



UNIVERSIDAD PERUANA
CAYETANO HEREDIA

Facultad de
MEDICINA

USO DE INTELIGENCIA ARTIFICIAL PARA LA ESTIMACIÓN DE EDAD
ÓSEA MEDIANTE RADIOGRAFÍAS DIGITALES EN POBLACIÓN
PEDIÁTRICA EN ESTABLECIMIENTOS DE SALUD: UNA REVISIÓN DE
ALCANCE

USE OF ARTIFICIAL INTELLIGENCE FOR BONE AGE ESTIMATION
USING DIGITAL X-RAYS IN THE PEDIATRIC POPULATION IN HEALTH
FACILITIES: A SCOPE REVIEW

TESIS PARA OPTAR POR EL TÍTULO PROFESIONAL DE LICENCIADO
EN TECNOLOGÍA MÉDICA EN LA ESPECIALIDAD DE RADIOLOGÍA

AUTORES

ANA CRISTINA CHAUCA VALVERDE
SHANNIA SCARLETT MARTINEZ HUARANCCA
DAYANNA MARIYOL REINA ZUMAETA

ASESORA

CECILIA LUIS MIRANDA

CO - ASESOR

ARQUIMEDES MANSUETO GAVINO GUTIERREZ

LIMA - PERÚ

2026

JURADO

Presidente: DR. GUILLERMO BRAVO PUENTE

Vocal: LIC. JOSE FERNANDO MARQUEZ PACHAS

Secretario: LIC. MARCO ANTONIO RIVERO MENDOZA

Fecha de Sustentación: 20 de Abril del 2026

Calificación: Aprobado

ASESORES DE TESIS

ASESORA

MG. CECILIA LUIS MIRANDA

Departamento Académico de Tecnología Médica

ORCID: 0009-0003-4631-2210

CO - ASESOR

DR. ARQUIMEDES MANSUETO GAVINO GUTIERREZ

Unidad de Epidemiología Clínica de la Facultad de Medicina

ORCID: 0000-0002-3325-1004

DEDICATORIA

A mis padres, Franklin y Marybel, quienes con su ejemplo, paciencia y entrega me enseñaron a no rendirme incluso en los momentos más desafiantes. Este trabajo lleva todo el reflejo de lo que han sembrado en mí. A mis hermanos, Rogger y Julieth, por estar siempre presentes, por su apoyo y por recordarme, incluso en los días difíciles, que nunca estoy sola. A Copito, por ser ese pequeño refugio de alegría que hizo más ligeros los días de esfuerzo. Con todo mi cariño, este logro también les pertenece.

-Ana Cristina Chauca Valverde

A mi mamá, por su apoyo incondicional, esfuerzo constante y por brindarme la oportunidad de alcanzar mis metas. A mi familia, por su acompañamiento constante, su comprensión y por ser mi soporte en cada momento de este proceso. A Dios, por darme la vida, la fortaleza y la perseverancia para culminar esta etapa. A ustedes, que fueron mi impulso en cada paso, les dedico este logro.

- Shannia Scarlett Martinez Huarancca

A mis padres, por ser el cimiento de mi vida y el ejemplo más puro de sacrificio y amor; gracias por creer en mí y por trabajar incansablemente para darme las oportunidades que hoy se ven reflejadas. A mi hermano, mi compañero de vida y mi motivación para superarme y con el deseo de ser siempre un ejemplo para él. A mi tía Juani, quien me tomó de la mano en el inicio de este camino; aunque la pandemia nos separó físicamente, su apoyo y aliento en aquellos primeros pasos fueron la semilla de este logro. Siempre estará presente en cada éxito de mi vida.

- Dayanna Mariyol Reina Zumaeta

AGRADECIMIENTO

Expresamos nuestro más profundo agradecimiento a la Universidad Peruana Cayetano Heredia, nuestra casa de estudios, por brindarnos el entorno y las herramientas necesarias para nuestra formación profesional.

De manera especial, manifestamos nuestra gratitud a nuestros asesores de tesis. Su guía experta, rigor científico y constante apoyo fueron los pilares fundamentales para el desarrollo y la culminación con éxito de esta investigación.

Asimismo, agradecemos a los docentes del área de Radiología, quienes con su experiencia y dedicación contribuyeron significativamente a nuestro crecimiento académico y profesional a lo largo de estos años.

FUENTES DE FINANCIAMIENTO

Autofinanciado

DECLARACIÓN DE CONFLICTO DE INTERÉS

Las autoras declaran no tener conflicto de interés

DECLARACIÓN DE ORIGINALIDAD



UNIVERSIDAD PERUANA
CAYETANO HEREDIA

DECLARACIÓN DE ORIGINALIDAD

Los egresados:

N°	APELLIDOS Y NOMBRES
1.	CHAUCA VALVERDE ANA CRISTINA
2.	MARTINEZ HUARANCCA SHANNIA SCARLETT
3.	REINA ZUMAETA DAYANNA MARIYOL

Pertenecientes al programa de la **CARRERA PROFESIONAL DE TECNOLOGÍA MÉDICA EN LA ESPECIALIDAD DE RADIOLOGÍA**, autores del trabajo titulado: **USO DE INTELIGENCIA ARTIFICIAL PARA LA ESTIMACIÓN DE EDAD ÓSEA MEDIANTE RADIOGRAFÍAS DIGITALES EN POBLACIÓN PEDIÁTRICA EN ESTABLECIMIENTOS DE SALUD: UNA REVISIÓN DE ALCANCE** el cual ha sido elaborado, sustentado y aprobado, según corresponda, para optar por el **TÍTULO PROFESIONAL DE LICENCIADO EN TECNOLOGÍA MÉDICA EN LA ESPECIALIDAD DE RADIOLOGÍA** bajo la modalidad de **TESIS**.

En calidad de docentes asesores de la Universidad Peruana Cayetano Heredia:

N°	APELLIDOS Y NOMBRES DEL DOCENTE	FACULTAD	NIVEL DE ASESORÍA
1.	LUIS MIRANDA CECILIA	MEDICINA	ASESORA
2.	GAVINO GUTIERREZ ARQUIMEDES MANSUETO	MEDICINA	CO-ASESOR

Declaramos que el contenido del presente documento es original y que las citas y referencias a otros autores cumplen con las normas académicas establecidas. En ese sentido, hacemos constar que:

- El documento presenta un porcentaje de similitud de **13%**, según el reporte emitido por el software **Turnitin®** (identificador de entrega: **trn:oid:::1:3558824252**); fecha de entrega: **04-05-2026**
- Tras una revisión detallada del reporte y del contenido del trabajo en cuestión, no se han identificado indicios de plagio.
- Se certifica que el documento respeta los principios de integridad académica y cumple con los requisitos institucionales de originalidad.

Lugar y fecha: **Lima, 04 de mayo de 2026**

Firma del asesor
N° DNI: 07168077
ORCID: 0009-0003-4631-2210

Firma del Co-asesor
N° DNI: 40048216
ORCID: 0000-0002-3325-1004



TABLA DE CONTENIDO

RESUMEN

ABSTRACT

I.	INTRODUCCIÓN	1
II.	OBJETIVOS:	6
III.	MATERIALES Y MÉTODOS	7
IV.	RESULTADOS	17
V.	DISCUSIÓN	22
VI.	LIMITACIONES	25
VII.	CONCLUSIONES	26
VIII.	REFERENCIAS BIBLIOGRÁFICAS	28
IX.	TABLAS	33
	ANEXOS	

RESUMEN

Introducción: La estimación de la edad ósea es fundamental para evaluar el desarrollo infantil, pero los métodos tradicionales presentan limitaciones que pueden afectar su precisión. Esto motiva el uso de la inteligencia artificial, cuyos modelos recientes muestran resultados esperanzadores. **Objetivo:** Mapear la evidencia disponible del uso de la inteligencia artificial para la estimación de la edad ósea en la población pediátrica mediante radiografías digitales en establecimientos de salud. **Metodología:** Se realizó una revisión de alcance siguiendo PRISMA-ScR. La búsqueda incluyó Embase, PubMed, IEEE Xplore, LILACS, Web of Science, Cochrane, SciELO y como motor de búsqueda complementario Google Scholar, considerando estudios publicados entre enero de 2013 y octubre de 2025. Tras aplicar los criterios de elegibilidad, los resultados se analizaron de forma descriptiva y por frecuencias mediante Microsoft Excel 2021, reportados en gráficos y tablas. **Resultados:** Se analizaron nueve estudios entre 2020 y 2025, de diseño retrospectivo, que incluyeron más de 3,500 radiografías digitales de mano izquierda en población pediátrica de 1 a 18 años. Los modelos de IA mostraron alta precisión: los errores absolutos medios entre 0.3 y 0.7 años, mientras que las correlaciones intraclases reportadas fueron de 0.99, alcanzando concordancias casi perfectas con los métodos tradicionales Greulich & Pyle y Tanner–Whitehouse 3, se redujo 87% el tiempo de lectura y la variabilidad inter e intraobservador. **Conclusiones:** La IA en radiografías pediátricas logra precisión comparable a especialistas reduciendo el tiempo y variabilidad, dependiendo de la calidad de imagen y puede presentar sesgos en poblaciones distintas a las de entrenamiento.

