Abstract title

Camera-agnostic self-annotating Artificial Intelligence (AI) system for blastocyst evaluation

Biography

Dr. Matthew (Tex) VerMilyea is the Vice President of IVF Lab Operations at Ovation Fertility and oversees 11 network IVF laboratories. Dr. VerMilyea received his PhD in Epigenetics from The University of Birmingham in the UK and completed a fellowship at the RIKEN institute in Kobe, Japan. After serving as Scientific Director in New Zealand, he returned to Texas as Scientific Director of Texas Fertility Centers in Austin and San Antonio, Texas. Matthew has a strong interest in non-invasive embryo selection including artificial intelligence, time-lapse technologies and spent media analysis.

M. VerMilyea[1,2], J.M.M. Hall[3,4], S. Diakiw[4], A. Johnston[3,4], T. Nguyen[4], M.A. Dakka[4], A. Lim[5], W. Quangkananurug[6], D. Perugini[4], A.P. Murphy[4], M. Perugini[4].
[1]Ovation Fertility, Laboratory, Austin, U.S.A..
[2]Texas Fertility Center, IVF Laboratory, Austin, U.S.A..
[3]Australia/Australian Research Council Centre of Excellence for Nanoscale BioPhotonics, The University of Adelaide, Adelaide, Australia.
[4]Presagen, Life Whisperer, Adelaide, Australia.
[5]Alpha Fertilty Centre, IVF Laboratory, Petaling Jaya, Malaysia.
[6]Safe Fertility Center, IVF Laboratory, Bangkok, Thailand.

Study question:

Can computer vision image annotation techniques be used alongside machine learning to provide reliable blastocyst evaluation that is robust to different camera or microscope types?

Summary answer:

AI that combines automated embryo annotation, trained on optical microscope images alone, generalises to yield high accuracy and consistency for time-lapse derived images.

What is known already:

Recent studies have shown that AI and computer vision can improve embryo selection and accurately predict clinical pregnancy from images of human embryos at a fixed time point (e.g. Day 5). These results are expanded to consider techniques that are robust to camera and microscope type, and objective focal length, including snapshots taken from cameras used in time-lapse incubators. Computer vision detection and segmentation techniques are able to improve the distribution of AI ranking scores, showing consistent accuracy when using EmbryoScope or (preliminary) GERI time-lapse incubator data, with only 2.2% sample deviation of accuracy across six different focal lengths.

Study design, size, duration:

The original Life Whisperer model (VerMilyea et al, 2020), was retrained using extensive image augmentation, 2,530 non-time-lapse incubator microscope images of Day 5 blastocyst embryos, and related clinical pregnancy outcomes, from four US laboratories, two Australian laboratories and one New Zealand laboratory. The AI includes embryo-detection and segmentation to maximize generalizability across different imaging modalities. The AI was applied to

double-blind datasets of optical microscope, EmbryoScope (Malaysia/Thailand) and GERI images from the US.

Participants/materials, setting, methods:

3,470 separate optical microscope, 221 EmbryoScope, and 38 GERI images from patients undergoing fertility treatment at 12 IVF laboratories in five countries were used to train, validate and test the AI accuracy, distribution, robustness to camera/microscope type, and objective focal length. Only images of Day 5 blastocysts for which pregnancy (heartbeat at first scan) outcome was known, were used. This study was determined exempt from IRB review by Sterling IRB, USA (#6467).

Main results and the role of chance:

This is the first study to show that AI trained on standard Day 5 microscope images can generalize to time-lapse incubator images, demonstrating the robust and camera agnostic nature of this approach.

The AI accuracy for prediction of clinical pregnancy (fetal heartbeat) was 65.4% when averaged over a blind test set of Day 5 blastocyst images from 10 IVF clinics in four countries. The sensitivity of the AI was 86.2%, while the specificity varied depending on the curation of the dataset. New blinded datasets of Day 5 single images of blastocyst stage embryos from the EmbryoScope and GERI time-lapse systems were then assessed with the AI. The AI generalized well to the time-lapse derived images and consistent overall accuracy (59.1%) was achieved, with a sensitivity of 77.9%. Multiple focal lengths were also considered and showed only a 2.2% deviation of accuracy. Together these results suggest that this method of pre-processing and automated annotation, as well as AI trained on a globally diverse dataset, creates a generalizable AI that is robust to camera type and focal setting.

Limitations, reasons for caution:

The GERI data set is small and therefore analysis of the distribution of scores from this set should be expanded. Additional consideration of camera effects and focal lengths from both EmbryoScope and GERI devices across a wider range of clinics should also be considered.

Wider implications of the findings:

Constructing AI that is robust to image variation represents an advancement in the computer vision field. Applying these techniques to blastocyst-stage embryos demonstrates that AI can be robust and generalizable to different clinical environments. This suggests that AI in the clinical embryology setting is practical and scalable, regardless of hardware.