

Embryologist versus machine

Measuring efficiency with AI for embryo evaluation

A. Papatheodorou¹, Y. Paserman², C. Curchoe², R. Maor², D.S. Seidman^{2,3}, D. Gilboa²

Study question

Does the adoption of automated embryo evaluation with EMA™ [AIVF, Israel] improve efficiency in an IVF setting?

Summary answer

Embryo evaluation using the EMA AI model was 45 times faster than evaluation done by an embryologist.

What is known already

Embryo grading and annotation is a time-consuming part of the embryologist day, which could be replaced by AI. Automated AI-driven embryo evaluation systems have been validated clinically for accuracy.

Inter-laboratory variability during embryo assessment is a common occurrence resulting in various reasons.

The results are:

- Lower laboratory efficiency
- Risk of manual data transfer errors
- Occupational stress/fatigue
- Embryologist burnout
- Inter-laboratory/inter-clinic communication failure
- Unnecessary repetitive workflows



Hunger



Fatigue



Training



Stress



Personality



Weather



Previous experiences



Mood

Study design, size, duration

We developed a simulator to accurately quantify embryologist's embryo evaluation and annotation timing and pitted it against EMA AI platform [AIVF, Israel], that records a continuous score for each embryo. We calculated each embryologist's performance time for embryo evaluation with and without EMA. A total of 526 embryos from 49 cycles were included. This analysis was restricted to cycles where ≥ 2 good-quality blastocysts were eligible for evaluation.

Participants/materials, setting, methods

EMA™ by AIVF is an AI software platform that assesses embryo quality and developmental competency using a set of deep-learning-based algorithms. The average time spent by three senior embryologists to evaluate embryo quality and record their decisions [transfer/freeze/discard] was measured using published consensus guidelines versus EMA.

¹ Embryolab Fertility Clinic, IVF Lab, Thessaloniki, Greece
² AIVF, Tel Aviv, Israel
³ Sackler School of Medicine Tel-Aviv University, Tel Aviv, Israel

Main results and the role of chance:

The efficiency of EMA under typical clinical patient workload was assessed. Embryo evaluation done by an embryologist consists of 3 phases: manual grading, recording and decision making. Embryo grading done by an embryologist took 190 seconds per embryo. Recording time took 10 seconds per embryo and decision per cohort took 301 seconds [for 10 embryos]. Using EMA grading time was instantaneous, recording was instantaneous [both done by AI] and decision by an embryologist took 51 seconds per cohort. Overall, embryo evaluation time was reduced by 97.8%.

	Conventional Embryo Evaluation by Embryologist Alone [seconds]	Embryo Evaluation by Embryologist + EMA [seconds]
Embryo grading and annotation time [per embryo]	190	0
Recording time [per embryo]	10	0
Decision-making time by embryologist [per cohort*]	301	51
Total embryo evaluation time [per cohort]	2,302 seconds	51 seconds

*10 embryos per cohort

Limitations, reasons for caution:

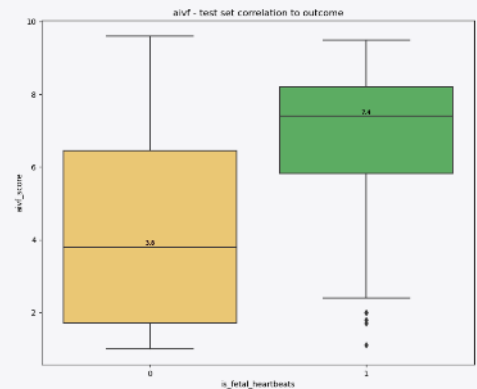
This is a single center study. It should be repeated over wide geographical locations and patient populations.

Wider implications of the findings:

The results from this study show that clinic efficiency can be greatly improved using EMA, as the time-consuming process of manual embryo evaluation can be replaced by EMA. The AI-based automated evaluation and ranking also markedly shortens the decision time per cohort.

Distribution of EMA AI Scores Relative to Pregnancy Outcome

- EMA score distributions are consistently higher for embryos that result in clinical pregnancy when compared to those that result in nonpregnancy with appropriately minimized overlapping of the interquartile range [p<0.05].
- In this way, EMA scores are robustly associated with clinical outcome [nonpregnancy/pregnancy]



Distribution of EMA Scores Relative to ASEBIR Grading Scale

- Boxplot diagram evaluating the association of EMA with embryo quality [listed E-A; ascending order]. Embryo quality is based on the ASEBIR grading scale.
- There is a significant positive linear correlation between scores and embryologist-annotated embryo qualities as well as strong ability to differentiate between discrete qualities with high classification.

