EMA™ AI-Powered Embryo Evaluation for Superior Decision-Making In the IVF Clinic: **Performance Overview**

Executive Summary

EMA™ (AIVF, Israel) artificial intelligence (AI) powered embryo evaluation was developed through rigorous scientific research, tailored to the needs of the embryology laboratory.

Its AI embryo evaluation scoring system provides quantitative assessments for each embryo that reflect its quality and developmental competence.

EMA's AI is intended to enhance decision-making in the IVF clinic by optimizing embryo evaluation workflow and efficiency.

The result: data-driven decision-support that optimizes treatment efficiency, efficacy, and quality of care.

Introduction

Since its inception, in-vitro fertilization (IVF) technology has resulted in over 10 million births[1]. Though initially fraught with low pregnancy rates due to limited treatment protocols and technological inefficiencies, current IVF success rates have nearly doubled in the last 30 years[2,3]. Embryo evaluation in IVF involves the process of manually evaluating embryos under the microscope and selecting the highest-quality embryo for implantation.

Embryo quality is assessed by embryologists according to standard classification systems that assign a grade and/or number to each embryo based on predefined categories of embryo quality (example: degree of blastocyst expansion, morphological appearance of inner cell mass [ICM] and trophectoderm [TE] cells)[4]. These grading systems have also been shown to correlate with pregnancy rates[5,6].

Although manually grading embryos is considered the gold standard for embryo quality assessments, this method is time-consuming and fraught with intra- and inter-observer variability[7-9]. Furthermore, embryo classification systems discretize a continuous scale of embryo quality[10].
Characterizing embryo quality by discrete grades limits the ability of the embryologist to assess morphologically-similar looking embryos by their chances of pregnancy success. There is a need for objective, data-driven tools to improve embryo evaluation and pregnancy success rates.

EMA intends to overcome these challenges in the field. EMA's automated AI-driven embryo evaluation model determines a continuous score (ranging from 1-9.9) for each embryo according to its quality and potential for pregnancy success.

Performance Overview

EMA's embryo evaluation model is trained and validated on a large, demographically diverse database of hundreds of thousands of embryos and pregnancy outcomes, in collaboration with leading healthcare institutions across the US, Europe and Asia.

EMA undergoes rigorous performance evaluation to assess its clinical utility prior to use. Strong emphasis is placed on preclinical validation of the model architecture and data to verify its unbiased performance across all clinic datasets. This representative study (N=18,700) evaluates how EMA's AI embryo evaluation scores correlate with embryo quality, ploidy, and likelihood of achieving clinical pregnancy to objectively and robustly optimize clinical decision-making in the IVF clinic.

EMA scores are positively correlated with embryo morphological assessments carried out by embryologists

EMA scores demonstrate precise ability to discriminate between top (A grade), medium (B grade), and fair-quality (C grade) embryos with high classification accuracy, as expected. EMA score values increase according to ascending embryo morphology grades, as assessed by the embryologist. Importantly, EMA's continuous scoring scale (ranging from 1-9.9) correlates linearly with the conventional quality scales for each of the top three morphological parameters known to affect overall embryo quality: inner cell mass quality, trophectoderm quality, and embryo development stage ($p<0.05$) **(Fig 1)**. We thereby display robust, consistent ability of the AI to reflect on known parameters of morphological embryo quality.

The linear increase in average EMA scores relative to the qualities for each known morphology parameter demonstrates EMA's ability to reflect an objective representation of embryo quality without being boundaried by predefined embryo morphology grades. In this way, EMA's embryo evaluation AI overcomes the limitations of the discretized embryo scoring scale used by embryologists. EMA's optimized embryo assessment method shows significant and repeatable correlations between its score distribution and biologically meaningful parameters of embryo quality for objective, not relative, scoring of all embryos. This highlights the model's biological explainability and clinical relevance inside the IVF clinic.