Palabras clave: Inteligencia artificial, edad ósea, maduración esquelética, radiografía, pediatría.

ABSTRACT

Introduction: Bone age estimation is fundamental for assessing child development, but traditional methods have limitations that can affect their accuracy. This motivates the use of artificial intelligence, whose recent models show promising results. **Objective:** To map the available evidence on the use of artificial intelligence for bone age estimation in the pediatric population using digital radiographs in healthcare facilities. **Methodology:** A scoping review was conducted following the PRISMA-ScR guidelines. The search included Embase, PubMed, IEEE Xplore, LILACS, Web of Science, Cochrane, SciELO, and, as a complementary search engine, Google Scholar, considering studies published between January 2013 and October 2025. After applying the eligibility criteria, the results were analyzed descriptively and by frequency using Microsoft Excel 2021 and reported in graphs and tables. **Results:** Nine retrospective studies conducted between 2020 and 2025, including more than 3,500 digital radiographs of the left hand in a pediatric population aged 1 to 18 years, were analyzed. AI models showed high accuracy: mean absolute errors ranged from 0.3 to 0.7 years, while reported intraclass correlations were 0.99, achieving near-perfect agreement with the traditional Greulich & Pyle and Tanner–Whitehouse methods. Reading time and inter- and intra-observer variability were reduced by 87%. **Conclusions:** AI in pediatric radiography achieves accuracy comparable to specialists, reducing time and variability, depending on image quality. However, it may exhibit bias in populations other than the training population.

Keywords: Artificial intelligence, bone age, skeletal maturation, radiography, pediatrics.

I. INTRODUCCIÓN

El crecimiento óseo es uno de los indicadores más útiles para comprender cómo avanza el desarrollo de niños y adolescentes. Aunque la genética marca una base importante, este proceso se ve influido por la alimentación, las hormonas y el entorno en el que los menores crecen (1,2). Cuando alguno de estos factores se altera, pueden aparecer dificultades en el crecimiento, por lo que evaluar este proceso con atención resulta esencial para detectar problemas a tiempo y orientar un tratamiento adecuado.

En este marco, la estimación de la edad ósea es una herramienta clave para conocer el nivel de maduración del esqueleto infantil. Generalmente se obtiene mediante radiografías de la mano y la muñeca, regiones que contienen varios centros de osificación que permiten evaluar su progreso. Esta evaluación permite comparar la edad biológica con la cronológica, e identificar posibles alteraciones en el crecimiento y respaldar decisiones clínicas en especialidades como pediatría, endocrinología y ortopedia (1,3).

El tecnólogo médico en radiología cumple un rol clave en la determinación de la edad ósea, siendo responsable de la preparación del paciente, la adquisición de imágenes radiológicas y el control de calidad para garantizar su valor diagnóstico. Asimismo, participa en el procesamiento y registro de las imágenes, y brinda apoyo técnico en la evaluación mediante métodos como Greulich & Pyle y Tanner-Whitehouse. Aunque no emite diagnóstico, su labor es esencial dentro del trabajo interdisciplinario.

El atlas de Greulich & Pyle es uno de los métodos tradicionales para estimar la edad ósea, basado en la comparación de una radiografía de la mano con imágenes de referencia que representan distintas etapas de maduración según la edad y el sexo (4,5). Por su parte, el método Tanner-Whitehouse propone un enfoque más detallado: evalúa cada hueso de forma individual, asignando una puntuación y obteniendo una estimación final a partir de su suma. Ambos métodos han sido ampliamente utilizados en la práctica clínica.

Sin embargo, pese a su uso extendido, estos métodos presentan diversas limitaciones. Su precisión depende de la experiencia del evaluador, lo que puede generar variabilidad en los resultados. Además, requieren un tiempo de análisis amplio y se basan en atlas desarrollados a partir de poblaciones que no siempre representan la diversidad actual. Estas limitaciones han impulsado el desarrollo de alternativas más objetivas y estandarizadas, como los métodos basados en inteligencia artificial (6). Los cuales permiten automatizar el análisis, reducir la variabilidad y mejorar la reproducibilidad en la estimación de la edad ósea.

La inteligencia artificial (IA) ha emergido como una alternativa para superar limitaciones de los métodos tradicionales en la estimación de la edad ósea. De manera amplia, la IA abarca diversas aproximaciones: los sistemas simbólicos o basados en reglas, que operan siguiendo algoritmos definidos por especialistas, capaces de reconocer patrones a partir de grandes volúmenes de datos; y las técnicas de aprendizaje profundo, que emplean redes neuronales con múltiples capas. Dentro de estas últimas, las redes neuronales convolucionales (CNN) han demostrado un alto desempeño en el análisis de radiografías digitales debido a su capacidad para

identificar características complejas con gran precisión (6,7). Estas herramientas permiten optimizar el análisis de imágenes médicas y mejorar la consistencia en los resultados.

En este contexto, los modelos de IA intervienen principalmente en las fases de procesamiento e interpretación de la imagen médica. Su participación en la adquisición es limitada y se orienta a la optimización de parámetros técnicos y al control de calidad automatizado. En la etapa de procesamiento, estos sistemas permiten la segmentación de estructuras óseas, la detección de centros de osificación y la extracción de características relevantes. Posteriormente, en la fase de interpretación, facilitan la estimación automatizada de la edad ósea.

Es importante considerar que los rangos de edad utilizados por estos modelos de IA no son universales, sino que dependen del software y de las bases de datos empleadas durante su entrenamiento definidas por las características específicas de la población analizada.

Las redes neuronales convolucionales (CNN) han demostrado alta eficiencia en el análisis de radiografías para la estimación de la edad ósea. Un ejemplo es el desafío organizado por la Radiological Society of North America (RSNA), donde distintos modelos de IA alcanzaron correlaciones superiores a 0.99 con la edad cronológica y un error absoluto medio menor a 4.5 meses. Estos resultados evidencian que la IA puede lograr niveles de precisión similares o incluso superiores a los de especialistas con amplia experiencia. Sin embargo, es fundamental resaltar que estas tecnologías no reemplazan la labor profesional, sino que actúan como un soporte que refuerza y facilita las decisiones clínicas. (8)

El uso de la inteligencia artificial en radiología pediátrica empezó a ganar protagonismo alrededor de 2013, cuando aparecieron los primeros algoritmos de aprendizaje profundo y comenzaron a integrarse herramientas automatizadas en la práctica clínica (8,9). Entre ellas sobresalió BoneXpert, un software que actualmente opera en más de 150 centros europeos y que marcó un cambio importante al pasar de métodos tradicionales a sistemas digitales más consistentes. Desde entonces, las investigaciones se han multiplicado, evaluando no solo cómo se compara su precisión con técnicas clásicas como Greulich & Pyle, (10,11) sino también cómo responde en distintos entornos clínicos y en diversas poblaciones pediátricas.

