Abstract citation ID: dead093.147

O-120 Predicting Embryo Ploidy Status Using Time-lapse Images

S. Rajendran[1], M. Brendel[1], J. Barnes[1], Q. Zhan[2], J. Malmsten[2], Z. Rosenwaks[2], M. Meseguer[3], N. Zaninovic[2], O. Elemento[1], I. Hajirasouliha[1]

[1]Weill Cornell Medicine, Department of Physiology & Biophysics, New York City, U.S.A.
[2]Weill Cornell Medicine, The Ronald O. Perelman and Claudia Cohen Center for Reproductive Medicine, New York City, U.S.A.
[3]Health Research Institute la Fe, IVI Valencia, Valencia, Spain

Study question: Can deep learning models using time-lapse images of embryo development be used to predict embryo ploidy and provide supplemental information to embryologists for clinical decision-making?

Summary answer: We developed a general MDBS-Ploidy model that uses time-lapse images and maternal age to predict embryo quality scores and ploidy status.

What is known already: Ploidy status, or the presence or lack of chromosomal abnormalities, is an important factor in a successful pregnancy. Embryos with abnormalities are classified as aneuploid, whereas those without are euploid. With the advent of artificial intelligence in computer vision and the collection of large IVF-related datasets combining images, videos, and clinical outcomes, a variety of methods have been developed to automatically assess embryo quality and other characteristics using images from time-lapse sequences. To our knowledge, a video classification model validated by heterogeneous data to predict both embryo quality score and ploidy is lacking.

Study design, size, duration: The training dataset for the model consisted of 1,998 time-lapse sequences captured by the Embryoscope®. Time-lapse image sequences consisted of around 360–420 frames captured during 5 days of development. PGT-A results were used as ground truth labels for all ploidy prediction tasks, with embryos classified as euploid or aneuploid. The dataset also included clinical information such as blastocyst score (BS) and maternal age at the time of oocyte retrieval.

Participants/materials, setting, methods: The MDBS-Ploidy consists of two steps. The first step is quality score prediction from day-5 time-lapse video input using a Bidirectional Long Short-Term Memory architecture with added output layers for multitasking. In the second step, the predicted score, in addition to maternal age, is used to predict the ploidy status of the embryo using a logistic regression model. We evaluated the performance of the MDBS-Ploidy using area under the receiver-operating-characteristic (AUROC) on a validation dataset.

Main results and the role of chance: For the MDBS-Ploidy quality score prediction module, the Pearson correlation between scores predicted by the MDBS-Ploidy and ground truth quality scores is 0.70 on the validation dataset, suggesting moderate correlation strength. As for the aneuploidy prediction module, the MDBS-Ploidy can discriminate between euploidy and aneuploidy with an AUROC of 0.76 $\pm$ 0.002. The MDBS-Ploidy performs comparably with a model trained on the embryologist-annotated blastocyst score. We replicate these comparative results in external validation datasets as well, namely the Embryoscope+® dataset with 1,000 embryos and the IVI Valencia Spain dataset with 543 embryos. Moreover, the MDBS-Ploidy is completely automated, requiring only time-lapse images from 96 to 112 hpi and maternal age to predict embryo ploidy status. This allows the MDBS-Ploidy to be adapted clinically without interrupting ongoing workflows. The MDBS-Ploidy also provides a certain level of explainability; embryologists can use the intermediate quality score (from the first module) to determine why an embryo is classified as a certain ploidy status. With a recall of 0.84 $\pm$ 0.004 for the validation dataset, the MDBS-Ploidy shows promise for successfully selecting euploid embryos. We believe that the MDBS-Ploidy can provide supplemental information for selecting the most viable embryo.

Limitations, reasons for caution: Our model primarily uses data from time-lapse microscopy. Clinics without access to this technology will be unable to use our model.

Wider implications of the findings: This model is clinically relevant for making decisions as to whether an embryo may be chromosomally normal. Automated quality score prediction is instrumental to embryologists who are

currently annotating embryos manually. Future iterations of the model can theoretically be adopted into clinical practice because it is end-to-end and fully automated.

Trial registration number: not applicable