

## Multicentre validation of CHLOE-EQ: An embryo assessment assistant based on Artificial Intelligence (AI)

List the authors: Reprofert; Reprofert; Fertility; Fertility; Fertilitat; Fertilitat; Primordia; Primordia; Moura, B.; Zepeda, A.; Brualla, A.; Hickman, C. 1 Reprofert, Brazil 2 Fertility FIV, Brazil 3 Fertilitat, Brazil 4 Primordia, Brazil 5 Embriologica, Brazil 6 Fairtility, Israel

### Objective:

CHLOE-EQ is an embryo assessment assistant that automatically processes time-lapse videos using AI with the objective to increase consistency, efficacy of prediction whilst saving valuable embryologist time. The purpose of this study was to compare the assessment of embryoscope time-lapse videos by experienced embryologists with CHLOE-EQ (Fairtility) across four independent clinics.

### Methods:

Following culture of embryos in a time-lapse incubator (Embryoscope, Vitrolife) at four clinics (Clinic A N=147, Clinic B N=40, Clinic C N=143, Clinic D N=462); experienced embryologists prospectively assessed the number of pronucleates, morphokinetics, inner cell mass (ICM) and trophoderm quality and determined which embryos should be utilised or discarded as per routine clinical practice. The same time-lapse videos were retrospectively assessed by CHLOE-EQ (Fairtility), blind to the human assessments. Intra-Correlation Coefficient (ICC) was used to quantify the level of agreement between Embryologist and CHLOE for morphokinetics: Very weak (0-0.2), weak (0.2-0.4), moderate (0.4-0.6), strong (0.6-0.8), very strong (0.8-1). Agreement of PN assessment by CHLOE and Embryologists was assessed using Kappa score. Efficacy of prediction of blastulation, utilisation, selection for transfer and ploidy was assessed against CHLOE-Blast Score and CHLOE EQ Score and CHLOE RANK using Binary logistic regression. Each of the assessments was analysed per clinic, and overall across all five clinics.

### Results:

As outlined in the table below, when combining the data from all clinics, all morphokinetics had a very strong agreement between CHLOE and embryologist annotations. At the individual clinic level, the lowest level of agreement was moderate for tPNa in clinic A and strong for t4 in clinic A; all other clinics had a very strong level of agreement between embryologist and CHLOE for all remaining morphokinetics.

The overall accuracy of PN assessment was 96%, with a kappa agreement 0.87 (very strong). Across all clinics, CHLOE BLAST Score was predictive of blastulation (AUC=0.77-0.99,  $p < 0.001$ ), whilst CHLOE EQ Score was predictive of utilisation (AUC 0.81-0.91), selection for transfer (AUC 0.70-0.85), euploidy (AUC=0.65-0.75) and CHLOE Ranking was predictive of utilisation (AUC=0.70-0.86) and selection for transfer (AUC=0.85-0.86).

### Conclusion:

CHLOE-EQ can automatically annotate morphokinetics, count pronucleates and identify blastulation with a strong level of agreement with experienced embryologists across different clinics, bringing a consistent language of embryo assessment that can be generalised to different clinics using time-lapse incubation. Incorporating AI based tools such as CHLOE in a time-lapse clinics can help improve consistency in embryo assessment, efficacy of prediction of embryo viability whilst saving valuable embryologist time.

Keywords: Artificial Intelligence; Embryo Selection; Morphokinetics; Embryo Quality; CHLOE.