BoneXpert analiza radiografías de la mano izquierda y combina el método de Greulich & Pyle con algoritmos de inteligencia artificial capaces de entregar resultados en menos de 15 segundos. Su desempeño ha sido probado en múltiples grupos poblacionales, obteniendo un error cuadrático medio (RMS) de 0.62 años, lo que refleja un alto nivel de exactitud. Además, cuenta con un sistema automático que descarta imágenes digitales de mala calidad o inadecuadas para el análisis, lo que mejora la fiabilidad del procedimiento y aporta mayor seguridad en la práctica clínica. (12,13)

El uso de equipos digitales de radiología es fundamental para la aplicación de la IA en estos estudios, permitiendo una evaluación rápida y precisa, optimizando tiempos y mejorando la exactitud del diagnóstico. Estos equipos están diseñados para asegurar la calidad de las imágenes, descartando automáticamente aquellas que no sean adecuadas para el análisis, lo que aumenta la confiabilidad del proceso.

A pesar de los avances recientes, la información sobre el uso de la inteligencia artificial para estimar la edad ósea en población pediátrica aún se encuentra dispersa. Si bien existen estudios y revisiones que abordan modelos o tecnologías específicas, no se dispone de una revisión de alcance que sintetice la evidencia disponible. Esta limitación dificulta contar con una visión clara sobre los modelos utilizados, los contextos clínicos en los que se aplican, las poblaciones en las que han sido validados y las principales limitaciones metodológicas.

En este contexto, una revisión de alcance permitiría orientar futuras investigaciones, favoreciendo una adopción más segura y efectiva de estas herramientas en la práctica pediátrica. En este sentido, se plantea la siguiente pregunta de investigación: ¿Cuál es la evidencia disponible sobre el uso de inteligencia artificial para la estimación de la edad ósea mediante radiografías digitales para la población pediátrica en establecimientos de salud?

II. OBJETIVOS:

II.1. Objetivo General:

- Mapear la evidencia disponible del uso de la inteligencia artificial para la estimación de edad ósea mediante radiografías digitales en población pediátrica en establecimientos de salud.

II.2. Objetivos Específicos

- Identificar los modelos y técnicas de la inteligencia artificial utilizados para estimar la edad ósea mediante radiografías digitales en población pediátrica de los establecimientos de salud.
- Describir los diseños de estudios utilizados en la literatura sobre inteligencia artificial para la estimación de la edad ósea en población pediátrica.
- Describir el impacto y la utilidad clínica de la inteligencia artificial en la estimación de la edad ósea mediante radiografías digitales en población pediátrica.

III. MATERIALES Y MÉTODOS

3.1 Diseño del estudio

El presente estudio corresponde a una revisión de alcance (scoping review), desarrollada siguiendo las directrices del Joanna Briggs Institute (JBI), con el objetivo de organizar, mapear y sintetizar la evidencia disponible sobre el uso de inteligencia artificial en la estimación de la edad ósea mediante radiografías digitales en población pediátrica atendida en establecimientos de salud. Su elaboración y presentación se ajustaron a la guía PRISMA-ScR. Asimismo, el protocolo fue registrado en el Sistema Descentralizado de Información y Seguimiento a la Investigación (SIDISI) de la Dirección Universitaria de Investigación, Ciencia y Tecnología (DUICT) para su aprobación.

Este diseño fue seleccionado porque permite describir la evidencia existente e identificar brechas de conocimiento en torno al uso de inteligencia artificial para la estimación de la edad ósea en niños y adolescentes de 0 a 18 años. De acuerdo con las recomendaciones del JBI, no se realizó evaluación del riesgo de sesgo, ya que este tipo de revisión se centra en mapear la evidencia disponible más que en valorar su calidad.

Se diseñó una estrategia de búsqueda sistemática en diversas bases de datos electrónicas, entre ellas PubMed, Embase, IEEE Xplore, LILACS, SciELO, Web of Science y Cochrane, complementada con Google Scholar. Se emplearon descriptores controlados y palabras clave en inglés y español, combinados mediante operadores booleanos: (“bone age assessment” OR “bone age estimation”) AND

(“artificial intelligence” OR “deep learning” OR “machine learning”) AND (“pediatric” OR “children”) AND (“radiograph” OR “hand radiography”). La búsqueda se restringió a artículos publicados entre 2013 y 2025, revisados por pares y disponibles en texto completo. El proceso se realizó entre el 9 y el 25 de octubre de 2025.

La recopilación de datos fue realizada de manera independiente por las tres revisoras (AC, SM, DR), considerando variables como año de publicación, tipo de estudio, población, área anatómica, modalidad diagnóstica y tipo de inteligencia artificial. Los resultados fueron exportados a Zotero, donde se eliminaron los duplicados.

El proceso de selección se desarrolló en tres etapas: revisión de títulos, evaluación de resúmenes y análisis de textos completos. Este procedimiento permitió identificar nueve estudios que cumplieron con los criterios establecidos, los cuales fueron incluidos para la extracción de datos. Todo el proceso fue representado mediante un diagrama de flujo PRISMA.

Se incluyeron artículos en español e inglés publicados entre el 1 de enero de 2013 y el 8 de octubre de 2025, considerando el crecimiento de la inteligencia artificial en el ámbito del diagnóstico por imágenes durante este periodo. La selección de estudios se realizó según el marco PCC:

- **Población:** niños y adolescentes de 0 a 18 años, sin distinción de sexo o condición clínica.

- **Concepto:** uso de inteligencia artificial aplicada a la estimación de la edad ósea mediante radiografías.
- **Contexto:** establecimientos de salud con servicio de radiología y equipos de rayos X digitales.

3.2 Criterios de inclusión

- Estudios que involucran población pediátrica entre 0 a 18 años, sin distinción de sexo.
- Estudios que utilizan el uso de inteligencia artificial, incluyendo *machine learning*, *deep learning*, redes neuronales y visión computacional, para la estimación, evaluación o determinación de la edad ósea mediante radiografías digitales.
- Estudios realizados en establecimientos de salud que cuentan con servicio de radiología con equipos digitales.
- Estudios primarios que incluyen diseños observacionales (casos, cohortes y controles, estudios analíticos transversales), además informes técnicos, reportes y series de casos que presentan resultados sobre la estimación de la edad ósea mediante inteligencia artificial.
- Literatura gris que incluye informes técnicos, tesis, documentos emitidos por agencias de salud o universidades y disertaciones, siempre que cumplan con los criterios de inclusión.
- Estudios secundarios los que incluyeron, revisiones narrativas y mapas de evidencia cuyo propósito es exclusivamente identificar investigaciones

primarias relevantes y reconocer los vacíos de conocimiento existentes en el área.

- Artículos publicados entre enero de 2013 y octubre de 2025, en línea con el desarrollo acelerado que tiene la inteligencia artificial aplicada a la salud durante este período.
- Artículos publicados en español, inglés y portugués.

3.3 Criterios de exclusión

- Artículos sobre inteligencia artificial aplicada a la estimación de la edad ósea que no presentan resultados claros o cuyos datos fueran insuficientes para realizar un análisis adecuado.
- Artículos cuya metodología no permitiera determinar si cumplían adecuadamente con los criterios de inclusión establecidos.
- Artículos realizados con equipos radiográficos analógicos, ya que su calidad y formato no son comparables con los estándares digitales necesarios para el uso de inteligencia artificial.

3.4 Definición de variables

Las variables corresponden a los elementos centrales que permiten analizar cómo se utiliza la inteligencia artificial para estimar la edad ósea en población pediátrica. Cada una fue definida considerando su dimensión conceptual, su forma de medición y operacionalización, los indicadores utilizados y el tipo de variable al que pertenece. (Anexo 1)

3.5 Fuente de información

La búsqueda de información se llevó a cabo considerando publicaciones entre enero de 2013 y octubre de 2025, de acuerdo con los criterios de inclusión establecidos. Para ello, se consultaron diversas bases de datos biomédicas y regionales, incluyendo MEDLINE (vía PubMed), Embase (vía Ovid), Scopus, Web of Science, LILACS, SciELO, IEEE Xplore y Cochrane. Asimismo, se limitaron los resultados a estudios publicados en inglés, español y portugués, utilizando estos idiomas como filtros en cada plataforma.

