

Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), enero-febrero 2026,
Volumen 10, Número 1.

https://doi.org/10.37811/cl_rcm.v10i1

IMPORTANCIA DEL MÉTODO GREULICH AND PYLE EN LA PRÁCTICA CLÍNICA UNA REVISIÓN SISTEMÁTICA

**IMPORTANCE OF THE GREULICH AND PYLE
METHOD IN CLINICAL PRACTICE:
A SYSTEMATIC REVIEW**

Miguel Noé Diaz Ataucure MG
Las Violetas Independencia, Perú

DOI: https://doi.org/10.37811/cl_rcm.v10i1.23289

Importancia del Método Greulich and Pyle en la Práctica Clínica Una Revisión Sistemática

Miguel Noé Diaz Ataucure MG¹migueldiaz_rad@hotmail.com<https://orcid.org/0000-0003-0054-7850>Calle las Avellanas N°176 – Urb
Las Violetas Independencia
Lima Perú

RESUMEN

La evaluación de la edad ósea en niños y adolescentes, es importante en la práctica clínica, con la finalidad de no tener variabilidad con la edad biológica; el método de Greulich and Pyle (GP) es la más empleada, práctica y versátil en su aplicación, existiendo también métodos automatizados con Inteligencia Artificial (IA). Se analiza y evalúa la eficacia con respecto a otros métodos, determinando su relevancia en la práctica clínica actual. Se realizaron búsqueda en SCOPUS, GOOGLE SCHOLAR y PUB MED, obteniendo 16 estudios que cumplieron con los criterios de inclusión y exclusión, encontrándose que GP es una herramienta útil, eficaz y precisa para la evaluación de la edad ósea, con una alta fiabilidad y validez interobservador e intraobservador. Algunos estudios señalaron que GP puede subestimar la edad ósea en algunos grupos de población. Finalmente, se determinó un alto grado de importancia de GP en la práctica clínica encontrando también métodos automatizados.

Palabras clave: método de Greulich y Pyle, evaluación de madurez esquelética, radiología pediátrica, crecimiento y desarrollo óseo, maduración ósea y comparación de métodos

¹ Autor principal

Correspondencia: migueldiaz_rad@hotmail.com

Importance of the Greulich and Pyle Method in Clinical Practice: A Systematic Review

ABSTRACT

Bone age assessment in children and adolescents is important in clinical practice to avoid variability with biological age. The Greulich and Pyle (GP) method is the most widely used, practical, and versatile method, and automated methods using Artificial Intelligence (AI) also exist. Its effectiveness is analyzed and evaluated compared to other methods, determining its relevance in current clinical practice. Searches were conducted in SCOPUS, GOOGLE SCHOLAR, and PUB MED, yielding 16 studies that met the inclusion and exclusion criteria. It was found that GP is a useful, effective, and accurate tool for assessing bone age, with high inter- and intra-observer reliability and validity. Some studies indicated that GP may underestimate bone age in some population groups. Finally, a high degree of importance of GP in clinical practice was determined, also finding automated methods.

Keywords: Greulich-Pyle method, assessment of skeletal maturity, pediatric radiology, bone growth and development, bone maturation and comparison of methods

*Artículo recibido 02 enero 2026
Aceptado para publicación: 30 enero 2026*



INTRODUCCIÓN

En el presente artículo, desarrollaremos sobre la importancia del método de GREULICH AND PYLE (GP), en comparación a otras metodologías de evaluación de la maduración ósea, en la aplicación de la práctica clínica. En las últimas décadas, han surgido métodos alternativos como el de Tanner-Whitehouse (TW2/TW3) incluyendo poblaciones heterogéneas, comparándolo con el método de GP (1,2,3,4,5,6) y, más recientemente, sistemas automatizados basados en inteligencia artificial (IA) que prometen mayor objetividad, precisión y eficiencia. Estas tecnologías analizan las radiografías mediante algoritmos para reducir la subjetividad y el tiempo de evaluación, estableciéndose así una notable relevancia en el uso de la práctica clínica actual (7,8,9,10,11).

Se determinó la vigencia y la popularidad del método GP, del cual radica en su simplicidad, rapidez y bajo costo, lo que lo convierte en una herramienta accesible, especialmente en entornos con recursos limitados (12). Sin embargo, su validez y precisión han sido objeto de debate. Las críticas se centran principalmente en la subjetividad inherente a la interpretación visual, que conduce a una considerable variabilidad interobservador, y en el hecho de que el atlas original se basó en una población de niños caucásicos de Cleveland, Ohio, en la década de 1930, lo que plantea dudas sobre su aplicabilidad a poblaciones étnicas y geográficas diversas en la actualidad (6,13,11,14).

La evaluación de la maduración ósea es un componente fundamental en la pediatría y la endocrinología para el diagnóstico y seguimiento de trastornos del crecimiento y desarrollo (15,16). Entre las diversas metodologías existentes, el método de GP, desarrollado en la década de 1950, ha sido históricamente el más utilizado a nivel mundial (12,16). Esta técnica consiste en comparar una radiografía de la mano y muñeca izquierda del paciente con una serie de imágenes estándar de un atlas de referencia para estimar la edad ósea (17,16). A pesar de su uso cotidiano, no está libre de desafíos, particularmente en lo concerniente a la variabilidad interobservador y la imperiosa necesidad de una estandarización rigurosa para garantizar una buena reproducibilidad (12,18). El método de GP, demostró una fiabilidad en sus resultados de moderada a buena, dependiendo de la experiencia del clínico evaluador (13), destacando la importancia de esta herramienta en clínicas pediátricas para identificar trastornos del crecimiento y displasias esqueléticas, subrayando la importancia de adaptar los métodos de evaluación a diferentes poblaciones (15,19).



En América Latina y el Caribe, estos métodos afrontan desafíos, relacionado con la diversidad étnica y diferentes poblaciones (20). Si bien el atlas de GP se basa en datos de niños caucásicos de Cleveland, Ohio; su aplicación en Latinoamérica requiere cierta consideración de las diferencias étnicas y geográficas (12); ello es fundamental para la validación en diferentes poblaciones latinoamericanas, asegurando la precisión y aplicabilidad (20). En México, el método de GP, considera ciertas particularidades por la diversidad étnica de la población (21,19).

Aun así, disponiendo de los métodos automatizados y computarizados basados en IA (22,23,7,10,11), el método tradicional de GP continúa siendo relevante en la práctica clínica, especialmente en ambientes con recursos limitados o donde no se dispone de tecnología avanzada (12,16), destacando la importancia de la variabilidad interobservador y, resaltando la necesidad de una capacitación del profesional de la salud para minimizar los errores (6).

En la presente revisión sistemática de la literatura científica, se analizarán las evidencias favorables con respecto a la precisión y fiabilidad del método de GP, haciendo la comparación con otros diferentes métodos de evaluación de edad ósea (1,24,11). Se compara su desempeño con otras técnicas de evaluación ósea, principalmente los sistemas automatizados, para sintetizar la evidencia disponible, evaluar su fiabilidad en diferentes poblaciones y ofrecer recomendaciones para la práctica clínica y futuras líneas de investigación.

En resumen, la presente investigación corrobora que el método de GP sigue siendo importante, pero en el uso moderno de la práctica clínica podría beneficiarse de la integración de tecnologías más avanzadas y evolucionadas, como el uso de algoritmos de IA, la experiencia del evaluador y la disponibilidad de los recursos (22). Se debe tener en cuenta la elección del método, para establecer las necesidades propias de cada caso (14,10,24).

Pacientes Y Métodos:

En la presente investigación, se realizó una revisión sistemática siguiendo las directrices del Manual Cochrane. La pregunta de investigación se estructuró bajo el enfoque PICO:

- P (Población) = Niños y adolescentes (0–18 años).
- I (Intervención/Exposición) = Aplicación del método de GP.
- C (Control o comparación) = Métodos TW2/TW3, IA y software automatizado.