EMA scores accurately reflect pregnancy outcomes in the IVF clinic

Analysis of EMA scores and clinical pregnancy rate demonstrates a significant linear correlation between EMA scores and the percentage of embryos that result in clinical pregnancy **(Fig 2)**. To show this, we evaluate the proportion of embryos resulting in pregnancies relative to defined AI score categories ('score bins'). This linearity reflects EMA's ability to apply its learned features robustly across all embryo cohorts to quantitatively assess likelihood of achieving clinical pregnancy. Similarly, average EMA scores are consistently higher for embryos that result in clinical pregnancy when compared to embryos that result in nonpregnancy ($p<0.05$). In this way, there is an association between EMA scores and clinical outcome (nonpregnancy/pregnancy) in the IVF clinic.

To assess EMA's ability to improve embryologists' decision-making inside the IVF clinic, we shift our analysis to assess EMA's performance on high-quality, morphologically-similar looking embryos. This embryo subset complicates conventional embryo evaluation processes, since quality differences between A grade embryos that may influence pregnancy potential are more difficult to distinguish by the embryologist's naked eye alone. Average EMA scores are consistently higher for A grade embryos that result in clinical pregnancy when compared to those that result in nonpregnancy ($p<0.05$) across all evaluation datasets.

EMA scores accurately reflect pregnancy outcomes in the IVF clinic (Cont.)

EMA's ability to differentiate between high-quality embryos by their clinical outcomes reflects its ability to improve upon conventional embryo evaluation processes, especially for the subset of top-quality embryos that are eligible for transfer.

Further analysis shows that EMA demonstrates 70% probability for a successful pregnancy[11] across all clinical datasets and is up to 38% more accurate than an embryologist alone[12]. Independent clinic holdout testing during model validation shows comparable performance across all datasets; EMA robustly generalizes to all clinic datasets without influence from varying demographics or clinic characteristics.

Conclusions

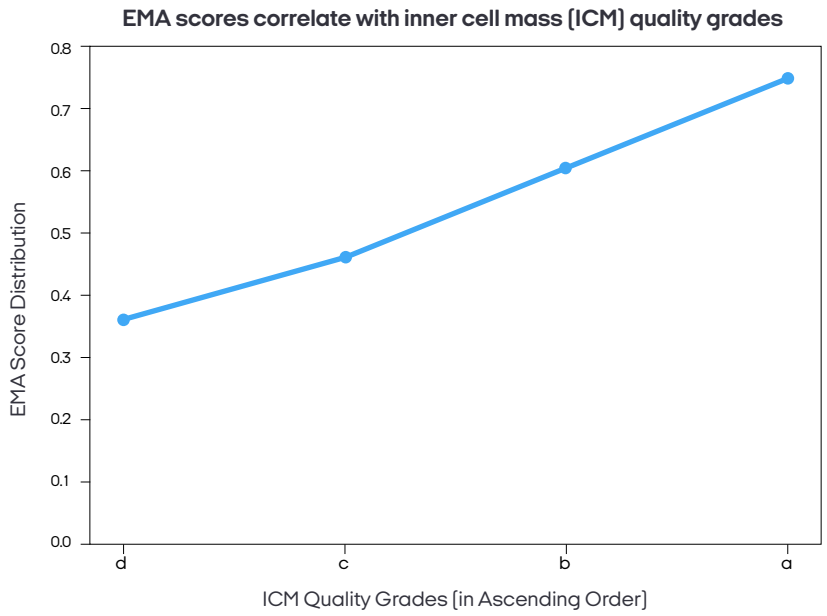
EMA's AI-powered embryo evaluation is a fully automated, objective deep learning model that has the potential to significantly optimize embryo evaluation workflow in the IVF clinic without compromising on clinic confidence, transparency, or outcomes. The model's strong performance highlights its ability to apply its learned features objectively and robustly to optimize decision-making inside the IVF clinic.

