejo de privacidad de los datos. El estudio fue aprobado por el Comité de Ética para la Investigación y el Comité Local de Investigación con el número de registro institucional R-2024-1002-001.

Se llevó a cabo la obtención de las imágenes radiológicas, con las que se creó y entrenó el algoritmo de IA. Para la implementación de los algoritmos se utilizaron programas de código abierto, específicamente, Supervisely para el etiquetado de las imágenes, MicroDicom para la transformación de las imágenes, Python como lenguaje de programación y las librerías de Tensorflow para su implementación. La información personal de los pacientes fue anonimizada antes del análisis. Se utilizaron las radiografías de la mano

izquierda para el desarrollo del modelo de IA, distribuidas de la siguiente manera: 80% para entrenamiento y 20% para validación. Las imágenes que por defecto se almacenaron en formato DICOM fueron convertidas en formato JPEG y anonimizadas utilizando el programa MicroDicom para su manipulación. Las máscaras se realizaron utilizando el programa en línea de etiquetado Supervisely. Las imágenes se normalizaron para tener una escala de grises uniforme y se redimensionaron a 256 x 256 píxeles. Se aplicaron técnicas de rotación, *zoom* y *windowing* para aumentar la variabilidad del conjunto de datos y mejorar la generalización del modelo. Se utilizó ResNet-50, una red neuronal convolucional de 50 capas (48 capas convolucionales, una capa MaxPool y una capa AvgPool). Se usó Adam como optimizador debido a su eficiencia en convergencia y Mean Squared Error como función de pérdida. Se entrenó el modelo durante 50 épocas, ya que con base en experimentos previos se consideró como cantidad suficiente para alcanzar la convergencia en tareas similares. Se distribuyó el conjunto de datos en 80% para entrenamiento y 20% para validación; se utilizó la validación cruzada para monitorizar el *overfitting*. Se empleó el error cuadrático medio como métrica principal, dado que es un estándar en tareas de regresión. Una vez obtenido un resultado aceptable mayor, la IA estimó la EO de cada imagen en específico y se comparó su resultado con el obtenido por el método de Greulich y Pyle con 2 especialistas expertos independientes que desconocían la edad cronológica del paciente.

El cálculo del tamaño de muestra se realizó por el coeficiente de correlación, de acuerdo con lo reportado por Artioli *et al.*,<sup>11</sup> en el que se reportó una desviación de la variable de interés, edad de  $1.1 \pm 0.11$  años, amplitud del intervalo 0.22, amplitud estandarizada del intervalo obtenida con la fórmula:  $\text{amplitud total/DE}$ , que correspondió a  $0.22/1.1 = 0.2$ , con un nivel de confianza de 95%. Se requirió un total de 271 radiografías para este estudio.

Se realizó el análisis descriptivo de las variables numéricas, así como el reporte de frecuencias y porcentajes de variables nominales. Se utilizó la prueba de normalidad de Kolmogorov-Smirnov y las variables mostraron libre distribución. Las variables numéricas se analizaron con prueba de rangos de Wilcoxon entre grupos pareados, así como prueba de Kruskal-Wallis entre 3 grupos independientes, con corrección de Bonferroni. Se evaluó la concordancia de la medición de la EO con el coeficiente de correlación de concordancia (CCC) y con el gráfico de Bland-Altman. Se consideró significación estadística con un valor de  $p < 0.05$ . Se utilizó el programa estadístico SPSS versión 29.0.2.0 (20)<sup>®</sup>.

## Resultados

Se analizaron 271 radiografías de pacientes pediátricos con proyección anteroposterior del carpo. Se desarrolló y validó el modelo de aprendizaje profundo para predecir la EO de pacientes pediátricos mediante el análisis de radiografías de las manos. La metodología adoptada se dividió por etapas clave que incluyeron la recolección de datos, el preprocesamiento de imágenes, la generación de máscaras con un modelo basado en regiones de interés (RoIs), la aplicación de transformaciones para aumentar el conjunto de datos y la implementación del modelo VGG16 para la predicción de la edad.

Del total de imágenes analizadas, 60.89% ( $n = 165$ ) pertenecieron a mujeres y 39.11% ( $n = 106$ ) a hombres. La distribución por grupo de edad y sexo se muestra en el cuadro I.

