A NOVEL ARTIFICIAL INTELLIGENCE (AI) MODEL, TRAINED USING 2-DIMENSIONAL ENDOMETRIAL ULTRASOUND (US) IMAGES, PREDICTS CLINICAL PREGNANCY OUTCOMES

Authors: Jullin Fjeldstad[1], Weikai Qi[1], Lais Vanzella[1], Parisa Mojiri Forooshani[1], Natalie Mercuri[1], Alex Krivoi[1], Dan Nayot[1,2,3,4]

Affiliations: [1]Future Fertility, Toronto, ON, Canada; [2]The Fertility Partners, Toronto, ON, Canada [3]The Reproductive Care Centre, Mississauga, ON, Canada [4]Bird&Be Company, Toronto, ON, Canada

Background

Endometrial receptivity (ER) is a key determinant of embryo transfer success in IVF. However, quantifying ER has been challenging and unreliable with previous research displaying inconsistencies in efficacy for both non-invasive (e.g., endometrial thickness) and invasive (e.g., endometrial receptivity assays) methods. A more consistent, objective, and meaningful assessment of ER is required to improve overall chances of success.

Objective

To determine if an AI model, developed to non-invasively assess endometrial receptivity using 2D US images, can predict clinical pregnancy (CP) outcomes.

Materials and Methods

Retrospective dataset of 65,874 transvaginal endometrial US images from 29,851 patients (19-57 years) that underwent 48,184 FET cycles (2018-2022) across 70 clinics of a United States fertility network were utilized. US images were of mid-sagittal view per standard clinical protocol and taken at the last US assessment before administration of exogenous progesterone in medicated cycles, or prior to the luteal phase in natural cycles. FETs included all embryos suitable for transfer. CP outcomes for each transfer were recorded. A positive CP outcome was defined by presence of a gestational sac or heartbeat during a 6–7-week US following a positive β-hCG.

A segmentation model was first developed to create a mask (outline) of the endometrial area (myometrium and endometrium) in the US images, which were then split into 60:20:20 for train, validation, and test sets to develop an ensemble model to predict CP. The ensemble included an image model utilizing US images and known CP outcomes, and a feature model utilizing 3 clinical features and 2 features from the masks generated in the segmentation model (endometrium thickness, age at transfer, progesterone level at time of US assessment, endometrial mask longest:shortest axis diameter ratio and solidity (contour:convex hull areas)).

The ensemble model was applied to the test dataset (n=13,061) to analyze model performance based on AUC, sensitivity, and specificity. The model's probabilities were also converted to a score between 0-10 and further divided into 8 groups (A:0-1.25; B:1.26-2.5; C:2.51-3.75; D:3.76-5.0; E:5.1-6.25; F:6.26-7.5; G:7.51-8.75; H:8.76-10). Two proportions z-test was used to compare true CP rates between the endometrial assessments assigned to the 8 model score groups.

Results

The endometrial receptivity AI model achieved an AUC of 0.62, specificity of 0.59 and sensitivity of 0.59 in predicting CP. A stepwise increase in positive CP rate is observed when the data is divided into the 8 score groups: A: 21.8%; B: 30.9%; C: 38.0%; D: 42.3%; E: 45.9%; F: 50.6%; G: 54.6%; H: 61.3%. Significant differences were observed between positive CP rates of consecutive groups (all p<0.001), except for groups D vs E, which displayed a borderline significant p-value of 0.05 (Figure 1).

Conclusions
The AI model demonstrates favorable performance in predicting CP from US images as a marker of endometrial receptivity. It generates a meaningful score that correlates with higher reproductive success. By utilizing AI to predict CP from US images, clinicians may benefit from more efficient, standardized, and objective insights, aiding in embryo transfer decisions, managing patient expectations, and ultimately improving IVF outcomes.

Support
Employees of Future Fertility.

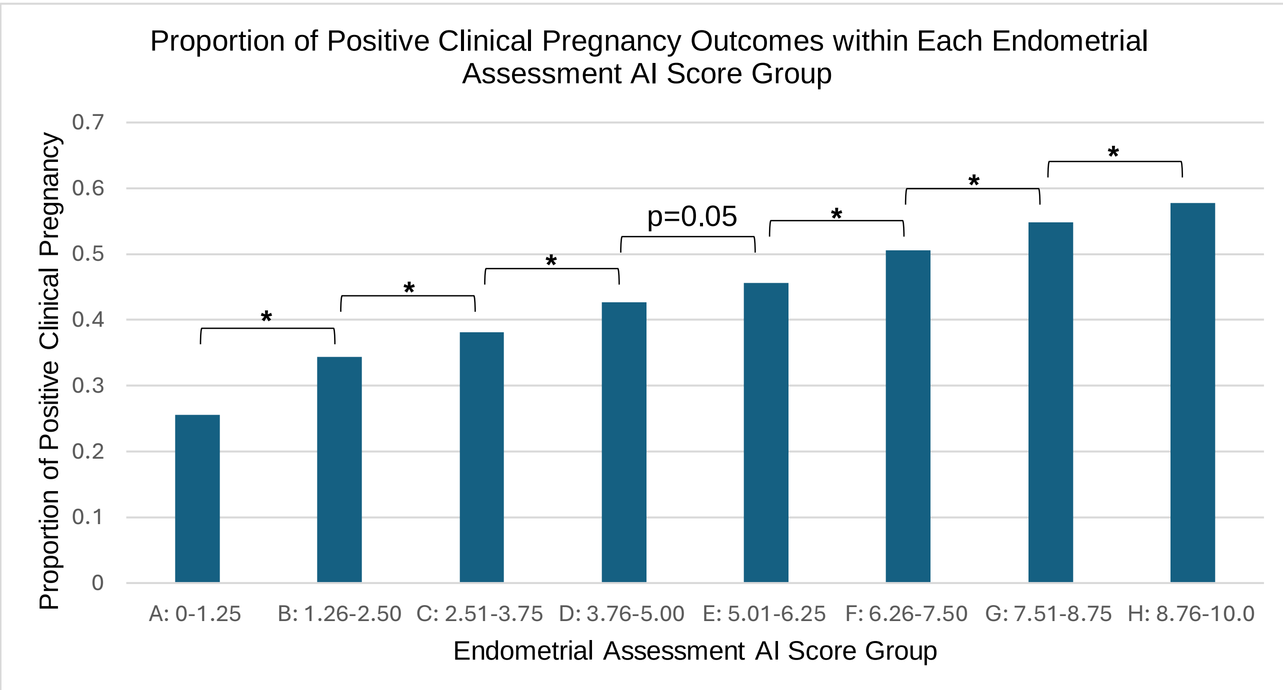

Figure 1. Proportion of positive clinical pregnancy outcomes for endometrial assessments assigned to each score group by the endometrial receptivity AI model. (*p<0.001).