Como complemento, se empleó Google Scholar para identificar literatura gris y recuperar estudios que no estuvieran indexados en las bases de datos principales. De igual manera, se revisaron repositorios de tesis, documentos emitidos por organismos de salud y otros recursos institucionales relevantes.

Adicionalmente, se realizó una revisión manual de las listas de referencias de los artículos incluidos (backward citation tracking), lo que permitió identificar estudios adicionales que no habían sido detectados durante la búsqueda electrónica.

Finalmente, todos los artículos seleccionados fueron revisados de manera manual por las autoras (AC, SM y DR), con el fin de verificar que cumplieran rigurosamente con los criterios definidos para esta revisión de alcance.

3.6 Estrategia de búsqueda

El presente estudio fue registrado el 06 de septiembre de 2025 en el Sistema Descentralizado de Información y Seguimiento a la Investigación (SIDISI) de la Dirección Universitaria de Investigación, Ciencia y Tecnología (DUICT), siendo aprobado el 03 de octubre del mismo año. Posteriormente, fue evaluado por el Comité de Ética de la Universidad Peruana Cayetano Heredia (CIE-UPCH), obteniendo su aprobación el 08 de octubre de 2025.

La estrategia de búsqueda se diseñó siguiendo las recomendaciones del Joanna Briggs Institute (JBI) y el marco PCC, el cual considera la población (niños y adolescentes de 0 a 18 años), el concepto (uso de inteligencia artificial en la estimación de la edad ósea) y el contexto (radiografías digitales en establecimientos de salud). Este proceso se desarrolló con el apoyo de un bibliotecario especializado, lo que permitió optimizar la calidad de la estrategia.

Se utilizaron términos controlados como DeCS, MeSH y Emtree, combinados con palabras clave relacionadas con inteligencia artificial, edad ósea, radiografía y población pediátrica, incluyendo sus respectivos sinónimos. Estos términos se integraron mediante operadores booleanos (AND, OR, NOT) y se adaptaron según las características de cada base de datos (Anexo 2).

Se aplicaron filtros por idioma (inglés, español y portugués) y por periodo de publicación (enero de 2013 a octubre de 2025). La búsqueda se realizó entre el 09 de octubre y el 09 de noviembre de 2025. Asimismo, se complementó con una revisión manual de referencias de los estudios incluidos y de revisiones secundarias relevantes, con el objetivo de identificar evidencia adicional.

Todo el proceso fue documentado siguiendo las recomendaciones de PRISMA-ScR, garantizando así la transparencia y la reproducibilidad de la búsqueda.

3.7 Selección de estudios

La selección de los estudios se desarrolló de manera sistemática utilizando el gestor bibliográfico Zotero, el cual facilitó la organización de las referencias y el seguimiento del proceso de revisión. En una primera etapa, todos los registros fueron importados al software y se eliminaron los duplicados mediante herramientas automáticas, complementadas con una verificación manual.

Posteriormente, las revisoras AC, SM y DR evaluaron de forma independiente los títulos, resúmenes y textos completos, aplicando los criterios de inclusión y exclusión previamente definidos bajo el marco PCC.

Se incluyeron únicamente estudios primarios que abordaran población pediátrica, emplearan inteligencia artificial y utilizaran radiografías digitales obtenidas en establecimientos de salud. Por otro lado, los estudios secundarios, como revisiones sistemáticas, revisiones de alcance y metaanálisis, no fueron considerados para la extracción de datos; sin embargo, sus referencias fueron revisadas para identificar posibles estudios primarios adicionales.

En caso de discrepancias entre las revisoras, estas se resolvieron mediante discusión conjunta y, cuando fue necesario, con la participación de una cuarta revisora (CL).

El proceso de selección se realizó en tres etapas: revisión de títulos, evaluación de resúmenes y análisis del texto completo. Los motivos de exclusión fueron registrados de manera estandarizada, siguiendo las recomendaciones de PRISMA-ScR.

Finalmente, los resultados del proceso se representaron mediante un diagrama de flujo PRISMA-ScR, en el que se detallan los estudios identificados, evaluados, incluidos y excluidos (Anexo 3).

3.8 Extracción y análisis de datos

3.8.1 Extracción de datos

La extracción de datos se realizó a partir de los estudios primarios incluidos, empleando una herramienta diseñada específicamente para esta revisión y basada en las directrices del Joanna Briggs Institute (JBI).

Dicha herramienta fue elaborada en Microsoft Excel (versión 2021) y sometida a una prueba piloto, con el fin de asegurar su claridad, consistencia y alineación con los objetivos del estudio.

Se recolectó información relacionada con los autores, año de publicación, país de origen, diseño metodológico, características de la población (edad, sexo y tamaño de muestra), tipo de inteligencia artificial utilizada, método de estimación de edad

ósea, tipo de radiografía digital, resultados principales, métricas de precisión y limitaciones reportadas por los autores.

El proceso fue llevado a cabo de manera independiente por las revisoras AC, SM y DR. Las discrepancias identificadas durante la extracción fueron resueltas mediante consenso y, en caso necesario, con la intervención de una cuarta revisora (CL).

Posteriormente, la información recopilada fue organizada en tablas y mapas de evidencia, lo que facilitó su análisis e interpretación.

3.8.2 Análisis de datos

Los datos extraídos fueron analizados mediante una síntesis narrativa complementada con representaciones visuales, con el objetivo de organizar y describir la evidencia disponible en relación con los objetivos de la revisión. Este proceso se desarrolló siguiendo las directrices del Joanna Briggs Institute y PRISMA-ScR.

Los estudios se agruparon de acuerdo con el marco PCC, considerando las características de la población, el tipo de intervención (inteligencia artificial) y el contexto de aplicación. Esta organización permitió identificar patrones, tendencias y vacíos en la literatura existente.

Asimismo, los resultados se presentaron a través de tablas y mapas de evidencia elaborados en Microsoft Excel, incorporando variables como país, año de publicación, diseño del estudio, tamaño de muestra, grupo etario, tipo de imagen utilizada, tipo de inteligencia artificial aplicada y métricas de precisión reportadas.

Las diferencias en la interpretación de los datos fueron discutidas entre las revisoras (AC, SM y DR) y, cuando fue necesario, con el apoyo de una revisora adicional (CL). Finalmente, los hallazgos fueron presentados en el cuerpo principal del estudio y complementados mediante anexos, siguiendo las recomendaciones de PRISMA-ScR.

IV. RESULTADOS

La tabla 1 muestra las características generales de los estudios.

De los nueve artículos estudiados, el 77.78% pertenece al continente asiático y 22.22% a Europa, siendo tres artículos que provienen de China y dos artículos de Alemania, mientras que Taiwán, Corea del Sur, Turquía e Irán están representados con un estudio cada uno. Respecto al año de publicación, la mayoría corresponden a 2020 y 2022 (dos estudios por año), seguidos de 2021, 2023, 2024 y 2025 (un estudio por año cada uno).

Respecto al tamaño muestral, los estudios presentaron variabilidad, con muestras que oscilaron entre menos de 200 y más de 1,000 radiografías, acumulando un total aproximado superior a 3,500 imágenes. El 66.67% (n=6) de los estudios incluyó muestras mayores a 300 pacientes, lo que evidencia un adecuado volumen de datos para la validación de modelos.

En relación con la población, el 100% (n=9) de los estudios evaluó población pediátrica, con rangos de edad entre 1 y 18 años, los cuales corresponden a las muestras analizadas y no a limitaciones propias de los modelos de inteligencia artificial.

La tabla 2 muestra los modelos y técnicas de inteligencia artificial utilizados.

En cuanto a los modelos de inteligencia artificial, el 50% (n=4) de los estudios utilizó el software BoneXpert como herramienta principal para la estimación automatizada de la edad ósea, mientras que el 22.22% (n=2) empleó VUNO Med-

BoneAge y el 11.11% (n=1) utilizó Lab Panda. Asimismo, el 16.67% (n=2) de los estudios desarrolló modelos propios basados en redes neuronales convolucionales (CNN), evidenciando el predominio de técnicas de aprendizaje profundo en este campo.

Booz (18) describen el uso de BoneXpert como un sistema completamente automatizado que emplea algoritmos avanzados para la segmentación de estructuras óseas y la evaluación sistemática de la maduración esquelética a partir de radiografías digitales.