Durante la determinación de la EO, se encontró diferencia significativa entre el tiempo empleado para la identificación de la EO del observador 1, observador 2 e IA (cuadro II).

Al comparar la edad cronológica (EC) con la EO en todos los sujetos de estudio, solo la IA mostró resultados similares entre ambas mediciones; sin embargo, en el grupo etario  $< 6$  años y en el de  $6 \text{ a } < 10$  años, la EC y EO fueron diferentes entre los 3 observadores; solo en el grupo  $\geq 10$  años, la EC y la EO mostraron resultados similares entre los 3 observadores (cuadro III).

El tiempo de estimación de la EO entre el observador 1, observador 2 e IA fue diferente en todos los casos, así como por grupo etario (cuadro IV).

La concordancia entre los observadores 1, 2 y la IA se muestra en el cuadro V en la totalidad de imágenes y por grupos etarios. También se muestra la relación entre las variables pareadas de EO entre los observadores y la IA en la totalidad de la muestra (figura 1).

## Discusión

En el presente estudio, se evaluó la EO mediante un

**Cuadro I** Distribución de las imágenes por edad y sexo

Variables	Todos $n = 271$	Hombres $n = 106$	Mujeres $n = 165$
	$n$ (%)	$n$ (%)	$n$ (%)
EC $< 6$ años	48 (17.71)	21 (19.81)	27 (16.36)
EC $6 \text{ a } < 10$ años	79 (29.15)	18 (16.98)	61 (36.97)
EC $\geq 10$ años	144 (53.14)	67 (63.21)	77 (46.67)

EC: edad cronológica

Prueba de chi cuadrada = 12.60,  $p = 0.001$

**Cuadro II** Comparación del tiempo para la estimación de la EO entre observador 1, observador 2 e IA

Variable	Observador 1 <i>n</i> = 271	Observador 2 <i>n</i> = 271	IA <i>n</i> = 271	<i>p</i>
	Mediana (RIC)	Mediana (RIC)	Mediana (RIC)	
Edad ósea (años)*	11.5 (7-13.5)	12 (6.8-13.5)	10.2 (7.4-13.4)	0.68
Tiempo (s)†	10.5 (6.5-14.7)	7 (5-9)	20.6 (19.6-22.1)	< 0.001

Para las medianas y los rangos intercuartílicos (RIC) se empleó prueba de Kruskal-Wallis

IA: inteligencia artificial

\*Prueba Bonferroni, observador 1 frente a IA, *p* = 1.00; observador 2 frente a IA, *p* = 1.00; observador 1 frente a Observador 2, *p* = 1.00

†Prueba Bonferroni, observador 1 frente a IA, *p* < 0.001; observador 2 frente a IA, *p* < 0.001; observador 1 frente a observador 2, *p* < 0.001

**Cuadro III** Comparación de EC y EO en los diferentes grupos etarios en el observador 1, observador 2 e IA

Variable	Todos <i>n</i> = 271			Grupo < 6 años <i>n</i> = 48			Grupo 6 a < 10 años <i>n</i> = 79			Grupo ≥ 10 años <i>n</i> = 144		
	EC	EO	<i>p</i>	EC	EO	<i>p</i>	EC	EO	<i>p</i>	EC	EO	<i>p</i>
	M (RIC)	M (RIC)		M (RIC)	M (RIC)		M (RIC)	M (RIC)		M (RIC)	M (RIC)	
O1	10.3 (7-13.1)	11.5 (7-13.5)	0.01	3.3 (2.4-5)	2.5 (1.5-3.7)	0.0004	9.4 (7.9-12.9)	9 (7-11.3)	< 0.0001	12.5 (11-14.7)	13.1 (12-14.3)	0.15
O2	10.3 (7-13.1)	12 (6.8-13.5)	0.01	3.3 (2.4-5)	2.5 (1.5-3.8)	0.0008	9.4 (7.9-12.9)	10 (6.8-11)	< 0.0001	12.5 (11-14.7)	13 (12-14)	0.25
IA	10.3 (7-13.1)	10.2 (7.4-13.4)	0.12	3.3 (2.4-5)	3.5 (1.8-4.9)	0.60	9.4 (7.9-12.9)	8.1 (7.4-9.8)	0.06	12.5 (11-14.7)	13.2 (11.3-15.0)	0.67

