Title: A flexible dual-branch AI for ovarian tumor risk assessment using grayscale and Doppler ultrasound: a prospective international multicenter study

Authors: Axel Geysels, Giulia Garofalo, Florian Ramakers, Cristina Taliento, Serena Guerra, Stefan Timmerman, Antonia Carla Testa, Francesca Moro, Francesca Buonomo, Daniela Fischerová, Lil Valentin, Povilas Sladkevicius, Ligita Jokubkiene, Marek Kudla, Caroline Van Holsbeke, Artur Czekierdowski, Elisabeth Epstein, Yvette Groszmann, Bart De Moor, Ben Van Calster, Wouter Froyman, Dirk Timmerman

Objectives: To develop a robust Artificial Intelligence (AI) model that leverages both grayscale and Doppler ultrasound data – and remains effective when either modality is unavailable – for accurate malignancy risk assessment of ovarian tumors.

Methods: We conducted a large, prospective, multicenter study (IOTA7-AI) from 2019 to 2024, enrolling patients with ovarian masses at 11 centers in 6 countries. Ultrasound images were obtained at inclusion, and definitive histology was recorded post-surgery. Images were retrieved either directly from scanners – allowing Doppler to be selectively included or omitted – or from PACS (with Doppler embedded, if present). A dual-branch AI model processing both modalities was developed. Model performance was evaluated via patient-level 5-fold cross-validation, with discrimination measured by the area under the receiver operating characteristic curve (AUC). The ADNEX model (without CA-125), based on ultrasound variables measured at the inclusion scan, served as a clinical benchmark. DeLong's test was used to compare AUCs.

Results: A total of 14,477 ultrasound images from 2,510 patients were analyzed, of which 956 (38%) had malignant histology. The AI model showed strong overall performance with an AUC of 0.922 (95% CI 0.910-0.932), while the ADNEX model achieved an AUC of 0.928 (0.918-0.938; p=0.23). In a paired sub-analysis of 1670 patients who had both Doppler and non-Doppler images, the model achieved an AUC of 0.916 (0.901-0.929) when Doppler data were included, compared to 0.905 (0.891-0.920; p=0.095) for grayscale alone. Notably, performance remained consistent across multiple ultrasound devices.

Conclusions: The dual-branch AI model reliably classifies ovarian tumors by integrating grayscale and Doppler ultrasound – even when one modality is unavailable – and achieves performance comparable to the widely used, externally validated ADNEX model. Its flexibility could benefit diverse clinical workflows, facilitate triage, and support timely referrals to specialized centers.

*The figures on the next page will not be included in the abstract submission due to character limitations.