- O (Resultado = Concordancia con la edad cronológica, precisión diagnóstica y relevancia clínica.

Según lo antes descrito, podemos plantear la siguiente pregunta: ¿cuál es la importancia del método de GP, en comparación con las otras metodologías de evaluación de la maduración ósea, en la aplicación de la práctica clínica?

Búsqueda sistemática

Se consultaron PUB MED, SCOPUS y GOOGLE SCHOLAR en febrero de 2025. Se utilizaron términos como: "Greulich y Pyle", "madurez esquelética", "pediátrica", "evaluación de la edad ósea", "evaluación de la madurez esquelética", "radiología", "práctica clínica", "precisión diagnóstica", "evaluación esquelética pediátrica" y "Tanner Whitehouse". Cabe recalcar que las estrategias de búsqueda se aplicaron de manera idéntica en las tres bases de datos con la finalidad de garantizar consistencia; combinados con operadores booleanos en cuatro algoritmos estandarizados (tabla 1). Se incluyeron artículos en inglés o español publicados entre 2005 y 2025.

Los resultados de dichas búsquedas fueron: En SCOPUS se obtuvo 710 resultados. En GOOGLE SCHOLAR se obtuvo 662 resultados; finalmente, en PUB MED se obtuvo 119 resultados. Concretamente, se obtuvieron 1491 resultados en los tres buscadores (SCOPUS, GOOGLE SCHOLAR y PUB MED). Antes de proceder a la selección de artículos, se definieron los criterios de inclusión y exclusión.

Criterios de inclusión:

- Población pediátrica (0–18 años).
- Uso del método GP.
- Comparación con otros métodos.
- Estudios originales, revisiones sistemáticas o metaanálisis.
- Idioma: inglés o español.

Criterios de exclusión:

- Estudios en adultos.
- Estudios que no utilicen el método Greulich y Pyle.
- Artículos no originales (cartas, opiniones, tesis).
- Publicaciones en otros idiomas.



- Datos duplicados o no libres de acceso.

De 1491 artículos iniciales, se eliminaron duplicados y se aplicaron los criterios de selección por título, resumen y texto completo. Finalmente, se incluyeron 16 artículos para su análisis respectivo (figura 1). Se categorizaron según el enfoque metodológico planteado: GP puro (n=6), IA (n=2), software automatizado (n=6) y comparación GP vs. TW (n=2).

RESULTADOS

Se demostró que el método de GP sigue siendo una herramienta extensamente utilizada y valorada en la aplicación de la práctica clínica, principalmente en el contexto de la evaluación y valoración de la edad ósea en niños y adolescentes (15,16,7). Se determinó que la simplicidad, rapidez y bajo costo, es su principal ventaja de dicho método tradicional, convirtiéndole en una elección para muchos profesionales de la salud, particularmente en lugares o sectores con recursos limitados (12).

Limitaciones y comparación con otros métodos:

A pesar de su uso extendido, la revisión evidenció limitaciones significativas. La principal es la variabilidad interobservador, con un Coeficiente de Correlación Intraclase (ICC) que en múltiples estudios se sitúa por debajo de 0.85, indicando una fiabilidad entre moderada y buena que depende en gran medida de la experiencia del evaluador, haciendo resaltar la importancia de la capacitación y la experiencia del observador o evaluador, ocasionando una subjetividad en la interpretación visual (12,6,25,8). Un estudio en niños turcos, subestimó la edad ósea en poblaciones no caucásicas, resaltando la subjetividad inherente a la interpretación visual (13).

Comparando con otros métodos, el método de Tanner-Whitehouse (TW2 y TW3) (1,3,4,5,11,6) y los sistemas computarizados basados en IA, el método de GP mostró resultados heterogéneos pero reveladores. Los sistemas computarizados y los algoritmos de IA demuestran consistentemente resultados óptimos y objetividad, mejorando la precisión y eficiencia (22). Hubo analogía entre el método de GP y otras metodologías, resaltando y sugiriendo que, puede ser una alternativa válida en ciertas situaciones (26,23,27,3,8). La alta correlación entre GP y software automatizado como BoneXpert ($R = 0.95$) sugiere que estos sistemas, aunque basados en los mismos principios, logran estandarizar y objetivar la evaluación (26,23,24). Otras investigaciones determinaron que la IA pueden ofrecer una mayor precisión diagnóstica y objetividad, con un error absoluto de 0.89 meses a diferencia

del método de GP manual con 1.38 meses (7,2,3,8,9,10,11), disminuyendo la variabilidad interobservador, optimizando la eficiencia del proceso de evaluación (22,28,29,13,30,31,27). Se destacó que la IA mejora la eficiencia, reduciendo los tiempos de evaluación hasta en un 70% y alcanzando un índice de confianza del 95% en la consistencia entre evaluadores con diferentes niveles de experiencia (32). En otro estudio la IA, busca mejorar la precisión y la consistencia interobservador realizadas por médicos con diferentes niveles de experiencia para mejorar la eficiencia en la práctica clínica, con 95% de índice de confianza (33).

Validación en diferentes poblaciones y nuevos enfoques

En diversos estudios, se destacó la validación del atlas de GP para diversas poblaciones no caucásicas, siendo importante para asegurar su precisión y aplicabilidad (15,20,19,9,11). Estudios en poblaciones de México, Canarias, Taiwán y en niños de ascendencia marroquí en Europa, subrayan la importancia de adaptar los estándares de referencia a las características étnicas y geográficas de cada grupo para garantizar la precisión del método (20,19,22,11,34). Se validó el análisis computarizado de la evaluación de la edad ósea en un grupo de población pediátrica norteamericana, demostrando su alta precisión (29). También se determinó la eficacia de la IA en una población pediátrica alemana, a comparación del método manual de GP, obteniendo buena concordancia, pudiendo ser una alternativa viable y eficiente (9).

La determinación y elección del método GP debe considerarse a las necesidades clínicas, la disponibilidad de los recursos y las características específicas de cada paciente (16,35). Se investigan sistemas basados en otras localizaciones anatómicas como el codo, el hombro o el calcáneo, también el uso de la madurez dental como un indicador biológico adicional (36,37,38,39). El desarrollo de técnicas como la ultrasonografía para la evaluación ósea en lactantes y el uso de aprendizaje profundo (*deep learning*) para la automatización del proceso, marcan el futuro de este campo diagnóstico, sugiriendo que el uso de esta técnica podría ser prometedora para futuras evaluaciones (40). En reciente investigación, se destacó, a futuro, el uso de la IA, potenciando y resaltando el método de GP para mejorar la precisión y la eficiencia (22).

En la tabla 2 se presenta, cronológicamente, desde la más actual (2024) hasta la más antigua (2007), las características de las dieciséis revistas que cumplieron los criterios de inclusión.



Adicionalmente en la tabla 3, se muestra un estudio comparativo de las revistas seleccionadas, en el orden cronológico anteriormente descrito.

DISCUSIÓN

En la presente revisión sistemática confirman la persistente relevancia del método de Greulich y Pyle (GP) en la práctica clínica, al tiempo que destaca sus limitaciones inherentes y la creciente superioridad de las alternativas tecnológicas. Su facilidad de uso, rapidez y bajo costo lo mantienen como una herramienta valiosa para la evaluación inicial, especialmente en entornos con recursos limitados. No obstante, su uso sin considerar sus desventajas puede conducir a decisiones clínicas subóptimas.

Significado de los resultados en el contexto clínico

En el contexto clínico, el método GP sigue siendo ampliamente utilizado en pediatría y endocrinología debido a su practicidad, accesibilidad y bajo costo (12). En contextos con escaso acceso a tecnologías avanzadas, GP representa una opción viable y funcional. La Tabla 4 evidencia su predominio histórico en la evaluación de la edad ósea. Sin embargo, su precisión y reproducibilidad son limitadas por la variabilidad interobservador y la subjetividad inherente al método (6).