Para la mediana y los rangos intercuartílicos se empleó prueba de rangos de Wilcoxon

O1: observador 1; O2: observador 2; EC: edad cronológica; EO: edad ósea; IA: inteligencia artificial; M: mediana; RIC: rangos intercuartílicos

**Cuadro IV** Tiempo de estimación de la EO del observador 1, observador 2 e IA

Variables	Todos <i>n</i> = 271			
	Mediana (RIC)	Mediana (RIC)	Mediana (RIC)	<i>p</i>
	Observador 1	Observador 2	IA	
Tiempo de medición EO*	10.5 (6.5-14.7)	7 (5-9)	20.6 (19.6-22.1)	< 0.001
	Grupo < 6 años <i>n</i> = 48			
	Observador 1	Observador 2	IA	<i>p</i>
Tiempo de medición EO†	13.3 (8.6-16.5)	8.6 (6.5-11.4)	19.9 (18.8-21.3)	< 0.001
	Grupo 6 a < 10 años <i>n</i> = 79			
	Observador 1	Observador 2	IA	<i>p</i>
Tiempo de medición EO‡	9.3 (5.7-11.3)	6.7 (4.8-9.1)	19.7 (19.2-21.1)	< 0.001
	Grupo ≥ 10 años <i>n</i> = 144			
	Observador 1	Observador 2	IA	<i>p</i>
Tiempo de medición EO§	7.4 (4.3-12.1)	5.4 (4.2-7.6)	20.6 (19.6-22.1)	< 0.001

Para la mediana y los rangos intercuartílicos (RIC) se usó la prueba de Kruskal-Wallis

EO: edad ósea; IA: inteligencia artificial; RIC: rangos intercuartílicos

\*Prueba Bonferroni, observador 1 frente a IA, *p* < 0.001; observador 2 frente a IA, *p* < 0.001; observador 1 frente a observador 2, *p* < 0.001

†Prueba Bonferroni, observador 1 frente a IA, *p* < 0.001; observador 2 frente a IA, *p* < 0.001; observador 1 frente a observador 2, *p* = 0.24

‡Prueba Bonferroni, observador 1 frente a IA, *p* < 0.001; observador 2 frente a IA, *p* < 0.001; observador 1 frente a observador 2, *p* = 1.00

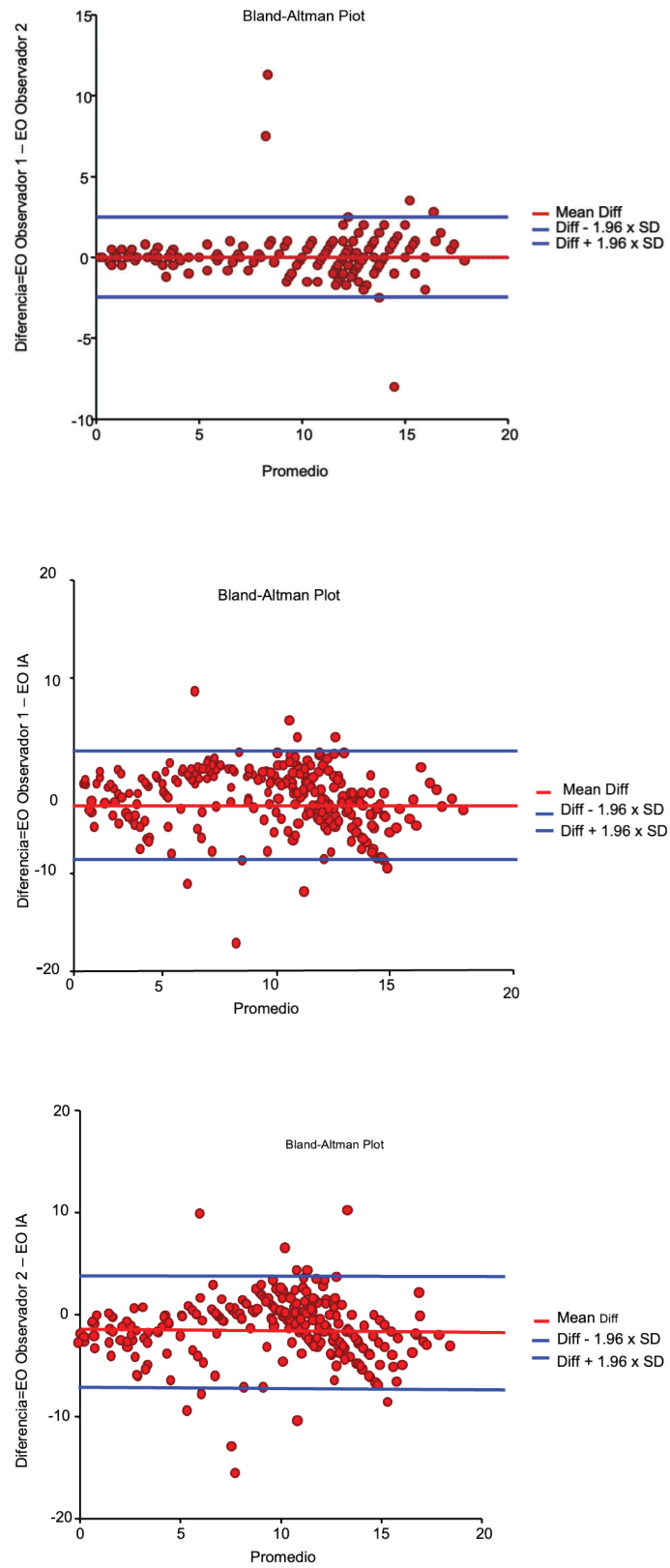
§Prueba Bonferroni, observador 1 frente a IA, *p* < 0.001; observador 2 frente a IA, *p* < 0.001; observador 1 frente a observador 2, *p* < 0.001