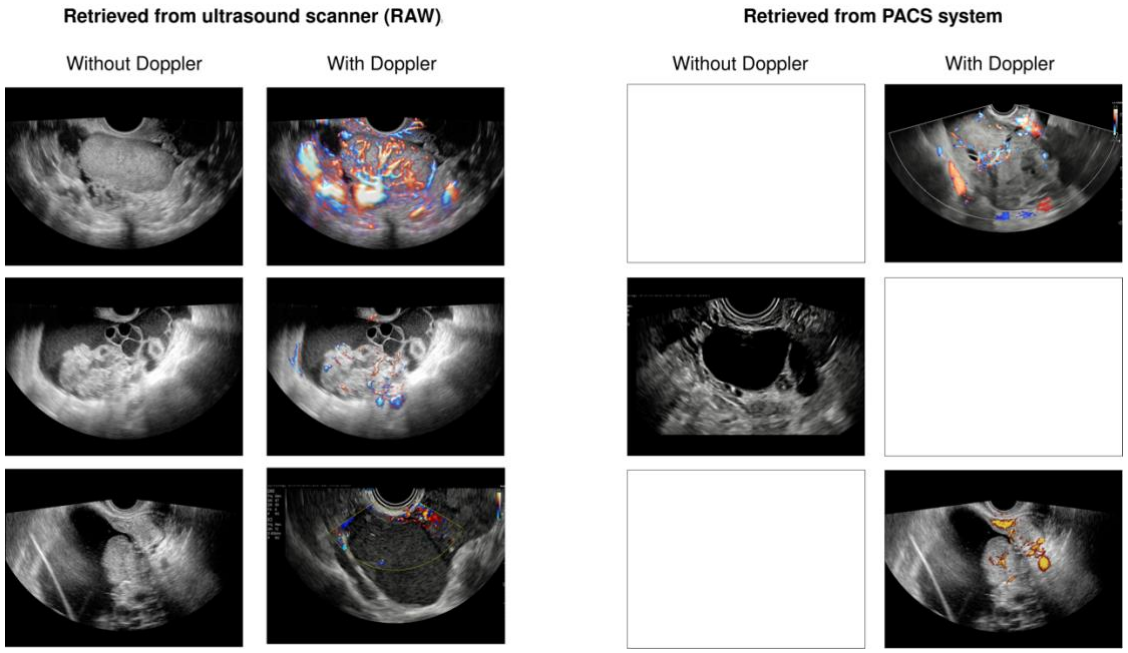


Figure 1*. (Left) Example workflow for images retrieved directly from the ultrasound scanner. If Doppler was present, it could be removed post-hoc, effectively yielding two versions: one with Doppler and one without. (Right) Images obtained from the PACS system, for which only a single version of each scan was available (either with or without Doppler).

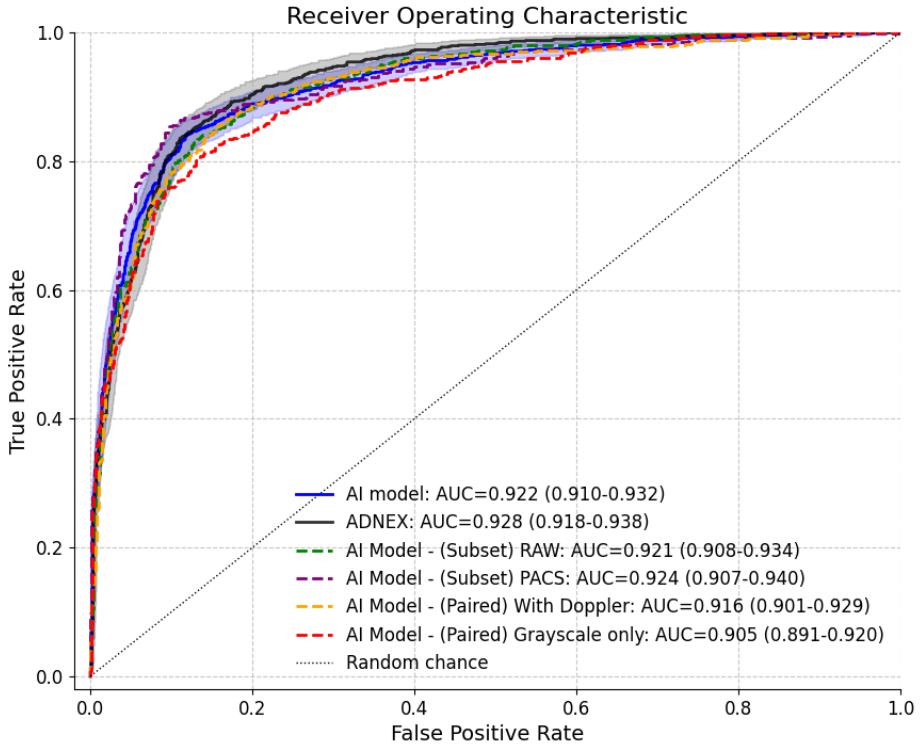


Figure 2*. ROC curves with AUC (95% CI) of the AI model against performance of ADNEX without CA-125. A sub-analysis based on the origin of the images (RAW vs PACS) and based on the presence of Doppler in the model input.