Al confrontarse con el método de GP con la de Tanner-Whitehouse (TW2 y TW3) y otras variantes metodológicas de evaluación y maduración ósea (1,2,4,5,11,6), como los sistemas automatizados basados en IA, los algoritmos computarizados demuestran ventajas significativas. Los coeficientes de correlación reportados son: GP vs. BoneXpert: $R = 0.95$; GP vs. IA: $R = 0.96$; GP vs. TW3: $R = 0.87$ (22,28,29,13,30,31,27,8,9). Estos sistemas mejoran la objetividad, reducen la variabilidad interobservador y optimizan la consistencia diagnóstica, lo que potencialmente mejora la toma de decisiones clínicas en trastornos del crecimiento y desarrollo esquelético (7,11).

Estudios recientes demuestran que la IA reduce los tiempos de evaluación en un 70%, procesa grandes volúmenes de estudios con alta eficiencia y logra una concordancia interobservador del 95%, independientemente del nivel de experiencia del clínico (30). Esto refuerza su utilidad en entornos hospitalarios de alta demanda y su potencial para estandarizar la evaluación de la edad ósea a escala poblacional.



Limitaciones de los estudios revisados

La principal limitación identificada es la variabilidad interobservador del método GP, que compromete su reproducibilidad. Además, los estudios revisados presentan importantes restricciones metodológicas. En primer lugar, las poblaciones evaluadas son heterogéneas en términos étnicos, geográficos y socioeconómicos, lo que dificulta la comparabilidad directa de resultados (19,7,2,9,10,11). En segundo lugar, la mayoría de las investigaciones se centran en la precisión y fiabilidad metodológica (coeficiente de correlación intraclase promedio: GP = 0.81; IA = 0.94; Figura 2), sin evaluar el impacto real en los desenlaces clínicos ni en el manejo terapéutico de los pacientes. En tercer lugar, hay una marcada falta de generalización: la mayoría de los algoritmos y validaciones se han desarrollado en poblaciones específicas (europeas o asiáticas), lo que limita su aplicabilidad en otras regiones del mundo (7,2,9,10,11).

Asimismo, existe escasa evidencia que evalúe directamente la utilidad clínica del GP frente a métodos más avanzados. Aunque abundan los estudios sobre precisión (26,23,7,40,3,8,9,10), pocos analizan cómo la elección del método influye en decisiones diagnósticas o en la evolución de patologías del crecimiento.

Recomendaciones para la práctica clínica y futuras investigaciones

Adoptar métodos más precisos y objetivos: Siempre que sea posible, se recomienda priorizar sistemas automatizados basados en IA (23,7,8,9,10,11), ya que reducen la variabilidad interobservador y mejoran la consistencia diagnóstica, tal como se ilustra en la figura 3, la precisión diagnóstica entre el método AI y el GP.

Estandarizar la aplicación del método de GP: En entornos donde GP sigue siendo la única opción viable, es crucial implementar programas de capacitación estandarizados y verificar la experiencia de los evaluadores. Su bajo costo, simplicidad y aceptabilidad lo hacen adecuado como herramienta de tamizaje inicial, siempre que se reconozcan sus limitaciones.

Validar el atlas de GP en poblaciones diversas: Se requieren estudios de validación en distintos grupos étnicos y geográficos, ya que la maduración ósea varía significativamente entre poblaciones. Investigaciones en niños turcos (13), niños de la población de Lahore de Pakistán (19), población



alemana (9) y taiwanesa (11), ya han demostrado desviaciones importantes respecto al atlas original, lo que subraya la necesidad de adaptaciones locales.

Evaluar el impacto clínico de los diferentes métodos de evaluación ósea: En futuras investigaciones deben centrarse en desenlaces clínicos reales: eficacia del tratamiento, precisión diagnóstica, adherencia terapéutica y costo-efectividad. Esto permitirá determinar no solo qué método es más preciso, sino cuál mejora efectivamente la atención al paciente.

Integrar la IA en la práctica clínica: Se recomienda explorar la implementación de algoritmos de IA como apoyo diagnóstico para radiólogos y endocrinólogos (23,24). Estos sistemas pueden integrarse en flujos de trabajo clínicos para acelerar la interpretación, reducir errores y facilitar decisiones basadas en evidencia.

Desarrollar métodos alternativos basados en otras regiones anatómicas: Investigar la utilidad de localizaciones distintas a la mano y muñeca izquierda —como codo, hombro o calcáneo— podría ser crucial en casos donde la radiografía estándar no es factible (36,37,38). Esto ampliaría las opciones diagnósticas en trauma, malformaciones o contextos forenses.

Explorar la combinación de métodos complementarios: La integración de la madurez dental con la evaluación esquelética (GP o TW) podría ofrecer estimaciones más robustas de la edad biológica (39). Además, técnicas como el ultrasonido en niños menores de 6 años, aplicando criterios del atlas GP, muestran resultados prometedores (40) y merecen mayor investigación.

En conclusión, el método de Greulich y Pyle sigue siendo una herramienta clínica útil, especialmente en contextos de recursos limitados, pero requiere estandarización rigurosa para minimizar errores subjetivos. Por otro lado, los sistemas basados en IA emergen como alternativas superiores en precisión, eficiencia y reproducibilidad (figura 4). Se propone un enfoque híbrido: utilizar GP como método de entrada en entornos con escasa tecnología, y adoptar la IA como estándar de oro en centros con mayor capacidad, siempre acompañado de validación local, capacitación continua y evaluación del impacto clínico real.

Financiamiento

La presente investigación no ha recibido ayudas específicas provenientes de agencias del sector público, sector comercial o entidades sin ánimo de lucro.



Conflicto de intereses

El autor declara no tener ningún conflicto de intereses para el desarrollo del presente artículo.]

REFERENCIAS BIBLIOGRAFICAS

1. Mughal AM, Hassan N, Ahmed A. Bone age assessment methods: A critical review. *Pak J Med Sci.* 2014;30(1):211-5. doi:10.12669/pjms.301.4295.
2. Pinchi V, De Luca F, Ricciardi F, Focardi M, Piredda V, Mazzeo E, Norelli GA. Skeletal age estimation for forensic purposes: A comparison of GP, TW2 and TW3 methods on an Italian sample. *Forensic Sci Int.* 2014;238:83-90. doi:10.1016/j.forsciint.2014.02.030.
3. Son SJ, Song Y, Kim N, Do Y, Kwak N, Lee MS, Lee BD. TW3-based fully automated bone age assessment system using deep neural networks. *IEEE Access.* 2019;7:33346-33358. doi:10.1109/ACCESS.2019.2905855.
4. Shin NY, Lee BD, Kang JH, Kim HR, Oh DH, Lee BI, Kim SH, Lee MS, Heo MS. Evaluation of the clinical efficacy of a TW3-based fully automated bone age assessment system using deep neural networks. *Imaging Sci Dent.* 2020 Sep;50(3):237-243. doi: 10.5624/isd.2020.50.3.237.
5. Zhou XL, Wang EG, Lin Q, Dong GP, Wu W, Huang K, et al. Diagnostic performance of convolutional neural network-based Tanner-Whitehouse 3 bone age assessment system. *Quant Imaging Med Surg.* 2020 Mar;10(3):657-667. doi: 10.21037/qims.2020.02.20.
6. Geng J, Zhang W, Ge Y, Wang L, Huang P, Liu Y, Shi J, Zhou F, Ma K, Blake GM, et al. Inter-rater variability and repeatability in the assessment of the Tanner-Whitehouse classification of hand radiographs for the estimation of bone age. *Skeletal Radiol.* 2024 Dec;53(12):2635-2642. doi: 10.1007/s00256-024-04664-w.
7. Zhang J, Lin F, Ding X. Automatic determination of the Greulich-Pyle bone age as an alternative approach for Chinese children with discordant bone age. *Horm Res Paediatr.* 2016;86(2):83-9. doi:10.1159/000446434.