**Cuadro V** Concordancia de la estimación de la EO entre los observadores 1, 2 y la IA

Todos, n = 271						
Observador 1	Observador 2					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B(0)	p	Pendiente B (1)	p
	0.96 (0.94-0.96)	0.92	0.43	0.02	0.95	< 0.001
Observador 1	IA					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B(0)	p	Pendiente B (1)	p
	0.84 (0.80-0.87)	0.70	1.38	< 0.001	0.87	< 0.001
Observador 2	IA					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B(0)	p	Pendiente B (1)	p
	0.82 (0.78-0.85)	0.67	1.54	< 0.001	0.85	< 0.001
< 6 años, n = 48						
Observador 1	Observador 2					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B(0)	p	Pendiente B (1)	p
	0.98 (0.97-0.99)	0.97	0.11	0.19	0.94	< 0.001
Observador 1	IA					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B(0)	p	Pendiente B (1)	p
	0.50 (0.27-0.67)	0.28	1.00	0.06	0.50	< 0.001
Observador 2	IA					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B(0)	p	Pendiente B (1)	p
	0.49 (0.26-0.67)	0.27	1.00	< 0.07	0.52	< 0.001
6 a < 10 años, n = 79						
Observador 1	Observador 2					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B (0)	p	Pendiente B (1)	p
	0.87 (0.80-0.91)	0.76	1.81	< 0.001	0.79	< 0.001
Observador 1	IA					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B(0)	p	Pendiente B (1)	p
	0.55 (0.38-0.68)	0.33	3.23	0.001	0.67	< 0.001
Observador 2	IA					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B (0)	p	Pendiente B (1)	p
	0.44 (0.26-0.59)	0.22	3.79	0.001	0.62	< 0.001
≥ 10 años, n = 144						
Observador 1	Observador 2					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B (0)	p	Pendiente B (1)	p
	0.77 (0.70-0.83)	0.60	3.00	< 0.001	0.77	< 0.001
Observador 1	IA					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B (0)	p	Pendiente B (1)	p
	0.38 (0.24-0.51)	0.15	9.07	< 0.001	0.31	< 0.001
Observador 2	IA					
	Coeficiente de regresión					
	CCC (IC 95%)	R <sup>2</sup>	Intecepto B (0)	p	Pendiente B (1)	p
	0.37 (0.23-0.50)	0.14	9.06	0.001	0.31	< 0.001

CCC: coeficiente de correlación de concordancia; IC: intervalo de confianza; EO: edad ósea; IA: inteligencia artificial