De manera similar, **Gao (19)** destacan el uso de modelos de inteligencia artificial basados en aprendizaje profundo, diseñados para reconocer patrones morfológicos en imágenes de mano y muñeca, mediante el análisis detallado de estructuras como carpos, metacarpos y falanges, lo que permite una evaluación automatizada de la edad ósea.

Por otro lado, **Pape (20)** señala que los sistemas de inteligencia artificial utilizan imágenes digitales de alta resolución para alimentar algoritmos que identifican características anatómicas específicas, lo que facilita el procesamiento uniforme de las imágenes y la detección de patrones complejos de desarrollo óseo.

Asimismo, otros estudios incluidos emplearon redes neuronales convolucionales (CNN) orientadas a tareas como la segmentación de regiones de interés y la extracción automática de características óseas, lo que mejora la capacidad de los modelos para analizar grandes volúmenes de imágenes médicas de manera eficiente.

Finalmente, el 33.33% (n=3) de los estudios combinó modelos de inteligencia artificial con métodos tradicionales como Tanner-Whitehouse (TW3) o RUS-CHN, utilizándolos como referencia para el entrenamiento y validación de los algoritmos. En conjunto, el 100% (n=9) de los estudios empleó técnicas basadas en inteligencia artificial, predominando el uso de aprendizaje profundo, lo que evidencia una tendencia hacia la automatización del análisis de imágenes radiográficas en la estimación de la edad ósea.

La tabla 3 muestra los diseños metodológicos de los estudios.

El 100% (n=9) de los estudios correspondió a investigaciones observacionales de tipo retrospectivo. Dentro de estos, el 88.89% (n=8) incluyó un enfoque de validación de modelos, comparando directamente los resultados de la inteligencia artificial con la evaluación realizada por especialistas.

Asimismo, el 66.67% (n=6) de los estudios utilizó métricas de concordancia como la correlación intraclass (ICC), mientras que el 77.78% (n=7) reportó el error medio absoluto (MAE) como indicador principal de precisión.

La tabla 4 muestra el impacto y utilidad clínica de la inteligencia artificial.

Los estudios reportaron una alta precisión de los modelos de inteligencia artificial en la estimación de la edad ósea. Los valores de error medio absoluto (MAE) oscilaron entre 0.3 y 0.7 años en el 77.78% (n=7) de los estudios, mientras que el 22.22% (n=2) reportó valores ligeramente superiores en poblaciones específicas.

Asimismo, el 88.89% (n=8) de los estudios evidenció una alta concordancia con los métodos tradicionales, con valores de correlación intraclase (ICC) cercanos a 0.99, lo que confirma la fiabilidad de estos sistemas en la práctica clínica.

Özmen (12) compararon los softwares BoneXpert v3.0.3 y VUNO Med-BoneAge v1.0.3 en una población pediátrica turca de 1 a 15 años, evidenciando una equivalencia técnica entre ambos modelos, con coeficientes de correlación intraclase cercanos a 0.995 y sin diferencias significativas entre ellos, lo que demuestra su estabilidad y confiabilidad en distintos grupos etarios y en ambos sexos.

Por su parte, **Wu-In Le (13)** analizaron el desempeño del software BoneAge en un entorno clínico real en Corea, evidenciando que la inteligencia artificial puede igualar e incluso superar la precisión de especialistas con diferentes niveles de experiencia. Reportaron valores de error cuadrático medio (RMSE) de 0.52 años en varones y 0.50 años en mujeres, así como errores medios absolutos (MAE) de 0.38 y 0.36 años respectivamente, con una alta correlación ($r = 0.983$) entre la IA y los evaluadores humanos.

Adicionalmente, **Fengdan Wang (14)** validaron un modelo basado en arquitectura ResNet (YITU Healthcare) en población Han, aplicándolo posteriormente a niños tibetanos de 4 a 18 años, demostrando una precisión aproximada del 86% dentro de un margen de ± 1 año. No obstante, se evidenció una menor precisión en niños de 4 a 6 años, atribuida a la asincronía en el desarrollo óseo en estas edades.

Por lo tanto, en términos de eficiencia, el 66.67% (n=6) de los estudios reportó una reducción significativa del tiempo de evaluación en comparación con el análisis manual, contribuyendo a optimizar el flujo de trabajo en los servicios de radiología.

Por otro lado, el 44.44% (n=4) de los estudios identificó variaciones en el rendimiento de los modelos al aplicarse en poblaciones distintas a aquellas utilizadas durante su entrenamiento, lo que resalta la necesidad de realizar validaciones locales para garantizar su aplicabilidad.

Finalmente, el 55.56% (n=5) de los estudios destacó que la calidad y estandarización de las imágenes radiográficas influyen directamente en el desempeño de la inteligencia artificial, siendo factores determinantes para obtener resultados confiables en la estimación de la edad ósea.

V. DISCUSIÓN

Los estudios incluidos en esta revisión permiten observar que la inteligencia artificial aplicada a la estimación de la edad ósea ha alcanzado un desempeño sólido y muy cercano al de los métodos tradicionales basados en los atlas de Greulich y Pyle (GP) y el sistema Tanner–Whitehouse (TW3). En la mayoría de los trabajos analizados, los modelos evaluados lograron niveles de precisión considerados adecuados para la práctica clínica, con errores absolutos medios dentro de rangos aceptables y con valores de correlación intraclase que evidencian una concordancia casi perfecta con las evaluaciones realizadas por especialistas. Esto muestra que la IA es capaz de reproducir el análisis manual con un alto grado de fiabilidad, especialmente cuando se trabaja con radiografías digitales obtenidas bajo condiciones técnicas estandarizadas.

Uno de los hallazgos que se repite en varios estudios es la reducción marcada de la variabilidad interobservador e intraobservador. Este aspecto es especialmente relevante si se considera que la interpretación manual puede variar entre evaluadores, incluso cuando se emplea el mismo método de referencia. En trabajos como el de Gao (19), se observó que los algoritmos basados en TW3 y RUS-CHN mejoraron notablemente la reproducibilidad de las mediciones, lo que contribuye a disminuir la subjetividad que caracteriza al análisis convencional de la edad ósea.

Otro elemento que destaca en la literatura es la eficiencia operativa que brindan los sistemas automatizados. En el estudio de Booz (18), por ejemplo, el uso de BoneXpert permitió reducir de manera considerable el tiempo requerido para analizar una radiografía, sin comprometer la precisión del diagnóstico. Este tipo de

ventajas resulta especialmente útil en servicios de radiología con alta carga de trabajo, donde la rapidez y la uniformidad en la interpretación pueden influir directamente en la calidad de atención.

En cuanto al flujo de trabajo en imagen médica, la IA interviene principalmente en las fases de procesamiento e interpretación diagnóstica. Durante el procesamiento, los algoritmos identifican la región de interés y analizan las estructuras óseas, mientras que en la interpretación estiman automáticamente la edad ósea a partir de patrones aprendidos. Su participación en la adquisición es limitada, ya que esta sigue dependiendo del tecnólogo médico, lo que evidencia que la IA actúa como un complemento y no como un reemplazo del proceso diagnóstico.

A pesar de estos beneficios, los estudios también muestran que el rendimiento de los modelos de IA no es uniforme en todas las poblaciones. En investigaciones realizadas en Turquía e Irán, se evidenció que algunos algoritmos entrenados con bases de datos occidentales presentaban diferencias en la estimación de la edad ósea cuando se aplicaban en grupos con características étnicas distintas. Estas variaciones subrayan la importancia de contar con bases de datos representativas y de validar los modelos en cada población antes de incorporarlos al uso clínico, ya que las diferencias en la maduración esquelética pueden influir de manera directa en la exactitud del sistema.

Los estudios revisados también coinciden en la importancia de las imágenes digitales como base para el funcionamiento adecuado de la IA. Casi todos los trabajos utilizaron radiografías directas (DR) en formato DICOM, lo que permitió un procesamiento uniforme y una menor pérdida de calidad. Incluso en

investigaciones que integraron imágenes con variaciones mínimas en la técnica, como el estudio de Pape (20), los modelos mantuvieron un desempeño aceptable, lo que evidencia cierta robustez ante condiciones reales de trabajo en radiología.