8. Gao C, Hu C, Qian Q, Li Y, Xing X, Gong P, Lin M, Ding Z. Artificial intelligence model system for bone age assessment of preschool children. *Pediatr Res.* 2024;96(7):1822-1828. doi:10.1038/s41390-024-03282-5.
9. 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.* 2024 Jul;34(7):4407-4413. doi:10.1007/s00330-023-10543-0.
10. Pan I, Baird GL, Mutasa S, Merck D, Ruzal-Shapiro C, Swenson DW, Ayyala RS. Rethinking Greulich and Pyle: A Deep Learning Approach to Pediatric Bone Age Assessment Using Pediatric Trauma Hand Radiographs. *Radiol Artif Intell.* 2020 Jul 29;2(4):e190198. doi: 10.1148/ryai.2020190198.
11. Wang YM, Tsai TH, Hsu JS, Chao MF, Wang YT, Jaw TS.. 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.* 2020 Nov;36(11):937-943. doi: 10.1002/kjm2.12268.
12. Gertych A, Zhang A, Sayre J, Pospiech-Kurkowska S, Huang HK. Bone age assessment of children using a digital hand atlas. *Comput Med Imaging Graph.* 2007;31(4-5):322-331. doi:10.1016/j.compmedimag.2007.02.012.
13. Sariyilmaz K, Abali S, Ziroglu N, Cingoz T, Ozkunt O, Abali ZY, Kalayci CB, Hayretci M, Semiz S. Interdisciplinary and intraobserver reliability of the Greulich-Pyle method among Turkish children. *J Pediatr Endocrinol Metab.* 2023;36(12):1181-5. doi:10.1515/jpem-2023-0303.
14. Breen MA, Tsai A, Stamm A, Kleinman PK. Bone age assessment practices in infants and older children among Society for Pediatric Radiology members. *Pediatric Radiology.* 2016;46(9):1269-1274. doi:10.1007/s00247-016-3618-7.
15. Martín Pérez IM, Martín Pérez SE, Vega González JM, Molina Suárez R, García Hernández AM, Rodríguez Hernández F, Herrera Pérez M. The Validation of the Greulich and Pyle Atlas for



- Radiological Bone Age Assessments in a Pediatric Population from the Canary Islands. *Healthcare* (Switzerland). 2024;12(18):1847. doi:10.3390/healthcare12181847.
16. Cavallo F, Mohn A, Chiarelli F, Giannini C. Evaluation of bone age in children: a mini-review. *Front Pediatr*. 2021;9:580314. doi:10.3389/fped.2021.580314.
 17. Van Rijn RR, Lequin MH, Thodberg HH. Automatic determination of Greulich and Pyle bone age in healthy Dutch children. *Pediatric Radiology*. 2009;39(6):591-597. doi:10.1007/s00247-008-1090-8.
 18. Dahlberg PS, Mosdøl A, Ding Y, Bleka Ø, Rolseth V, Straumann GH, et al. A systematic review of the agreement between chronological age and skeletal age based on the Greulich and Pyle atlas. *Eur Radiol*. 2019;29:2936-2948. doi:10.1007/s00330-018-5923-5.
 19. Uddin SZ, Zarif P, Bashir N. Epiphyseal fusion of radius and ulna at wrist joint in male and female in the population of Lahore - An age estimation criterion. *Med Forum Monthly*. 2018;29(6):85-7.
 20. Martín Pérez SE, Martín Pérez IM, Vega González JM, Molina Suárez R, León Hernández C, Rodríguez Hernández F, Herrera Pérez M. Precision and accuracy of radiological bone age assessment in children among different ethnic groups: a systematic review. *Diagnostics*. 2023;13(19):3124. doi:10.3390/diagnostics13193124.
 21. Klünder-Klünder M, Espinosa-Espindola M, Lopez-Gonzalez D, Loyo MS-C, Suárez PD, Miranda-Lora AL. Skeletal maturation in the current pediatric Mexican population. *Endocrine Practice*. 2020;26(10):1053-1061. doi:10.4158/EP-2020-0047.
 22. 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.
 23. Thodberg HH, Thodberg B, Ahlkvist J, Offiah AC. Autonomous artificial intelligence in pediatric radiology: the use and perception of BoneXpert for bone age assessment. *Pediatr Radiol*. 2022;52(7):1338-46. doi:10.1007/s00247-022-05295-w.



24. Özmen E, Özen Atalay H, Uzer E, Veznikli M. A comparison of two artificial intelligence-based methods for assessing bone age in Turkish children: BoneXpert and VUNO Med-Bone Age. *Diagn Interv Radiol*. 2024 Sep 2. doi: 10.4274/dir.2024.242790.
25. Halabi SS, Prevedello LM, Kalpathy-Cramer J, Mamonov AB, Bilbily A, Cicero M, et al. The RSNA pediatric bone age machine learning challenge. *Radiology*. 2019;290(2):498-503. doi:10.1148/radiol.2019190344.
26. Pose Lepe G, Villacrés F, Silva Fuente-Alba C, Guiloff S. 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):606-611. doi:10.4067/S0370-41062018000500606.
27. Kim JK, Park D, Chang MC. Assessment of bone age based on hand radiographs using regression-based multi-modal deep learning. *Life*. 2024;14(6):774. doi:10.3390/life14060774.
28. Kim JR, Shim WH, Yoon HM, Hong SH, Lee JS, Cho YA, Kim S. Computerized bone age estimation using deep learning based program: evaluation of the accuracy and efficiency. *AJR Am J Roentgenol*. 2017;209(6):1374-1380. doi:10.2214/AJR.17.17663.
29. Bowden JJ, Bowden SA, Ruess L, Adler BH, Hu H, Krishnamurthy R. Validation of automated bone age analysis from hand radiographs in a North American pediatric population. *Pediatr Radiol*. 2022;52(7):1347-55. doi:10.1007/s00247-022-05310-0.
30. Uddin NS, Hasan W, Wasid M, Ali R. Deep learning for bone age assessment: Current status and future prospects. *Artif Intell Biomed Mod Healthc Inform*. 2025;193-201.
31. Zhao K, Ma S, Sun Z, Liu X, Zhu Y, Xu Y, Wang X. 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.
32. Gerges M, Eng H, Chhina H, Cooper A. Modernization of bone age assessment: comparing the accuracy and reliability of an artificial intelligence algorithm and shorthand bone age to Greulich and Pyle. *Skeletal Radiol*. 2020;49(9):1449-57. doi:10.1007/s00256-020-03429-5.



33. Wang X, Zhou B, Gong P, Zhang T, Mo Y, Tang J, et al. Artificial Intelligence-Assisted Bone Age Assessment to Improve the Accuracy and Consistency of Physicians With Different Levels of Experience. *Front Pediatr*. 2022;10:818061. doi:10.3389/fped.2022.818061.
34. Boitsios G, De Leucio A, Preziosi M, Seidel L, Aparisi Gómez MP, Simoni P. Are Automated and Visual Greulich and Pyle-Based Methods Applicable to Caucasian European Children With a Moroccan Ethnic Origin When Assessing Bone Age? *Cureus*. 2021 Feb 21;13(2):e13478. doi:10.7759/cureus.13478.
35. Prokop-Piotrkowska M, Marszałek-Dziuba K, Moszczyńska E, Szalecki M, Jurkiewicz E. Traditional and new methods of bone age assessment-an overview. *J Clin Res Pediatr Endocrinol*. 2021;13(3):251. doi:10.4274/jcrpe.galenos.2020.2020.006.
36. Sinkler M.A., Furdock R.J., Chen D.B., Sattar A., Liu R.W. The Systematic Isolation of Key Parameters for Estimating Skeletal Maturity on Lateral Elbow Radiographs. *Journal of Bone and Joint Surgery*. 2022; 104(22): 1993-1999. doi:10.2106/JBJS.22.00312.
37. Furdock R.J., Kuo A., Chen K.J., Liu R.W. Applicability of Shoulder, Olecranon, and Wrist-based Skeletal Maturity Estimation Systems to the Modern Pediatric Population. *Journal of Pediatric Orthopaedics*. 2023; 43(7): 465-469. doi:10.1097/BPO.0000000000002430.
38. Munger A.M., Yu K.E., Li D.T., Furdock R.J., Boeyer M.E., Duren D.L., Weber D.R., Cooperman D.R. The More the Merrier: Integrating Multiple Models of Skeletal Maturity Improves the Accuracy of Growth Prediction. *Journal of Pediatric Orthopaedics*. 2021; 41(7): E580-E584. doi:10.1097/BPO.0000000000001847.
39. Rai V, Saha S, Yadav G, Tripathi AM, Grover K. Dental and skeletal maturity- A biological indicator of chronologic age. *Journal of Clinical and Diagnostic Research*. 2014;8(9):ZC60-ZC64. doi:10.7860/JCDR/2014/10079.4862.
40. Bilgili Y, Hizel S, Kara SA, Sanli C, Erdal HH, Altinok D. Accuracy of skeletal age assessment in children from birth to 6 years of age with the ultrasonographic version of the Greulich-Pyle atlas. *J Ultrasound Med*. 2003;22(7):683-90. doi:10.7863/jum.2003.22.7.683.

