Categorical decision concordance between ENDOSCAN AI and expert clinicians

7-mm endometrial transfer-readiness threshold. Secondary analysis of 2D and 3D ultrasound datasets

Siennicki M, Pawlik P, Zapała A, Sankowska U, Wygocki P, Tralik B, Pankiewicz K, Zak J, Pierzyński P

MIMFERTILITY.

Design & data

1. Multicentre retrospective study
Poland + United States IVF-cycle ultrasound data

2. Parallel imaging datasets
>150 sagittal 2D clips + paired 3D ultrasound dataset

3. 7-mm clinical dichotomisation
Transfer-ready vs not transfer-ready for embryo transfer

4. Agreement benchmark
AI–clinician κ tested against clinician–clinician κ

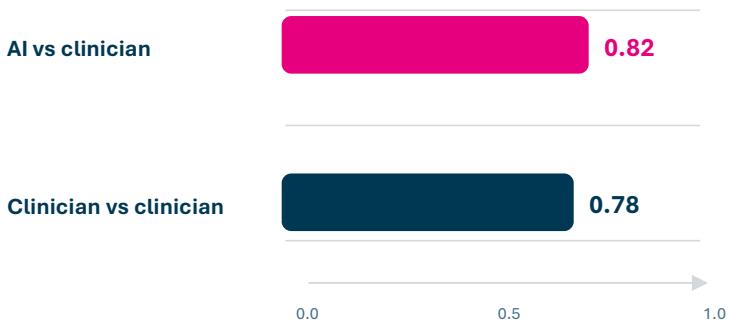
Input	Comparator	Endpoint
2D thickness	Reference clinician	κ at 7 mm
3D volume	Manual VOCAL	ICC + time

7 mm
ET decision threshold

>150
2D sagittal clips

Primary result: categorical agreement

Cohen's κ at the clinically operative 7-mm cutoff



AI classifications were no more divergent from expert judgement than experts were from one another

0.78
benchmark κ

→

0.82
AI–clinician κ

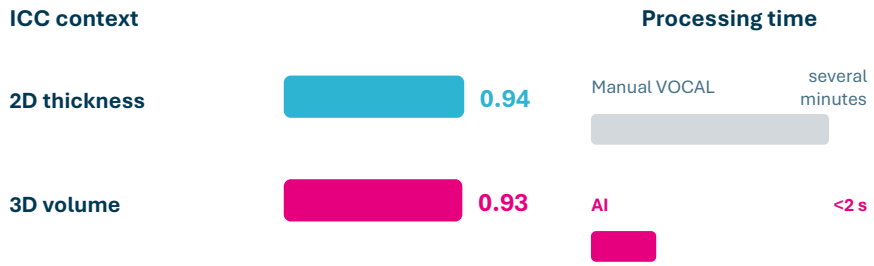
Equivalence criterion met

Continuous-performance context

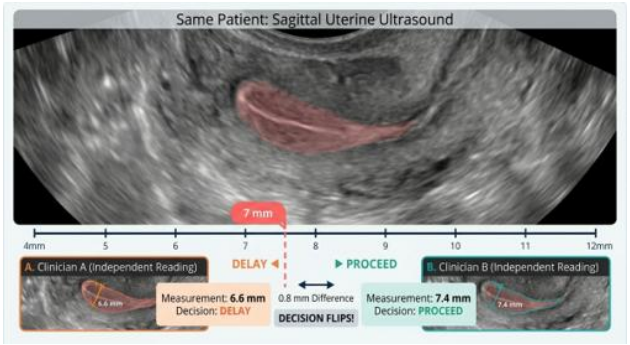
ICC 0.94
2D thickness vs clinicians

MAE 1.1 mm
2D mean absolute error

ICC 0.93
3D volume vs VOCAL



Clinical question	Decision rule	Poster finding
Transfer-ready?	≥7 vs <7 mm	κ=0.82 vs benchmark 0.78
Standardise?	paired same cases	preserved across clinics/platforms



Conclusion Expert-level categorical agreement at the 7-mm FET threshold.

Potential impact Reduced variability in transfer-readiness decisions
Evidence-based optimisation of FET protocols