Abstract title: 12/25

Using Artificial Intelligence (AI) and Time-Lapse to improve human blastocyst morphology evaluation

Study question: 25/25

Following from a bovine study presented at ESHRE 2017, we compared grading capabilities of human blastocysts using AI versus grading by five distinct experienced embryologists.

Summary answer: 16/25

As observed in the bovine model, AI outperformed embryologists when grading human blastocysts using time-lapse images.

What is known already: 100/100

Morphological grading of blastocysts is clinically used for embryo selection. Our group demonstrated that blastocyst grading in bovine and human embryos by embryolgists lead to wide inter- and intra- operator variation. We demonstrated using bovine blastocysts that an image analysis AI system can reduce variation in blastocyst grading and acquire additional parameters not detectable by operator assessment. Human blastocysts present additional challenges for AI image recognition compared to bovine, given the lower contrast of the inner cell mass. To date, AI technology, has not been demonstrated to improve operator human blastocyst grading, which was the focus of the current study.

Study design, size, duration: 75/75

394 human embryo time-lapse images taken at 110hpi were graded for Inner Cell Mass (ICM), Trophectoderm and Expansion using Gardner Grading by 5 different embryologists from 4 different countries. Of these, 171 were excluded as blastocysts that were too early (expansion 1 or earlier), hatched, out of focus or ICM not clearly visible. The mode of the remaining 223 images were used as output for the AI system (70% training, 15% validation, 15% blind assessment).

Participants/materials, setting, methods: 60/75

29 independent mathematical variables were extracted from the time-lapse images taken at precisely 110hpi from the central Z-stack and inputed into the AI system. The agreement was assessed using confusion matrices, ROC curves and Kappa Index. Embryologists originally trained in the same lab were compared to embryologists trained in different labs to assess inter and intra clinic variation in agreement.

Main results and the role of chance 125/200

Agreement between the 5 embryologists was low (Kappa agreement decreasing from Expansion 0.4, to trophectoderm 0.3 to ICM 0.3). There was no difference between the kappa agreement of embryologists trained in the same clinic to embryologists from different

countries. The low inter-operator agreement is likely to be due to the fixed time central stack image selection. Previously, we demonstrated that where images were selected according to focus of ICM, the operator agreement was considerably higher. Improved agreement was observed using AI to predict the mode of the embryologists with substancial agreement with Expansion (Kappa agreement 0.7) and ICM (0.7) and moderate agreement with trophectoderm (0.4). The AI's overall accuracy was almost perfect for prediction of blastocyst expansion (training 93.9% and blind validation 81.5%) and substancial for prediction of ICM (training 93% and validation 78.8%) and trophectoderm (78.8% training and 78.3% validation). The AI system was considerably more predictive of Expansion (AUC 0.888-0.956) compared to ICM (AUC 0.605-0.854) and trophectoderm (AUC 0.726-0.769).

Limitations, reasons for caution: 49/50

Our data suggests the AI is able to cope better than operators with the challenge of grading three dimensional embryos from a single fixed two-dimensional image. This technology has now been demonstrated in two independent centers. Further independent studies are required to demonstrate reproducibility before establishing its clinical application.

Wider implications of the findings: 49/50

Applying AI to human blastocyst grading is inexpensive, non-invasive, and more reliable than grading by an operator. Instead of a human looking at thousands of images, AI assesses them and continuously learns and quantifies additional information. As demonstrated, this technology can inherently enhance our capabilities of assessing embryo viability.

Trial registration number: 25