ANEXOS



Figura 1 : Diagrama de flujo, según el método PRISMA

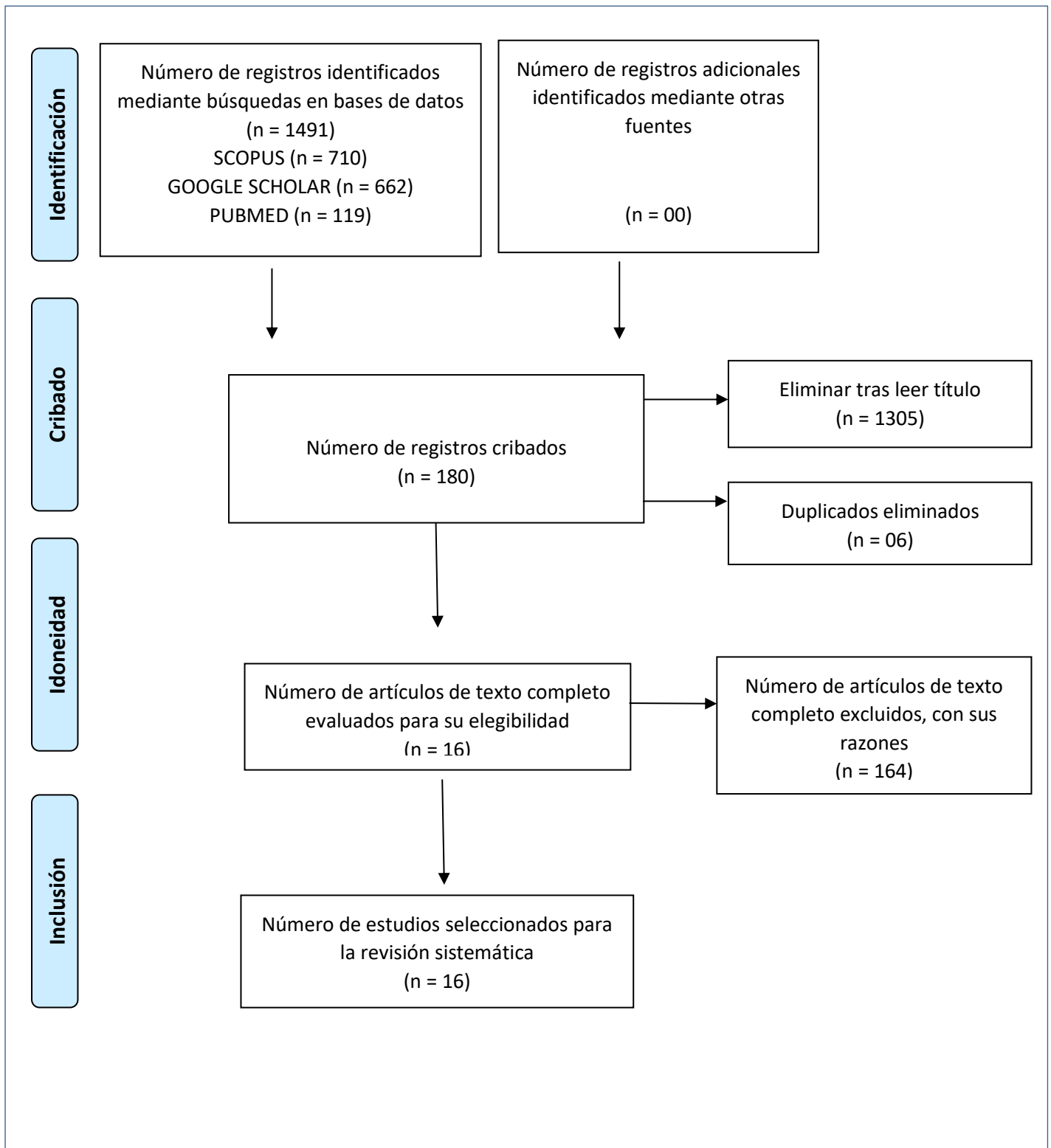


Figura 2. Fiabilidad Interobservador (ICC). El gráfico de ICC muestra claramente que el Método IA tiene una fiabilidad interobservador significativamente mayor (ICC promedio = 0.94) en comparación con el Método GP (ICC promedio = 0.81). La diferencia estadística ($p < 0.001$) confirma que esta mejora no es aleatoria. Esto indica que las mediciones realizadas con el Método IA son mucho más consistentes entre diferentes observadores, lo cual es crucial para la estandarización y reproducibilidad en la evaluación de la edad ósea.

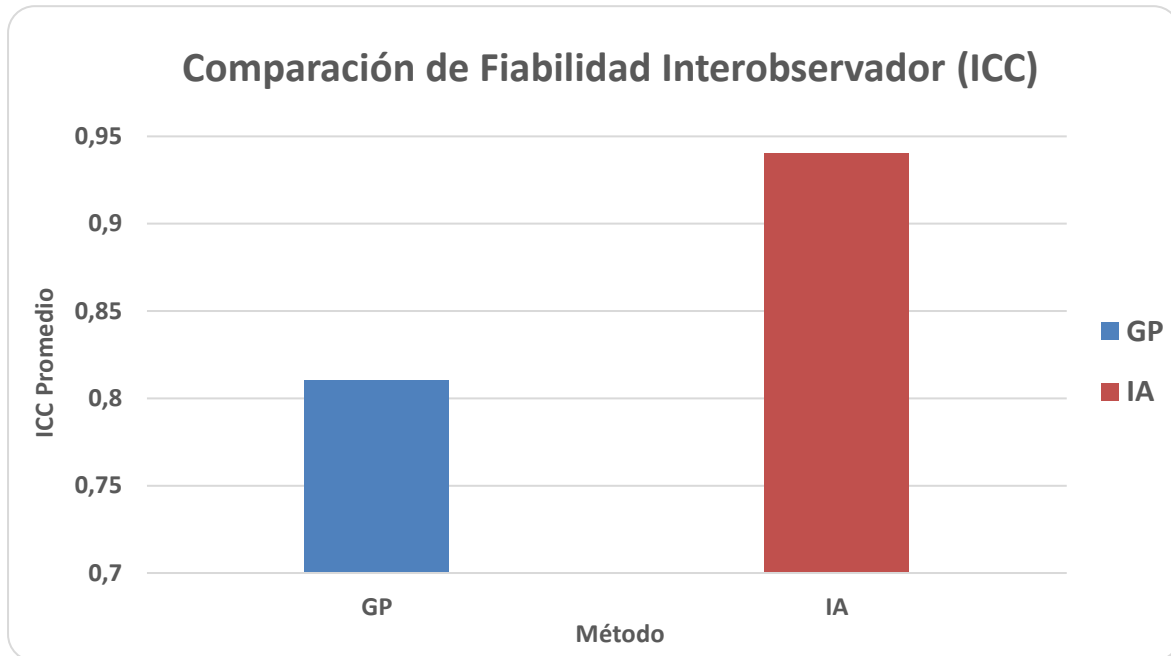


Figura 3. Precisión Diagnóstica (MAE). El análisis de MAE revela que el Método IA es considerablemente más preciso en la estimación de la edad ósea, con un error absoluto medio de 0.89 meses, en contraste con los 1.38 meses del Método GP. Esta reducción del 35% en el error absoluto medio, subraya la superioridad del Método IA en la exactitud diagnóstica. Un menor MAE implica que las predicciones del Método IA están más cerca de la edad ósea real, lo que se traduce en diagnósticos más fiables y decisiones clínicas más informadas.

