

Introduction

Successful implantation requires coordination between the myometrium and endometrium. However, current embryo transfer decisions rely predominantly on endometrial thickness (EMT). A recent endometrium AI-model (E-AI), trained on endometrial ultrasound images and clinical features, has improved this standard through objective insights. Given the molecular and physiological crosstalk between the endometrium and myometrium in preparing the uterus for implantation, incorporating both regions in model development holds promise for enhancing AI-driven predictions of receptivity.

Objective

To assess whether an AI-model incorporating myometrium and endometrium ultrasound regions (ME-AI) improves prediction of clinical pregnancy (CP) outcomes in frozen embryo transfer (FET) cycles, compared with E-AI alone.

Methods

Segmentation Models

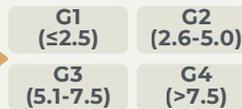


- Clinical details:
- EMT
 - Age
 - Progesterone level



E-AI vs ME-AI
 E-AI: endometrium only
 ME-AI: endo/myometrium

CP likelihood score (0-10)



CP score stratified as intervals

Dataset

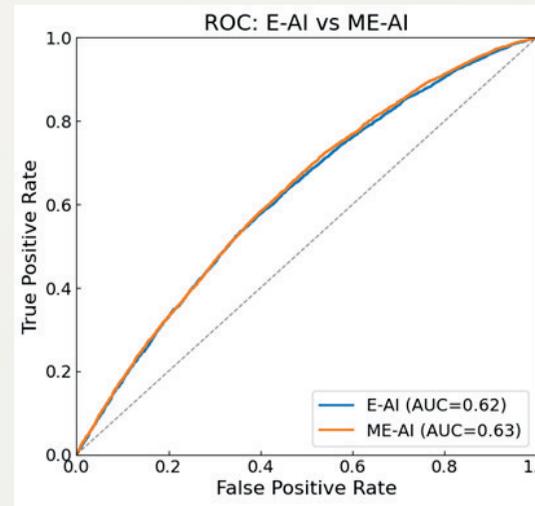
-  **70** US fertility clinics
 -  **48,184** autologous FETs
 -  **29,851** patients (19-57 years)
 -  **65,873** transvaginal ultrasound images
- Images taken before exogenous progesterone or luteal phase (retrospective data from 2018-2022)

Outcomes and Analysis

CP defined as gestational sac/fetal heartbeat at 6-7 week ultrasound.

CP outcomes compared via χ^2 with Holm-adjusted pairwise tests. Continuous variables compared via Welch's t-test. Model performance assessed via AUC, sensitivity, and specificity

Results



		E-AI Confusion Matrix		
True	Negative	4335 (59.0%)	3012 (41.0%)	Specificity 0.59 Sensitivity 0.59
	Positive	2355 (41.2%)	3359 (58.8%)	
		Negative	Positive	
		Predicted		
		ME-AI Confusion Matrix		
True	Negative	4927 (67.1%)	2420 (32.9%)	Specificity 0.67 Sensitivity 0.50
	Positive	2846 (49.8%)	2868 (50.2%)	
		Negative	Positive	
		Predicted		

ME-AI and E-AI predicted CP with comparable performance, each showing significant ($p < 0.001$) stepwise increases in CP rates across intervals G1-G4:

E-AI: G1: 30% → G4: 56%; **ME-AI:** G1: 29% → G4: 55%

The models agreed in 74% of cases. Within this subset, they were both correct in 5,926 cycles, yielding 1,708 live births (73% of CPs). Agreement with outcomes was strongest in older patients (37.2 vs 36.8 years, $p < 0.001$), thinner EMT (9.7 vs 10.0 mm, $p < 0.001$), and autologous oocyte use (89% vs 11% donor, $p < 0.05$).

In the 26% of cases where the models disagreed, E-AI identified a greater proportion of true-positive CP than ME-AI (32% vs 14%, $p < 0.001$), accounting for higher live birth rates (45% vs 19%, $p < 0.001$). By contrast, ME-AI more effectively identified true-negatives (38% vs 17% for E-AI, $p < 0.001$), indicating non-receptive cycles. Progesterone levels were slightly higher in cycles correctly classified by E-AI compared with ME-AI (0.351 vs 0.336, $p < 0.05$), with no other clinical differences identified.

Conclusions

ME-AI and E-AI achieved comparable overall performance. Notably, inclusion of the myometrium enhanced model specificity. While the endometrium remains the dominant predictor, myometrial context may offer selective value. Further investigation is needed to determine whether incorporating the myometrium alongside the endometrium enhances AI receptivity models, particularly in the presence of myometrial abnormalities, thereby improving embryo-transfer decision making.