

AI Foundation Model as an Educational and Interpretation Aid for Bone Age Assessment

Parisa Kaviani, Noel C. F. Codella, Suzannah E. McKinney, Naiteek Sangani, Kenneth Heckman, Julia Gong, Nishant Yadav, Bernardo C Bizzo,
 Lu Yuan, Keith J. Dreyer, Mannudeep K. Kalra
 Massachusetts General Hospital, Boston, MA, USA
 Microsoft Azure AI, USA



Background

- Assessing skeletal bone age is pivotal for diagnosing various endocrinological, genetic, and growth disorders in pediatric patients.
- Current methodologies, which primarily involve manual evaluation of radiographic images, are not only time-consuming and labor-intensive but also prone to observer variability.
- This has spurred interest in developing more efficient, automated methods utilizing advancements in technology, particularly in the field of deep learning and artificial intelligence.

Goal

We describe the use of an AI foundation model as an educational and interpretation aid for pediatric bone age assessment.

Method

Inclusion criteria:

- Patients with Less or equal 18 years old
- Patients underwent left-hand radiography for bone age assessment between January-December 2018

Final sample size
 (n=202)
 (Mean age 10±4 years)
 (Male: female 100:102)

Ground truth:

Pediatric radiologists' dictated radiology reports using patients' age for bone age (years ± standard deviation).

AI algorithm:

A universal medical image search AI base model (Microsoft Inc.) was utilized to catalog two external datasets:

- The GP index (comprising 59 GP reference images)
- The 12K index (containing 12,611 images from the RSNA training index)

Evaluation:

The AI tool's performance was evaluated based on the precision of the first four images to display the correct bone age.

A Chi-Square test assessed gender-based differences in AI performance.

Results



- ✓ The AI foundation model achieved precision rates of 0.76 and 0.86 for the 12K and GP index tools, respectively (TP: AI output within ± 1SD).
- ✓ Gender-wise (females, model precision 0.75; males, model precision 0.77) performance assessment revealed no significant differences (p-values ranging from 0.229 to 0.870).



The tool allowed image search with the entire radiograph and within a user-selected region of interest.

Conclusions

The AI foundation model as an educational and interpretation aid can help determine pediatric bone age assessment with high precision. Point-of-interpretation similar image search foundational AI model can help in trainee education and interpretation of imaging studies.