Figura 1 Representación de la relación de la EO pareada entre los observadores



Gráficos de Bland-Altman  
EO: edad ósea; IA: inteligencia artificial

algoritmo de IA creado para el área médica que atiende pacientes pediátricos. Es importante destacar que el algoritmo de IA desarrollado para determinar la EO aún se encuentra en etapa de perfeccionamiento. Se demostró que la evaluación de la EO por el algoritmo de IA fue similar al de los 2 observadores expertos. Además, los CCC más altos en la evaluación de la EO se identificaron entre los observadores 1 y 2, y entre los observadores 1 y 2 con la IA cuando se analizaron en conjunto todas las radiografías, sin categorizar por grupos etarios.

Si bien es adecuado comentar que el uso de la IA en áreas diferentes a la medicina lleva ya varias décadas desde su aplicación formal, en las ciencias médicas aún es un área en desarrollo e investigación.<sup>13,14</sup> Hassan *et al.*<sup>15</sup> han descrito que el desarrollo de IA para su uso en la práctica clínica es todo un reto; aunque los algoritmos tengan un buen rendimiento predictivo, esto no garantiza que incidan en la mejora de la toma de decisiones clínicas.

Existen pocas aunque crecientes investigaciones en cuanto al desempeño de algoritmos de IA para la estimación de la EO, incluso en población mexicana, como lo reportan Chávez-Vázquez *et al.*,<sup>16</sup> que compararon la estimación de la EO por observadores expertos con la de un algoritmo de IA comercial (BoneXpert®) y encontraron que existe una alta variabilidad interobservador al realizar la evaluación manual por el experto, comparada con el algoritmo de IA; esto a diferencia de la presente investigación en la que se encontró una alta concordancia entre los observadores expertos 1 y 2, y entre los observadores 1 y 2 con la IA cuando se incluyeron las imágenes de todas las edades; sin embargo, la concordancia disminuyó al clasificar por grupos etarios a las imágenes evaluadas, al explorar esta primera fase del algoritmo creado.

En un estudio realizado en China, Yang *et al.*<sup>17</sup> desarrollaron un algoritmo de estimación de EO en tiempo real y obtuvieron una estimación con exactitud de 0.39 años en hombres y 0.32 años en mujeres; en este estudio además los algoritmos fueron capaces de determinar la EO en un promedio de 26 milisegundos, lo cual difiere de los resultados del presente estudio, ya que el algoritmo de IA tomó de forma general una mediana de tiempo de 20.6 segundos para la evaluación y al analizarlo por grupo de edad una mediana de 19.9 segundos para los < 6 años, 19.7 segundos para las edades entre 6 a < 10 años y 20.6 segundos para las imágenes óseas de ≥ 10 años. Esto podría explicarse por la capacidad del equipo cómputo en el que se realizó el algoritmo y al mejorar su capacidad se podría reducir el tiempo de respuesta del modelo.

Otro uso que se ha dado a los algoritmos de IA de estimación de EO es en la predicción de la talla adulta mediante IA;

Huang *et al.*<sup>18</sup> utilizaron algoritmos de IA para estimar la EO de adolescentes mujeres en etapa puberal y la predicción de talla adulta, y los autores observaron hallazgos similares a los de la presente investigación, ya que las diferencias en las EO medidas aumentaron con el aumento de la EO.

Un tema central que hay que considerar, es la variabilidad interobservador en la predicción de la EO; en el estudio realizado en China por Nguyen *et al.*,<sup>12</sup> los autores demostraron que el algoritmo de IA creado tuvo rendimiento significativamente más alto que el mostrado por radiólogos en términos de error estimado medio en la predicción de la EO y en términos de correlación y sesgo.

El uso de IA para evaluar la EO se ha aplicado de manera global en todo tipo de poblaciones, como se demostró en el estudio realizado por Alshamrani *et al.*,<sup>19</sup> en cuyos resultados indicaron que la evaluación de la EO manual y con el programa BoneXpert®, en su versión derivada de métodos Greulich & Pyle y Tanner & Whitehouse, se pueden aplicar a población de mujeres y solo la BoneXpert® derivada de Tanner & Whitehouse se puede aplicar sin precaución en hombres de Arabia Saudita, por lo que concluyen que el número relativamente pequeño de radiografías incluidas en cada grupo de edad para el análisis de BoneXpert en comparación con la calificación manual pudo haber contribuido a las diferencias entre BoneXpert y la calificación EO derivada manualmente, situación similar a la identificada en la presente investigación al establecer las categorías por grupos etarios y que se podría considerar como área de oportunidad para futuras investigaciones.

