

# Development and Accuracy Validation of a Real-Time AI-Assisted Ultrasound Diagnostic System for Developmental Dysplasia of the Hip

Yusuke Ohashi<sup>1</sup>, Tomohiro Shimizu<sup>1</sup>, Hotaka Ishizu<sup>1</sup>, Daisuke Takahashi<sup>1</sup>, Katsuhisa Yamada<sup>1</sup>, Norimasa Iwasaki<sup>1</sup>  
I Department of Orthopedic Surgery, Faculty of Medicine and Graduate School of Medicine, Hokkaido University, Sapporo, Japan  
E-mail: yuk1274go@gmail.com (Y. Ohashi)

**Disclosures:** Y. Ohashi (N), T. Shimizu (N), H. Ishizu(N), D. Takahashi (N), K. Yamada (N), N. Iwasaki (N)

## INTRODUCTION:

Developmental dysplasia of the hip (DDH) is a pediatric condition in which early detection and intervention are critical to prevent long-term deformity and disability. Ultrasound is the global gold standard for infant hip screening; however, diagnostic accuracy is highly dependent on operator expertise in acquiring the correct standard plane and manually measuring the  $\alpha$ -angle,  $\beta$ -angle, and femoral head coverage (FHC). This operator dependency results in variability and limits diagnostic standardization. We hypothesized that an AI-based real-time diagnostic system analyzing ultrasound videos could automatically extract appropriate frames and measure key parameters, thereby reducing operator dependency and improving diagnostic accuracy. The aim of this study was to validate the accuracy of a newly developed AI-assisted real-time diagnostic system compared with manual measurements by orthopedic residents, using pediatric orthopedic surgeons as the reference standard.

## METHODS:

This prospective study included 116 infants (232 hips) who underwent Graf-based hip ultrasound during secondary or tertiary DDH screening at our institution between April 2024 and June 2025. For each case, raw ultrasound video data were collected. The AI system, trained on manually annotated datasets, incorporated bounding-box inference and line approximation algorithms to perform real-time frame-by-frame analysis. The algorithm sequentially (1) detected the iliac outer margin using local maxima and linear approximation, (2) defined a baseline parallel to the iliac cortex through the acetabular beak, and (3) automatically calculated  $\alpha$ -angle,  $\beta$ -angle, and FHC. To ensure diagnostic validity, only frames with baseline deviation within  $\pm 0.5^\circ$  from vertical were included in the analysis. Ground truth values were determined by consensus of experienced pediatric orthopedic surgeons. Manual measurements by orthopedic residents were used for comparison. Primary outcomes were mean absolute errors (MAE) for  $\alpha$ -angle,  $\beta$ -angle, and FHC relative to expert values. Secondary outcomes were Graf classification accuracy, analysis success rate, and failure characteristics.

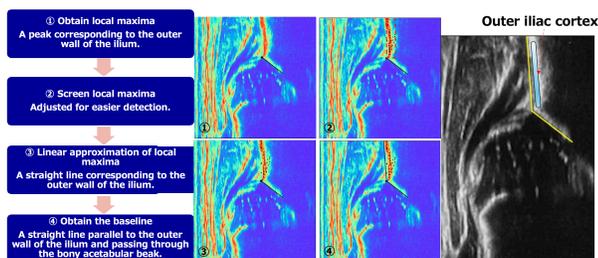
## RESULTS:

Valid diagnostic frames were identified in 199 hips (85.8% of all hips). The AI system successfully completed automatic measurements in 195 hips (98.0% of analyzable cases), while failures occurred exclusively in dislocated or severely subluxated hips where baseline detection was not feasible. The mean absolute errors compared with expert reference values were  $1.05^\circ$  for  $\alpha$ -angle,  $2.71^\circ$  for  $\beta$ -angle, and 2.02% for FHC, all well within clinically acceptable thresholds ( $<3^\circ$  or  $<5\%$  error). The intraclass correlation coefficients (ICC) between AI and expert measurements exceeded 0.90 for all parameters, indicating excellent agreement. Graf classification accuracy was 69.1% for orthopedic residents and 94.0% for the AI system, representing a 24.9-point improvement ( $p < 0.001$ ). Notably, AI maintained high accuracy even in borderline hips ( $\alpha$ -angle  $50\text{--}60^\circ$ ), where resident misclassification was common.

## DISCUSSION:

This study demonstrates that AI-based video analysis can significantly enhance the accuracy and reproducibility of DDH diagnosis. Unlike conventional still image-based systems, the present system integrates frame selection and measurement in a real-time workflow, providing both immediacy and standardization. The approximately 25-point improvement in Graf classification accuracy highlights its potential clinical impact. The newly developed AI-assisted real-time diagnostic system reduces operator dependency and offers highly reproducible diagnostic support for infant hip ultrasound. By standardizing DDH screening, this system holds strong promise to improve the quality and equity of pediatric orthopedic care. Future work should focus on multi-center studies and external validation across different ultrasound devices to establish generalizability and clinical adoption.

**CLINICAL RELEVANCE:** By automating both frame selection and measurement, this AI-assisted real-time system minimizes operator dependency and provides highly reproducible diagnostic support. It has the potential to standardize DDH screening worldwide, reduce regional disparities in pediatric care, and enable earlier and more accurate treatment interventions to preserve hip function and long-term quality of life.



**Fig 1. Automated Detection of the Outer Iliac Cortex in Hip Ultrasound**  
Stepwise process for identifying the outer iliac cortex in ultrasound images:  
① detect local maxima, ② screen maxima, ③ apply linear approximation, and ④ define a baseline parallel to the cortex through the bony acetabular beak. The right panel shows anatomical correspondence in the original image.