En términos generales, los resultados sugieren que la inteligencia artificial no pretende sustituir la evaluación del especialista, sino complementarla. Su mayor contribución radica en mejorar la precisión del análisis, reducir los tiempos de lectura y ofrecer un método más estandarizado para interpretar la edad ósea. En algunos casos, como el estudio de Huang (16), los modelos incluso se integraron con herramientas clínicas adicionales, como la predicción de la talla adulta, lo que demuestra el potencial de la IA para aportar información útil en escenarios específicos.

En conjunto, los hallazgos de esta revisión muestran que los sistemas de inteligencia artificial representan una alternativa válida y prometedora para la estimación de la edad ósea en población pediátrica. No obstante, para garantizar su adecuada aplicación, es necesario mantener procesos de validación continua, especialmente en poblaciones con patrones de maduración distintos a los empleados para entrenar los modelos. Su incorporación a la práctica clínica debe realizarse de forma gradual, asegurando siempre la supervisión del profesional especializado para interpretar adecuadamente los resultados automatizados.

VI. LIMITACIONES

Se identificó la heterogeneidad metodológica de los estudios revisados. Los artículos emplearon modelos distintos de inteligencia artificial, diferentes métodos de referencia (GP, TW3 o BoneXpert) y métricas no uniformes para evaluar precisión, como MAE, RMSE o ICC. Esta diversidad impidió realizar comparaciones directas y dificultó la síntesis de los resultados en una misma escala.

Otra limitación particularmente relevante fue la falta de investigaciones sobre el uso de la inteligencia artificial para la estimación de edad ósea en población pediátrica latinoamericana, lo que representa un vacío científico y dificulta valorar la aplicabilidad clínica de los hallazgos en nuestro contexto. La ausencia de datos locales impide determinar si los algoritmos mantienen su precisión frente a la diversidad étnica, las desigualdades en el acceso a servicios de salud y los diferentes estándares radiológicos. Esta carencia limita la posibilidad de proponer recomendaciones clínicas inmediatas basadas en la evidencia disponible.

Finalmente, la calidad de las radiografías utilizadas no fue homogénea. Las diferencias en los equipos digitales, su resolución y los protocolos de adquisición pueden influir en el rendimiento de los modelos de IA. La falta de estandarización en estos aspectos afectó la comparabilidad entre estudios y limitó la solidez de los resultados.

VII.CONCLUSIONES

La evidencia disponible demuestra que la inteligencia artificial aplicada a la estimación de la edad ósea en población pediátrica presenta un alto nivel de precisión y concordancia con los métodos tradicionales, como el atlas de Greulich y Pyle y Tanner-Whitehouse. Esto confirma su potencial como herramienta de apoyo en la práctica clínica, especialmente en el análisis de radiografías digitales de mano y muñeca, donde su intervención se centra en las fases de procesamiento e interpretación diagnóstica, optimizando el análisis de las imágenes sin reemplazar la adquisición, que depende del tecnólogo médico.

Se identificó que la mayoría de los sistemas utilizados en la literatura se basan en técnicas de aprendizaje automático y aprendizaje profundo, especialmente redes neuronales convolucionales, las cuales han demostrado una alta capacidad para reconocer patrones óseos y automatizar el proceso de estimación de la edad ósea, particularmente durante el procesamiento de la imagen y la generación de resultados diagnósticos.

La mayoría de los estudios incluidos correspondieron a investigaciones observacionales, principalmente de tipo retrospectivo y estudios de validación de modelos, en los cuales se comparó el desempeño de los sistemas de inteligencia artificial con la evaluación realizada por especialistas evidenciando resultados consistentes y reproducibles.

La inteligencia artificial demuestra un impacto positivo en la práctica clínica al reducir el tiempo de evaluación, mejorar la precisión diagnóstica y disminuir la

variabilidad interobservador e intraobservador. Sin embargo, su rendimiento puede verse influenciado por el grupo etario y las características poblacionales, lo que resalta la necesidad de realizar validaciones en contextos locales.

VIII. REFERENCIAS BIBLIOGRÁFICAS

1. Cheung C, Chan JPK, Ng CWK, Lai WT, Fung KKF, Kan EYL. Validation of artificial intelligence for bone age assessment in Hong Kong children. *Hong Kong J Radiol* [Internet]. 2024; Disponible en: https://www.hkjr.org/system/files/v27n4_Validation.pdf
2. Lopez.I, Gamarra.A. Valoraciones médico legales sobre la determinación de la edad cronológica mediante pruebas radiológicas entorno a los 18 años. [Internet]. España. Disponible en:<https://www.endocrinologiapediatrica.org/modules.php?name=articulos&idarticulo=77&idlangart=ES>
3. Hwa.P, Mang,H. Evaluación de la edad ósea mediante inteligencia artificial en la población pediátrica coreana: una comparación de modelos de aprendizaje profundo entrenados con edades cronológicas saludables y edades de Greulich-Pyle como etiquetas. [Internet]. Corea.PubMed;2023. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/37899524/>
4. Gräfe D, Beeskow AB, Pfäffle R, Rosolowski M, Chung TS, DiFranco MD. Automated bone age assessment in a German pediatric cohort: agreement between an artificial intelligence software and the manual Greulich and Pyle method. *Eur Radiol* [Internet]. 2024;34(7):4407–13. Disponible en: <http://dx.doi.org/10.1007/s00330-023-10543-0>
5. García CJ, Zúñiga CE, Illanes SE. Bone age: A review of the most common methods of assessment. *Rev Chil Pediatr* [Internet]. 2018

- [citado el 10 de Julio 2025];89(1):79–89. Disponible en:
https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0370-41062018005000705&lng=en
6. Nguyen T, Hermann A-L, Ventre J, Ducarouge A, Pourchot A, Marty V, et al. High performance for bone age estimation with an artificial intelligence solution. *Diagn Interv Imaging* [Internet]. 2023;104(7–8):330–6. Disponible en:
<https://www.sciencedirect.com/science/article/pii/S221156842300075X?via%3Dihub>
 7. Kim JR, Shim WH, Yoon HM, Hong SH, Lee JS, Cho YA, et al. Computerized bone age estimation using deep learning based program: Evaluation of the accuracy and efficiency. *AJR Am J Roentgenol* [Internet]. 2017;209(6):1374–80. Disponible en:
<https://pubmed.ncbi.nlm.nih.gov/28898126/>
 8. Shahein H, Hyunkwang L, Randheer S, Brezo I, Jie C, Sjirk J, Ruth L, Sehyo Y, Michael S, Synho D, et al. La interpretación asistida por inteligencia artificial de las radiografías de la edad ósea mejora la precisión y reduce la variabilidad [Internet] PubMed; 2018. Disponible desde: <https://pubmed.ncbi.nlm.nih.gov/30069585/>
 9. Brais G, Carla P, Jeronimo P, Alfonso S, Enrique F, Alejandra P. Inteligencia artificial en pediatría: actualidad y retos [Internet] España, Asociación Española de Pediatría: 2024. Disponible desde: <https://analesdepediatría.org/es-inteligencia-artificial-pediatría-actualidad-retos-artículo-S1695403324000377>

- 10.** Wenhao Y, Pei F, Pan W, Li Wei Z. Evaluación de la edad ósea mediante diversas técnicas de imágenes médicas mejoradas por inteligencia artificial [Internet]. China: Mdpi; 2025. Disponible desde: <https://www.mdpi.com/2075-4418/15/3/257>
- 11.** Koumakis L. Deep learning models in genomics comput Struct Biotechnol J [Internet]. 2020;18:1466–73. Disponible en: <https://www.sciencedirect.com/science/article/pii/S2001037020303068>
- 12.** Özmen E, Özen Atalay H, Uzer E, Veznikli M. Una comparación de dos métodos basados en inteligencia artificial para evaluar la edad ósea en niños turcos: BoneXpert y VUNO Med-Bone Age [Internet]. 2024. Disponible desde: <https://www.dirjournal.org/articles/a-comparison-of-two-artificial-intelligence-based-methods-for-assessing-bone-age-in-turkish-children-bonexpert-and-vuno-med-bone-age/doi/dir.2024.242790>
- 13.** Lea WW-I, Hong S-J, Nam H-K, Kang W-Y, Yang Z-P, Noh E-J. External validation of deep learning-based bone-age software: a preliminary study with real world data. Sci Rep [Internet]. 2022;12(1):1232. Disponible en: <http://dx.doi.org/10.1038/s41598-022-05282-z>
- 14.** Fengdan W. Performance of an IA system for bone age assessment in Tibet. [Internet] 2021. Disponible en: <https://academic.oup.com/bjr/article/94/1120/20201119/7452296>