Por ende, con el incremento del uso de IA para estimación de EO cada vez hay más evidencia de que tiene una mayor capacidad predictiva y un menor riesgo de sesgo.<sup>20</sup>

Entre las fortalezas de la presente investigación destaca que constituye uno de los primeros esfuerzos para el desarrollo de un modelo de IA para la estimación de EO en población pediátrica mexicana de la región del Bajío, con la proyección de que una vez consolidado pueda ponerse a disposición de manera gratuita en el IMSS. Cabe señalar que el algoritmo presentado es aún un prototipo inicial que requiere ser perfeccionado y entrenado con un mayor número de imágenes para optimizar su desempeño. Este estudio preliminar representa, por tanto, un punto de partida valioso para futuras investigaciones orientadas a incrementar la precisión y la capacidad del *software* en la determinación de EO en pacientes pediátricos.

Si bien los resultados preliminares son prometedores, entre las limitantes del estudio se encuentra que la precisión y la robustez del modelo aún pueden mejorar con la incorporación de un mayor volumen de datos para entrenar ade-

cuadramente al algoritmo de IA y este proceso de ampliación del conjunto de datos permita que el modelo aprenda una variedad más amplia de casos y, así, mejorar su capacidad para generar resultados más precisos, con rendimiento óptimo del algoritmo de IA para la estimación de EO.

## Conclusiones

El algoritmo de IA desarrollado constituye un prototipo inicial que permite estimar de manera objetiva la EO en niños y adolescentes. Aunque corresponde a un primer entrenamiento del modelo, sus resultados sugieren el potencial de esta herramienta como apoyo en la práctica clínica. La continuidad en su perfeccionamiento y entrenamiento con bases de datos más amplias permitirá optimizar su precisión y, eventualmente, alcanzar de forma independiente un nivel de desempeño comparable o incluso superior al obtenido mediante la evaluación experta con el método de Greulich-Pyle.

Se espera que el *software* inicial, desarrollado como prototipo en esta etapa del proyecto, pueda ser de utilidad en el IMSS y que su uso no implique costo alguno para la institución. Al haber sido diseñado a partir de pacientes mexicanos de la región del Bajío, se proyecta que este sistema pueda implementarse con plena accesibilidad en las clínicas del IMSS, lo cual facilitaría la estimación de la EO mediante IA. No obstante, es fundamental reconocer las limitaciones actuales del modelo y la necesidad de perfeccionarlo para incrementar su aplicabilidad clínica. Asimismo, es importante subrayar que este algoritmo no sustituye la valoración médica, sino que debe considerarse como una herramienta complementaria en el proceso de determinación de la EO.

---

**Declaración de conflicto de interés:** los autores han completado y enviado la forma traducida al español de la declaración de conflictos potenciales de interés del Comité Internacional de Editores de Revistas Médicas, y no fue reportado alguno relacionado con este artículo.