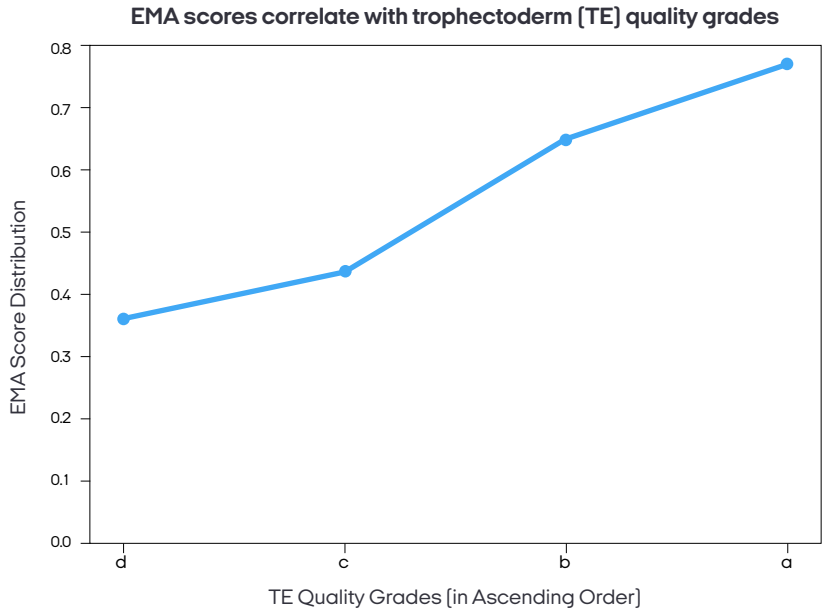
References

1. https://www.eshre.eu/Europe/Factsheets-and-infographics
2. Lee AM, Connell MT, Csokmay JM, Styer AK. Elective single embryo transfer- the power of one. Contracept Reprod Med. 2016;1(1):1-7. doi:10.1186/s40834-016-0023-4
3. Wang J, Sauer M V. In vitro fertilization (IVF): A review of 3 decades of clinical innovation and technological advancement. Ther Clin Risk Manag. 2006;2(4):355-364. doi:10.2147/tcrm.2006.2.4.355
4. Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Hum Reprod. 2011;26(6):1270-1283. doi:10.1093/humrep/der037
5. Nazem TG, Sekhon L, Lee JA, et al. The correlation between morphology and implantation of euploid human blastocysts. Reprod Biomed Online. 2019;38(2):169-176. doi:10.1016/j.rbmo.2018.10.007
6. Lou, H., Li, N., Guan, Y. et al. Association between morphologic grading and implantation rate of Euploid blastocyst. J Ovarian Res 14, 18 (2021). https://doi.org/10.1186/s13048-021-00770-8
7. Storr A, Venetis CA, Cooke S, Kilani S, Ledger W. Inter-observer and intra-observer agreement between embryologists during selection of a single Day 5 embryo for transfer: a multicenter study. Hum Reprod. 2017;32(2):307-314. doi:10.1093/humrep/dew330
8. Sundvall L, Ingerslev HJ, Breth Knudsen U, Kirkegaard K. Inter- and intra-observer variability of time-lapse annotations. Hum Reprod. 2013;28(12):3215-3221. doi:10.1093/humrep/det366
9. Adolfsson E, Andershed AN. Morphology vs morphokinetics: a retrospective comparison of inter-observer and intra-observer agreement between embryologists on blastocysts with known implantation outcome. JBRA Assist Reprod. 2018;22(3):228-237. Published 2018 Sep 1. doi:10.5935/1518-0557.20180042
10. Kragh MF, Rimestad J, Berntsen J, Karstoft H. Automatic grading of human blastocysts from time-lapse imaging. Comput Biol Med. 2019;115:103494.
11. Presented by Marcos Meseguer, IVIRMA Global conference, Malaga 2023
12. IVIRMA-AIVF validated studies

Figures

Figure 1.
Method evaluating the correlation between EMA scores and inner cell mass (ICM) quality, trophectoderm (TE) quality, and development stage.