- 15.** Wang Y-M, Tsai T-H, Hsu J-S, Chao M-F, Wang Y-T, Jaw T-S. Automatic assessment of bone age in Taiwanese children: A comparison of the Greulich and Pyle method and the Tanner and Whitehouse 3 method. *Kaohsiung J Med Sci* [Internet]. 2020;36(11):937–43. Disponible en: <http://dx.doi.org/10.1002/kjm2.12268>
- 16.** Huang S, Su Z, Liu S, Chen J, Su Q, Su H, 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* [Internet]. 2023;53(6):1108–16. Disponible en: <http://dx.doi.org/10.1007/s00247-022-05569-3>
- 17.** Lotfi M, Abolpour N, Ghasemi M, Heydari H, Pourghayumi R. Potential of artificial intelligence for bone age assessment in Iranian children and adolescents: An exploratory study. *Arch Iran Med* [Internet]. 2025;28(4):198–206. Disponible en: <http://dx.doi.org/10.34172/aim.32070>
- 18.** Booz C, Yel I, Wichmann JL, Boettger S, Al Kamali A, Albrecht MH, et al. Artificial intelligence in bone age assessment: accuracy and efficiency of a novel fully automated algorithm compared to the Greulich-Pyle method. *Eur Radiol Exp* [Internet]. 2020;4(1):6. Disponible en: <http://dx.doi.org/10.1186/s41747-019-0139-9>
- 19.** Gao C, Hu C, Qian Q, Li Y, Xing X, Gong P, et al. Artificial intelligence model system for bone age assessment of preschool

children. *Pediatr Res* [Internet]. 2024;96(7):1822–8. Disponible en:
<http://dx.doi.org/10.1038/s41390-024-03282-5>

- 20.** Pape J, Hirsch FW, Deffaa OJ, DiFranco MD, Rosolowski M, Gräfe D. Applicability and robustness of an artificial intelligence-based assessment for Greulich and Pyle bone age in a German cohort. *Rofo* [Internet]. 2024;196(6):600–6. Disponible en:
<http://dx.doi.org/10.1055/a-2203-2997>

IX. TABLAS

Tabla 1 Características generales de los estudios.

N°	Autor	País	Año	Objetivo	N° imágenes / pacientes	Grupo Etario	Tecnología
1	Ming Wang (15)	Taiwán	2020	Evaluar aplicabilidad de GP y TW3 mediante BoneXpert y establecer referencias para Taiwán	563 niños	301 varones 262 mujeres Niños de 3 a 17 años	Radiografía digital (DR)
2	Wu-in Lea (13)	Corea del Sur	2022	Evaluar el desempeño del software BoneAge VunoMed en datos reales	474 radiografías	35 varones 439 mujeres Niños de 4 a 17 años	Radiografía digital (DR).
3	Fengdan Wang (14)	China (Región del Tibet)	2021	Validar si un sistema de IA entrenado con niños de la ciudad Han es aplicable en Tibet	385 radiografías	Niños de 4 a 18 años	Radiografía digital (DR)
4	Pape (20)	Alemania	2023	Evaluar la aplicabilidad y robustez del sistema IA basado en G	1253 niños y adolescentes	Niños de 3 a 16 años	Radiografía digital (DR)
5	Booz (18)	Alemania	2020	Evaluar la precisión y eficiencia de BoneXpert versión 2.1	514 radiografías	Niños de 3 a 17 años	Radiografía digital (DR)
6	Gao (19)	China	2024	Evaluar impacto de IA en variabilidad inter e intraobservador usando TW3 y RUS-CHN	94 niños	41 varones 53 mujeres Niños de 3 a 6 años	Radiografía digital (DR)

7	Shurong Huang (16)	China	2022	Evaluar el rendimiento de los métodos de la evaluación de la edad ósea apoyado por IA en comparación a los métodos que corresponden para la altura adulta real con la altura adulta pronosticada	1663 radiografías	762 mujeres Niñas de 6 a 11 años	Radiografía digital (DR)
8	Özmen (12)	Turquía	2025	Evaluar la aceptabilidad de dos programas de evaluación de la edad ósea apoyados en la IA en comparación al método GP en niños turcos	292 niños	Niños entre 1 y 15 años	Radiografía directa (DR)
9	Lofti (17)	Irán	2025	Evaluar si la edad ósea en niños puede determinarse con precisión mediante un modelo de inteligencia artificial basado en aprendizaje profundo y compararlo con el método de referencia GP	555 radiografías	Niños entre 1 a 18 años	Radiografía directa (DR)

Nota: Realizado por las autoras

Tabla 2 Modelos y técnicas de inteligencia artificial utilizados.

Autor	Tipo de IA	Método utilizado	Principales Hallazgos
Booz (18)	Software BoneXpert	Atlas de Greulich & Pyle (GP)	El sistema basado en IA mostró mayor precisión y eficiencia que el método de Greulich-Pyle, reduciendo el tiempo de lectura en un 87% y mostrando una alta correlación con los valores de referencia.
Gao (19)	Sistema Deepwise	Tanner Whitehouse 3 y RUS - CHN (TW3)	La IA redujo la variabilidad interobservador y mejoró la reproducibilidad intraobservador en la evaluación de la edad ósea en niños preescolares. La precisión del diagnóstico mejoró considerablemente.
Pape (20)	Software IB LabPANDA V.1.06	Atlas de Greulich & Pyle (GP)	El software de IA para la determinación de la edad ósea (EO) mostró buen rendimiento, alineándose con la edad cronológica de la población, aunque presenta errores notables en niños pequeños. El sistema fue robusto en proyecciones oblicuas, aun cuando la mano no estaba perfectamente posicionada.
Ming Wang (15)	Software BoneXpert v2.5.4.1.	Greulich & Pyle (GP) y Tanner Whitehouse (TW3)	El estudio evidenció que los métodos de estimación de edad ósea, como GP y TW3, presentan diferencias sistemáticas respecto a la edad cronológica, tendiendo a subestimar en niños pequeños y a sobreestimar en adolescentes, con variaciones según el sexo.

Wu-in Lea (13)	BoneAge v1.0.3. (VUNO Med)	Greulich & Pyle (GP)	Demostró que un software basado en deep learning presenta un rendimiento comparable, e incluso superior, al de especialistas humanos, mostrando una alta correlación con sus evaluaciones ($r \approx 0.98$). Además, el modelo evidenció menores errores frente a la edad cronológica en comparación con algunos evaluadores, lo que respalda su utilidad clínica.
Fengdan (14)	Sistema IA basado en ResNet (YITU Healthcare)	Greulich & Pyle (GP)	Evidenció que un sistema de inteligencia artificial para la estimación de la edad ósea alcanzó una alta precisión en comparación con expertos, con una concordancia dentro de ± 1 año del 84.67% en niños tibetanos y 89.41% en niños Han, y un error absoluto medio de aproximadamente 0.65 y 0.56 años, respectivamente.
Shurong Huang (16)	Deepwise Intelligent Diagnosis System	Greulich– Pyle(GP), Tanner– Whitehouse (TW3– RUS) y China 05 RUS–CHN	Los modelos automatizados presentan una alta concordancia con los métodos tradicionales, con diferencias mínimas en la estimación de la edad ósea. La combinación de la evaluación automatizada con métodos de predicción de talla mejora la precisión en la estimación de la altura final, lo que resulta clínicamente útil en el manejo de trastornos del crecimiento.