## Referencias

1. Jani G, Patel B. Charting the growth through intelligence: A SWOC analysis on AI-assisted radiologic bone age estimation. *Int J Legal Med.* 2025;139(2):679-94. doi: 10.1007/s00414-024-03356-3
2. Pose Lepe G, Villacres F, Silva Fuente-Alba C, et al. Correlación en la determinación de la edad ósea radiológica mediante el método de Greulich y Pyle versus la evaluación automatizada utilizando el software BoneXpert. *Rev. Chil. Pediatr.* 2018;89(5):606-11. doi: 10.4067/S0370-41062018005000705
3. Núñez-Enríquez JC, Arias-Gómez J, Nishimura-Meguro E. Proceso diagnóstico en talla baja. *Rev Med Inst Mex Seguro Soc.* 2012;50(6):623-30.
4. Kowo-Nyakoko F, Gregson CL, Madanhire T, et al. Evaluation of two methods of bone age assessment in peripubertal children in Zimbabwe. *Bone.* 2023;170:116725. doi: 10.1016/j.bone.2023.116725
5. Akin Kağızmanlı G, Deveci Sevim R, Besci Ö, et al. Which method is more effective in predicting adult height in pubertal girls treated with gonadotropin-releasing hormone agonist? *Hormones (Athens).* 2023;22(3):501-6. doi: 10.1007/s42000-023-00466-2
6. Haghnegahdar A, Pakshir HR, Zandieh M, et al. Computer Assisted Bone Age Estimation Using Dimensions of Metacarpal Bones and Metacarpophalangeal Joints Based on Neural Network. *J Dent (Shiraz).* 2024;25(1):51-8. doi: 10.30476/dentjods.2023.95629.1882
7. Lee BD, Lee MS. Automated Bone Age Assessment Using Artificial Intelligence: The Future of Bone Age Assessment. *Korean J Radiol.* 2021;22(5):792-800. doi: 10.3348/kjr.2020.0941
8. Larson DB, Chen MC, Lungren MP, et al. Performance of a Deep-Learning Neural Network Model in Assessing Skeletal Maturity on Pediatric Hand Radiographs. *Radiology.* 2018;287(1):313-22. doi: 10.1148/radiol.2017170236
9. Jiménez Alés R. Artificial Intelligence. Challenges and concerns. *Rev Pediatr Aten Primaria.* 2023;25:205-10.
10. Zhao K, Ma S, Sun Z, et al. Effect of AI-assisted software on inter- and intra-observer variability for the X-ray bone age assessment of preschool children. *BMC Pediatr.* 2022;22(1):644. doi: 10.1186/s12887-022-03727-y
11. Artioli TO, Alvares MA, Carvalho Macedo VS, et al. Bone age determination in eutrophic, overweight and obese Brazilian children and adolescents: a comparison between computerized BoneXpert and Greulich-Pyle methods. *Pediatr Radiol.* 2019;49(9):1185-91. doi: 10.1007/s00247-019-04435-z
12. Nguyen T, Hermann AL, Ventre J, et al. High performance for bone age estimation with an artificial intelligence solution. *Diagn Interv Imaging.* 2023;104(7-8):330-6. doi: 10.1016/j.diii.2023.04.003
13. Garrido-Pino CA, López-Montero LM, López-Lozano L, et al. Edema macular en fotografías de fondo de ojo mediante un algoritmo computacional. *Rev Med Inst Mex Seguro Soc.* 2024;62(2):1-7. doi: 10.5281/zenodo.10711610
14. Rodríguez-Esquivel M, Mendoza-Rodríguez MG, Hernández-Quijano T, et al. La innovación sensorial para detección no invasiva del cáncer de mama. *Rev Med Inst Mex Seguro Soc.* 2020;58(Supl 1):S104-15. doi: 10.24875/RMIMSS.M20000121
15. Hassan N, Slight R, Morgan G, et al. Road map for clinicians to develop and evaluate AI predictive models to inform clinical decision-making. *BMJ Health Care Inform.* 2023;30(1):e100784. doi: 10.1136/bmjhci-2023-100784
16. Chávez-Vázquez AG, Klünder-Klünder M, Garibay-Nieto NG et al. Evaluation of height prediction models: from traditional methods to artificial intelligence. *Pediatr Res.* 2024;95:308-15. doi: 10.1038/s41390-023-02821-w
17. Yang C, Dai W, Qin B, et al. A real-time automated bone age assessment system based on the RUS-CHN method. *Front Endocrinol (Lausanne).* 2023;14:1073219. doi: 10.3389/fendo.2023.1073219

18. Huang S, Su Z, Liu S, et al. Combined assisted bone age assessment and adult height prediction methods in Chinese girls with early puberty: analysis of three artificial intelligence systems. *Pediatr Radiol.* 2023;53(6):1108-16. doi: 10.1007/s00247-022-05569-3
19. Alshamrani K, Hewitt A, Offiah AC. Applicability of two bone age assessment methods to children from Saudi Arabia. *Clin Radiol.* 2020;75(2):156.e1-9. doi: 10.1016/j.crad.2019.08.029
20. Lea WW, Hong SJ, Nam HK, et al. External validation of deep learning-based bone-age software: a preliminary study with real world data. *Sci Rep.* 2022;12(1):1232. doi: 10.1038/s41598-022-05282-z