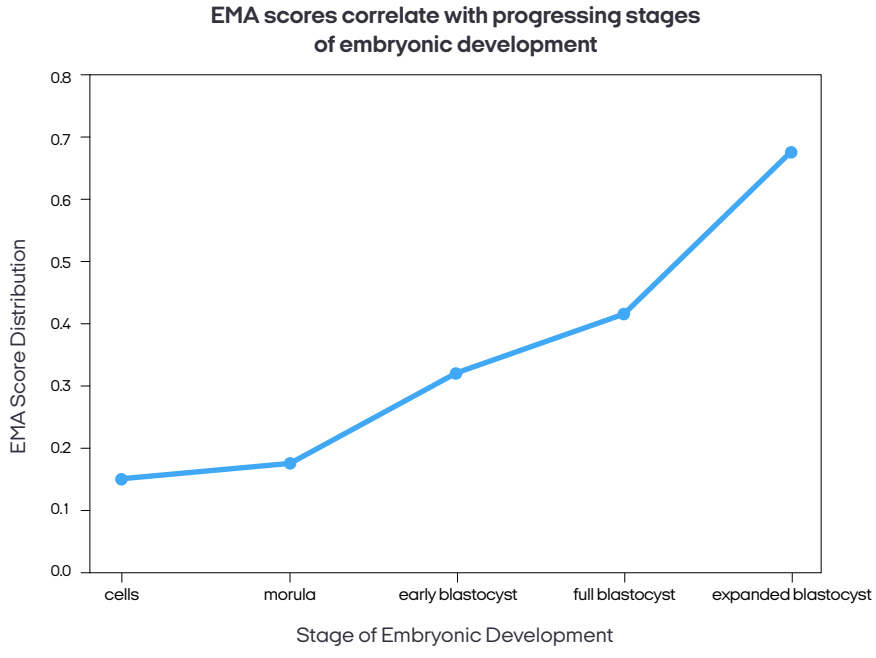
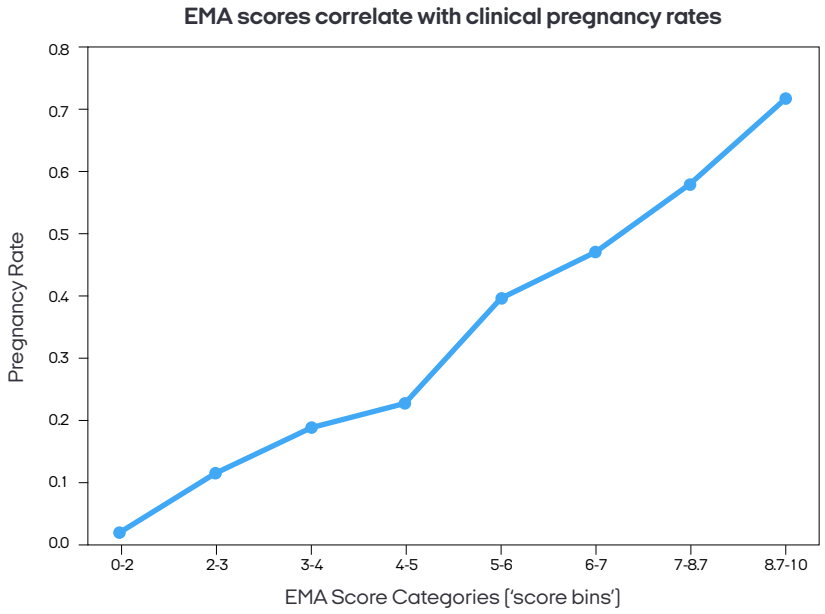


Figure 2.
Method evaluating the correlation between EMA scores and pregnancy rates.
A significant positive linear correlation is observed between ascending EMA scores and the rate of embryos resulting in clinical pregnancy.



AIVF™

AIVF is a reproductive technology company driving the next generation of IVF. The company's proprietary suite of digital solutions harnesses data and AI to empower the fertility care team and their patients with knowledge and transparency. The core technology is evidence-based and driven by real-world clinical use to help patients on a smoother, quicker and more accessible path to parenthood.