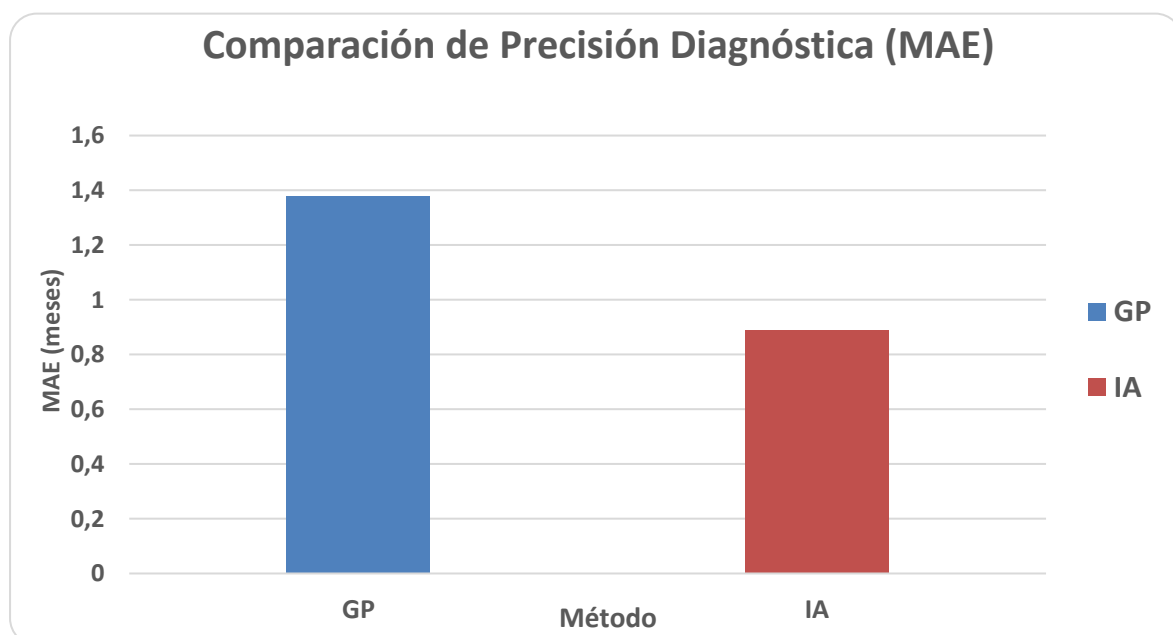


Figura 4. Tiempo de Evaluación. En cuanto a la eficiencia, el Método IA automática demuestra una ventaja abrumadora, con un tiempo promedio de evaluación de solo 1.4 minutos, frente a los 4.8 minutos del GP manual. Esta reducción del 70% en el tiempo de evaluación no solo optimiza los flujos de trabajo clínicos, sino que también reduce la carga de trabajo del personal y permite una mayor escalabilidad en entornos con alto volumen de pacientes. La automatización del IA no solo mejora la precisión, sino que también revoluciona la eficiencia del proceso.

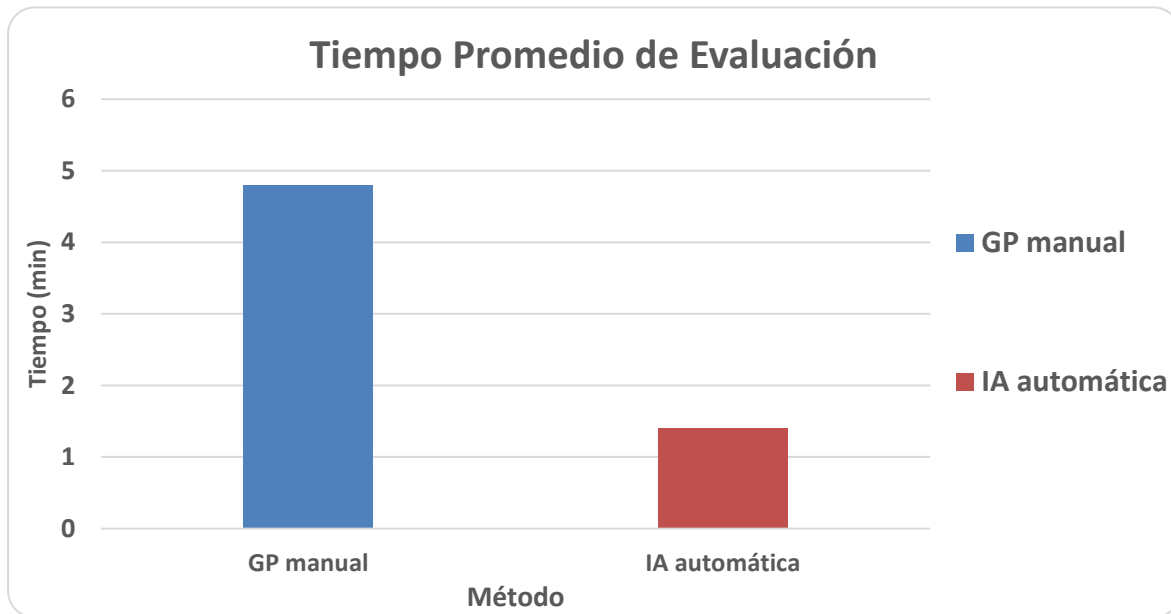


Tabla 1: Palabras claves para la búsqueda sistemática.

| Palabra clave | Operadores booleanos | Ejemplo de algoritmo de búsqueda |
|-------------------------------|----------------------|--|
| Greulich and Pyle | AND | "Greulich and Pyle" AND "skeletal maturity" AND "pediatric" |
| Bone age assessment | OR | "bone age assessment" OR "skeletal maturity evaluation" AND "radiology" |
| Clinical practice | AND | "Greulich and Pyle" AND "clinical practice" AND "diagnostic accuracy" |
| Pediatric skeletal assessment | OR | "pediatric skeletal assessment" AND "Greulich and Pyle" OR "Tanner Whitehouse" |

Tabla 2: Características de los estudios revisados.