(RUS-CHN)

Özmen (12)	BoneXpert v3.0.3 y VUNO Med-BoneAge v1.0.3	Greulich-Pyle (GP)	Los modelos basados en deep learning, especialmente redes neuronales convolucionales, han demostrado gran capacidad para automatizar el proceso, reduciendo el tiempo de análisis y la variabilidad interobservador.
Lofti (17)	Modelo de red neuronal profunda (CNN)	Greulich-Pyle (GP)	El modelo muestra potencial como herramienta complementaria para los médicos, los estudios futuros deberían centrarse en mejorar la precisión de la predicción, reducir la variabilidad y validar el modelo en conjuntos de datos más grandes y diversos antes de considerar una implementación clínica generalizada.

Nota: Realizado por las autoras

Tabla 3 Diseños metodológicos de los estudios.

Autor	Tipo de estudio	Enfoque del estudio	Comparación realizada	Métricas utilizadas
Ming Wang (15)	Observacional retrospectivo	Validación de modelo	IA vs Greulich-Pyle	MAE, correlación
Wu-In Lea (13)	Observacional retrospectivo	Evaluación de desempeño	IA vs especialistas	MAE, RMSE, correlación
Fengdan Wang (14)	Observacional retrospectivo	Validación externa	IA vs método tradicional	Precisión (%)
Pape (20)	Observacional	Evaluación de robustez	IA vs edad cronológica	MAE

Booz (18)	retrospectivo Observacional retrospectivo	Validación del modelo	IA vs Greulich-Pyle	MAE, correlación
Shurong Huang (16)	Observacional retrospectivo	Comparacion de metodos	IA vs métodos manuales	MAE, ICC
Özmen (12)	Observacional retrospectivo	Comparación entre softwares	BoneXpert vs VUNO	ICC
Gao (19)	Observacional retrospectivo	Evaluación diagnóstica	IA vs TW3 y RUS-CHN	ICC
Lotfi (17)	Observacional retrospectivo	Desarrollo y validación de modelo	IA vs Greulich-Pyle	MAE, precisión

Nota: Realizado por las autoras

Tabla 4 Impacto y utilidad clínica de la inteligencia artificial.

AUTOR	ERROR MEDIO	ERROR ABSOLUTO MEDIO	DESVIACIÓN ESTÁNDAR DEL ERROR	ICC	COMPARACIÓN P.A Y OBLICUA	COMPARACIÓN ENTRE MÉTODOS (IA Y CLÍNICOS)
Alemania Pape (20)	• NR	• NR	• Varones: 3,0 (± 13,7) meses • mujeres: 1,7 (± 13,7) meses	• NR	• PA: 2,5 (± 13,7) meses • Oblicua: 1,8 (± 13,9) meses	La IA mejora la precisión y consistencia en la evaluación de la edad ósea, especialmente al manejar imágenes oblicuas que son menos estándar

Alemania Booz (18)	● NR	● NR	● MAD: 0,34 años, ● RSMD: 0,38 años	● NR	● NR	La IA aumenta la eficiencia al reducir significativamente el tiempo necesario para evaluar la edad ósea sin perder precisión, mejorando la consistencia y fiabilidad de los resultados entre diferentes observadores.
China Gao(19)	● 0,151 años	● 0,119 años	● NR	● 0,99	● NR	La IA mejora la precisión del diagnóstico de radiólogos, especialmente en niños en edad preescolar, reduciendo la variabilidad y aumentando la consistencia en las evaluaciones de la edad ósea.

Nota: Realizado por las autoras

ANEXOS

Anexo 1: Descripción de variables

VARIABLE	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	INDICADOR	TIPO Y ESCALA DE MEDICIÓN
POBLACIÓN PEDIÁTRICA	Conjunto de personas que se encuentran en la etapa de crecimiento y desarrollo comprendida entre la infancia y la adolescencia (6)	Se incluye en el estudio solo población pediátrica de 0 a 18 años de ambos sexos, que cuenten con imágenes radiográficas.	Edad cronológica (años)	Cuantitativa discreta / de razón
SEXO	Hace referencia a las características biológicas y fisiológicas que definen a hombres y mujeres (7)	Se incluye el sexo del paciente según lo registrado	Mujer Varón	Categoría dicotómica/nominal
EDAD ÓSEA	Es el grado de madurez esquelética que tiene un individuo, estimado con base en el desarrollo de los huesos comparado con estándares con los Greulich y Pyle (8)	Se mide a través de las radiografías de mano y muñeca izquierda, evaluadas por IA, obteniendo un valor estimado en años.	Edad ósea estimada (en años) Diferencia con edad cronológica	Cuantitativa continua/De Razón
IA (Inteligencia Artificial)	Es la capacidad de un sistema informático para ejecutar tareas que normalmente requieren inteligencia humana, como el análisis de imágenes médicas. (9)	Se considera el uso de un software o algoritmo automatizado (como redes neuronales convolucionales) capaz de estimar la edad ósea a partir de una imagen radiográfica.	Tipo de IA: ML, DL, CNN, BoneXpert, BoneAge, etc	Categoría politómica/nominal
EQUIPO DIGITAL	Equipos con una técnica que emplea placas sensibles a los rayos X, las cuales sirven para capturar las imágenes de manera directa durante	Se toma en cuenta si los establecimientos de salud cuentan con estos equipos de radiología digital.	Dispone de equipo DR	Categoría dicotómica/nominal

	los exámenes de la población pediátrica y luego transferirlos de manera inmediata a un sistema informático. (10)			
IMAGEN RADIOGRÁFICA	Representación visual obtenida por rayos X que muestra las estructuras óseas, útil para el diagnóstico médico. (11)	Se utilizan imágenes en proyección PA de mano y muñeca izquierda, en formato digital, utilizadas como insumo para la estimación ósea.	Número de estudios analizados Estructura anatómica/tipo de estudio	Categoría politómica

Anexo 2: Estrategia de búsqueda

FUENTE DE BÚSQUEDA	TÉRMINOS	FILTROS
PUBMED	"Child"[MeSH] OR "Pediatrics"[MeSH] OR "Adolescent"[MeSH] AND "Artificial Intelligence"[MeSH] OR "Machine Learning"[MeSH] OR "Deep Learning"[MeSH] AND "bone age" OR "Bone Age"[MeSH] OR "skeletal maturity" AND "Radiography"[MeSH] OR "X-Ray" OR "radiograph" OR "radiography"	<ul style="list-style-type: none"> ● Año: 2013 - 2025 ● Texto Completo ● Idioma Inglés
EMBASE	(deep learning.mp. or deep learning/) or (machine learning.mp. or machine learning/) or (artificial intelligence.mp. or artificial intelligence/) and (bone age.mp. or bone age/) or (skeletal age.mp. or skeletal age/) or (age estimation.mp.) and (radiography.mp. or radiography/) or (digital radiography.mp. or digital radiography/) or (medical imaging.mp. or medical imaging/) and (child.mp. or child/) or (children.mp. or children/) or (pediatric*.mp. or pediatrics/)	<ul style="list-style-type: none"> ● Año: 2013 - 2025 ● Texto Completo ● Artículos en revisión ● Humanos ● Idioma Inglés
IEE EXPLORE	"artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network*" AND "bone age" OR "skeletal age" OR "bone maturation" OR "age estimation" OR "bone age assessment" AND "digital radiograph*" OR "hand radiograph*" OR "medical imaging" OR "radiography" AND (child* OR pediatric* OR adolescent* OR youth)	<ul style="list-style-type: none"> ● Año: 2013 - 2025 ● Idioma: inglés ● Texto Completo
GOOGLE	("artificial intelligence" OR "machine learning" OR "deep	<ul style="list-style-type: none"> ● Año: 2013 - 2025

SCHOLAR	learning" OR AI) AND ("bone age" OR "skeletal maturity" OR "bone maturation" OR "age estimation") AND (child OR children OR pediatric OR adolescent) AND ("digital radiography" OR radiology OR "medical imaging" OR "health facilities" OR hospitals)	<ul style="list-style-type: none"> ● Idioma: inglés y español ● Texto Completo ● Humanos
----------------	--	---

Anexo 3: Diagrama Prisma