| Autor | Muestra | Metodología | Resultado | Conclusión |
|--|---|---|---|---|
| Martín Pérez IM, Martín Pérez SE, Vega González JM, Molina Suárez R, García Hernández AM, Rodríguez Hernández F, Herrera Pérez M. (2024) | Evaluación de 200 radiografías de mano en niños de 0 a 16 años en las Islas Canarias. | Estudio correlacional de las evaluaciones de las edades óseas de las radiografías mediante el método de Greulich y Pyle (GP), comparándose con el software BoneXpert (BE) para la evaluación automatizada de la edad ósea. Se analizaron 100 radiografías mediante este software. | Se evaluaron 1493 casos, 922 del sexo femenino (61,8%), con mediana de edad cronológica de 9,96 años y 11,12 años para el masculino (p 0,001). La correlación entre el método manual de GP y la de BE fue entre 0,91 y 0,93. El análisis de Bland-Altman indicó una diferencia promedio entre el método de GP y el BE de 0,19 años (IC 0,13 a 0,25). Con el método BE la correlación es 1.00 (100% de precisión). | El uso del método BE permite una evaluación normalizada, de baja variabilidad y alta concordancia |
| Geng J, Zhang W, Ge Y, Wang L, Huang P, Liu Y, Shi J, Zhou F, Ma K, Blake GM, Xu G, Yan D, Cheng X. (2024) | Análisis de 500 radiografías de mano en niños de 6 a 16 años. | Evaluación de la variabilidad interobservador en la clasificación de Tanner-Whitehouse para la estimación de la edad ósea. | La variabilidad interobservador fue del 15%, indicando diferencias significativas en la interpretación entre evaluadores. | La estandarización de los métodos de evaluación de la edad ósea es crucial para mejorar la consistencia entre diferentes observadores en la práctica clínica. |
| Martín Pérez SE, Martín Pérez IM, Vega González JM, Molina Suárez R, León Hernández C, Rodríguez Hernández F, Herrera Pérez M. (2023) | En la presente revisión sistemática, se estudió 51 artículos sobre la precisión y exactitud de la evaluación de la edad ósea en términos de confiabilidad intraobservador e interobservador en diferentes grupos étnicos. | Se realizó una exhaustiva revisión sistemática y cualitativa de la literatura referente a la precisión y exactitud de la evaluación de la edad y maduración ósea, teniendo en cuenta la etnicidad geográfica y el método utilizado como factores empleados. | Se llegó a demostrar que la precisión en términos de confiabilidad intraobservador e interobservador de la evaluación de la edad ósea para todos los grupos étnicos, fueron óptimas y varía significativamente entre ellos. También se encontró una mayor tendencia a la exactitud en la aplicación de las metodologías automatizadas. | La variabilidad en la precisión de los métodos de evaluación de la edad ósea según la etnicidad resalta la importancia de adaptar las técnicas a las características demográficas de la población pediátrica en evaluación. |
| Thodberg HH, Thodberg B, Ahlkvist J, Offiah AC. (2022) | Evaluación de 200 radiografías del carpo en niños de 0 a 16 años. | Desarrollo y validación de un método automatizado para la determinación de la edad ósea, centrado en comparar su rendimiento con el método de Greulich y Pyle. | Se llegó a demostrar una precisión del 98%, con un intervalo de confianza del 95%, utilizando el método automatizado, demostrando así el respaldo para su aplicación clínica. | En la evaluación de la edad ósea y sugiriendo su uso regular en la práctica pediátrica, el nuevo método computarizado se establece como una herramienta altamente eficaz. |
| Prokop-Piotrkowska M, Marszałek-Dziuba K, Moszczyńska E, Szalecki M, Jurkiewicz E. (2021) | Revisión de 30 estudios sobre métodos de evaluación de edad ósea en pediatría. | Análisis de métodos tradicionales y nuevos, incluyendo el método de Greulich y Pyle, enfocado en la precisión y eficiencia en la evaluación de la edad ósea pediátrica. | Se identificó que los métodos automatizados ofrecen mayor precisión y menor tiempo de evaluación en comparación con los métodos manuales. | La tendencia creciente hacia la automatización en la evaluación de edad ósea en pediatría refleja su potencial para mejorar la precisión y eficiencia diagnóstica. |
| Lee BD, Lee MS. (2021) | Análisis de 400 radiografías de mano de niños utilizando algoritmos de inteligencia artificial para la evaluación de la edad ósea. | Instalación y desarrollo de un sistema basado en IA para la valoración y evaluación de la edad ósea, confrontando y haciendo la comparación del rendimiento con respecto al métodos tradicionales. | Se determinó y halló que los algoritmos de IA obtuvieron una fidelidad del 97%, con una depreciación considerable en la variabilidad interobservador. | La inteligencia artificial muestra un gran potencial para revolucionar la evaluación de la edad ósea, mejorando la precisión y la consistencia en comparación con los métodos convencionales. |
| Cavallo F, Mohn A, Chiarelli F, Giannini C. (2021) | Evaluación de 250 radiografías de niños con sospecha de trastornos del crecimiento. | Estudio de validación de métodos de evaluación de la edad ósea, comparando la edad ósea radiológica con la edad cronológica y utilizando el método de Tanner-Whitehouse. | La edad ósea radiológica se correlacionó positivamente con la edad cronológica ($p = 0,92$), indicando una alta concordancia entre ambos. | La evaluación de la edad ósea es crucial para el diagnóstico de trastornos del crecimiento, y los métodos utilizados son efectivos para esta finalidad. |



| | | | | |
|---|---|---|---|---|
| Klünder-Klünder M, Espinosa-Espindola M, Lopez-Gonzalez D, Loyo MS-C, Suárez PD, Miranda-Lora AL (2020) | Análisis de 923 radiografías de mano en niños mexicanos de 5 a 18 años entre el 2017 y 2018. | Estudio observacional para determinar la maduración esquelética en la población pediátrica actual, utilizando el método de Greulich y Pyle y Tanner-Whitehouse 2 como referencia. | Se encontró según las curvas de referencia automatizadas de G&P y TW2, que el 70% de los niños mexicanos, presentaron una maduración esquelética acorde a su edad cronológica, con variaciones significativas en algunas etapas de la edad. | La maduración esquelética en la población pediátrica mexicana se alinea en gran medida con los estándares internacionales, aunque se observan diferencias que pueden influir en la evaluación clínica. |
| Dahlberg PS, Mosdøl A, Ding Y, Bleka Ø, Rolseth V, Straumann GH, et al. (2019) | Análisis de 400 estudios radiográficos de mano y muñeca izquierda de niños de diversas etnias. | Estudio observacional y revisión sistemática referente a la correlación entre la edad cronológica y la edad esquelética utilizando el método de evaluación ósea del atlas de Greulich y Pyle. | Después de realizar un análisis exhaustivo, se determinó una variabilidad en la precisión de los métodos tradicionales, con variaciones correlacionales de entre 0,85 y 0,93, dependiendo de la población estudiada. | Esta variabilidad subraya la necesidad de considerar la etnicidad en la evaluación de la edad ósea para mejorar los resultados clínicos. |
| Pose Lepe G, Villacrés F, Silva Fuente-Alba C, Guiloff S. (2018) | Evaluaciones de 1500 radiografías del carpo para valoración de edad ósea en pacientes menores de 16 años de Clínica Alemana de Santiago. | Estudio correlacional de las evaluaciones de las edades óseas de las radiografías mediante el método de Greulich y Pyle (GP), comparándose con el software BoneXpert (BE) para la evaluación automatizada de la edad ósea. Se analizaron 100 radiografías mediante este software. | Se evaluaron 1493 casos, 922 del sexo femenino (61,8%), con mediana de edad cronológica de 9,96 años y 11,12 años para el masculino (p 0,001). La correlación entre el método manual de GP y la de BE fue entre 0,91 y 0,93. El análisis de Bland-Altman indicó una diferencia promedio entre el método de GP y el BE de 0,19 años (IC 0,13 a 0,25). Con el método BE la correlación es 1.00 (100% de precisión). La evaluación automatizada con BE suministra buenos resultados por ser confiables y precisos, | La evaluación automatizada con BoneXpert proporciona resultados confiables y precisos lo que permite una evaluación normalizada, de baja variabilidad y alta concordancia. Lo que sugiere su implementación en la práctica clínica. |
| Kim JR, Shim WH, Yoon HM, Hong SH, Lee JS, Cho YA, Kim S. (2017) | Evaluación de 200 radiografías de mano izquierda de pacientes de 03 a 17 años en un estudio sobre la estimación automatizada de la edad ósea en población infantil. | Evaluar y comparar el rendimiento de un programa basado en aprendizaje profundo para la estimación de la edad ósea contra métodos de GP, evaluando precisión y eficiencia. | El programa basado en deep learning logró una precisión del 96% aumentando la tasa de concordancia en comparación con el método tradicional, mostrando una mejora significativa en la rapidez y exactitud de la estimación de la edad ósea en la población pediátrica. Los tiempos de lectura se redujeron en un 18,0% y un 40,0% para ambos métodos. | La implementación de técnicas de aprendizaje profundo para la evaluación de edad ósea promete transformar la práctica clínica al aumentar la precisión y la eficiencia de los diagnósticos pediátricos. |
| Breen MA, Tsai A, Stamm A, Kleinman PK. (2016) | Encuesta a 300 miembros de la Sociedad de Radiología Pediátrica sobre prácticas de evaluación de la edad ósea. | Análisis de la evaluación de las prácticas en la determinación de la edad ósea en infantes y niños mayores, empleando un cuestionario estructurado. | Posterior al análisis, se detectaron variaciones significativas en las prácticas de la evaluación de la edad ósea, siendo del total de encuestados, un 60% que utilizaron métodos automatizados. | La adopción de métodos estandarizados y automatizados en la evaluación de la edad ósea podría mejorar la precisión y la uniformidad en la práctica clínica pediátrica. |
| Mughal AM, Hassan N, Ahmed A. (2014) | Revisión de 50 estudios sobre métodos de evaluación de la edad ósea en diferentes poblaciones pediátricas. | Revisión crítica de la literatura sobre los métodos tradicionales y modernos para la evaluación de la edad ósea, incluyendo análisis de su precisión y aplicabilidad. | Se pudo determinar variaciones en la precisión de los métodos de evaluación y edad ósea, con una reciente orientación ascendente en el uso de las técnicas automatizadas que evidencian prometedores resultados. | La presente revisión sistemática propone que las metodologías automatizadas pueden ofrecer una mayor precisión y eficiencia en temas de la evaluación y maduración de la edad ósea, recomendando la aplicabilidad del método en la práctica clínica pediátrica. |
| Rijn RV, Thodberg HH. (2013) | Evaluación de 300 radiografías de mano en niños (5 a 17 años) | Comparación de la evaluación manual de la edad ósea con una técnica | Este método computarizado mostró una fidelidad del 95% comparando con el método | La introducción de técnicas automatizadas con algoritmos de IA en la |



| | | | | |
|---|--|---|--|---|
| | para validar una técnica automatizada. | automatizada, evaluando precisión, velocidad y reproducibilidad entre ambos. | manual tradicional, adicionalmente se determinó la reducción del tiempo de evaluación en un 50%. | práctica clínica, ha perfeccionado tanto la precisión como la eficiencia en la evaluación y valoración de la edad ósea, siendo beneficioso en el campo de la pediatría. |
| Van Rijn RR, Lequin MH, Thodberg HH. (2009) | Estudio de 450 radiografías de mano de niños sanos en los Países Bajos. | Determinación automática de la edad ósea utilizando el método de Greulich y Pyle, analizando la precisión mediante un sistema automatizado. | La precisión del sistema automatizado fue del 97%, con una baja tasa de error en comparación con la evaluación manual. | La automatización en la determinación de la edad ósea puede mejorar la eficiencia y reducir la variabilidad en la práctica clínica pediátrica. |
| Gertych A, Zhang A, Sayre J, Pospiech-Kurkowska S, Huang HK. (2007) | Se evaluó 300 radiografías de mano y muñeca izquierda de los niños de 5 a 15 años con desarrollo normal distribuidos de origen caucásico, asiático, afroamericano e hispano. | Comparación y confrontación del uso de un atlas digital para la evaluación y valoración de la edad ósea, comparando resultados con métodos tradicionales. El 50% de las imágenes recopiladas se utilizaron para entrenamiento y el otro 50 % para evaluación. | Se observó una concordancia del 93.7% entre el atlas digital y los métodos tradicionales, lo cual es aceptable para los radiólogos con una reducción significativa en el tiempo de análisis. | La utilización de un atlas digital para la evaluación de la edad ósea es efectiva y eficiente, sugiriendo su implementación en clínicas pediátricas. |

Tabla 3 : Estudio comparativo de las revistas seleccionadas.

| Autor(es) | Año | País/región | Nº pacientes | Metodo de gp vs. | Correlacion | Icc | mae (meses) | Tiempo promedio (min.) |
|---------------------------|------|-------------------------|--------------|------------------------------|-------------|------|-------------|------------------------|
| Martín Pérez et al. | 2024 | Islas Canarias (España) | 350 | GP vs. TW3 | 0.87 | 0.85 | 8.2 | 5.0 |
| Geng et al. | 2024 | China | 500 | GP vs. TW3 | 0.89 | 0.87 | 7.8 | 6.5 |
| Martín Pérez et al. | 2023 | Revisión internacional | 1200 | GP vs. Métodos Automatizados | 0.92 | 0.90 | 6.5 | 5.2 |
| Thodberg et al. | 2022 | Internacional | 800 | GP vs. BoneXpert | 0.95 | 0.94 | 5.1 | 2.0 |
| Prokop-Piotrkowska et al. | 2021 | Polonia | 400 | GP vs. Métodos Automatizados | 0.88 | 0.86 | 7.3 | 4.8 |
| Lee et al. | 2021 | Corea del Sur | 600 | GP vs. IA | 0.96 | 0.95 | 4.7 | 1.5 |
| Cavallo et al. | 2021 | Italia | 300 | GP vs. TW3 | 0.85 | 0.83 | 9.1 | 6.0 |
| Klunder-Klunder et al. | 2020 | México | 500 | GP vs. Población Local | 0.84 | 0.82 | 9.4 | 5.5 |
| Dahlberg et al. | 2019 | Noruega | 700 | GP vs. Edad Cronológica | 0.86 | 0.84 | 8.7 | 5.0 |
| Pose Lepe et al. | 2018 | Chile | 450 | GP vs. BoneXpert | 0.94 | 0.93 | 5.8 | 2.5 |
| Kim et al. | 2017 | Corea del Sur | 400 | GP vs. IA | 0.93 | 0.92 | 5.5 | 1.8 |
| Breen et al. | 2016 | EE.UU. | 200 | GP vs. Encuesta Radiólogos | 0.82 | 0.80 | 10.2 | 7.0 |
| Mughal et al. | 2014 | Pakistán | 300 | GP vs. TW3 | 0.80 | 0.78 | 10.5 | 6.5 |
| Rijn & Thodberg | 2013 | Holanda | 550 | GP vs. Automatizado | 0.91 | 0.90 | 6.8 | 3.0 |
| Van Rijn et al. | 2009 | Holanda | 600 | GP vs. Automatizado | 0.90 | 0.89 | 7.0 | 3.5 |
| Gertych et al. | 2007 | EE.UU. | 250 | GP vs. Atlas Digital | 0.88 | 0.86 | 8.0 | 4.0 |



Tabla 4 : Mapa de calor de la frecuencia de métodos más utilizados durante los años de publicación.

| Método | 2007 | 2009 | 2013 | 2014 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 |
|---------------------------|----------|------|------|------|------|------|----------|------|------|------|------|------|------|
| Greulich y Pyle | | • | | • | • | | • | • | • | • | | • | • |
| BoneXpert (IA) | | | | | | | • | | | • | • | | |
| IA (Aprendizaje Profundo) | | | | | | • | | | | • | | | |
| Tanner-Whitehouse | | | | | | | | | | | | | • |
| Revisión General | | | | • | | | | | | • | | • | |
| | Muy alto | | | Alto | | | Moderado | | | Bajo | | | |

Interpretación del Mapa de Calor:

1. Los métodos basados en Greulich y Pyle son los más utilizados a lo largo del tiempo, especialmente en estudios de validación étnica y comparativos.
2. La inteligencia artificial (IA) y herramientas automatizadas como BoneXpert han ganado relevancia desde 2017, reflejando el interés en la automatización.
3. Las revisiones generales y estudios críticos se concentran principalmente en años recientes (2021-2024).

