Comparison of artificial intelligence and traditional non-invasive methods in prioritizing euploid blastocysts for transfer: a retrospective intra-cohort analysis

Innocenti F [1], Cimadomo V [2], Taggi M [1], Saturno G [1], Chiappetta V [1], Casciani V [1], Albricci L [1], Coticchio G [3], Canosa S [4], Vaiarelli A [1], Ubaldi FM [1], Capalbo A [5,6], Rienzi L [1,7], Cimadomo D [1]

[1] IVIRMA Global Research Alliance, Genera, Clinica Valle Giulia, Rome, Italy; [2] DLL Financial Solutions Partner, Eindhoven, The Netherlands; [3] IVIRMA Global Research Alliance, 9.baby, Bologna, Italy; [4] IVIRMA Global Research Alliance, Livet, Turin, Italy; [5] Juno Genetics, Rome, Italy; [6] Unit of Molecular Genetics, Center for Advanced Studies and Technology (CAST), "G. d'Annunzio" University of Chieti-Pescara, Chieti, Italy; [7] Department of Biomolecular Sciences, University of Urbino "Carlo Bo", Urbino, Italy

Introduction

Embryo morphology and developmental kinetics are associated with chromosomal and reproductive competence. However, traditional assessments rely on subjective interpretations and exhibit low reproducibility. Time-Lapse Technology (TLT) has facilitated continuous monitoring of preimplantation development, yet it has not significantly improved inter-observer agreement. The integration of Artificial Intelligence (AI) with TLT presents an opportunity to enhance standardization and objectivity in embryo assessment. Nevertheless, whole-chromosome testing remains the gold standard for determining embryo viability. AI-driven models aspire to provide a non-invasive prediction of blastocyst (an)euploidy, yet their clinical effectiveness must be validated within embryo cohorts to ensure their reliability and applicability in IVF settings.

Objectives

This study aimed to compare the effectiveness of AI-powered tools versus traditional non-invasive assessments in prioritizing euploid blastocysts for transfer.

Methods

A retrospective blinded analysis was conducted on 786 PGT-A cycles (maternal age: 38.9 years; study period: 2013-2020; 2184 blastocysts). Traditional static assessments were performed by three embryologists using Gardner's grading at the time of biopsy (t-biopsy). Morphodynamic evaluations incorporated time of blastocyst expansion (tEB) and embryo area at tEB (embA). AI prioritization was performed using three commercially available models. The effectiveness of each approach in ranking euploid blastocysts as top-quality was assessed in cohorts containing at least three blastocysts with different chromosomal diagnoses, including at least one euploid and one aneuploid (N=279/786, 35.5%). Coefficients of Variation (CVs = SD/mean) were calculated within cohorts containing at least three blastocysts (N=363/786, 46.2%) to evaluate intra-cohort data dispersion and its association with euploid prioritization.

Results

The intra-cohort coefficients of variation (CVs) were calculated for Gardner's grading and three AI models, revealing distinct patterns of data dispersion. Gardner's grading showed the highest intra-cohort CV (0.71±0.35), while AI models 1, 2, and 3 exhibited CVs of 0.28±0.14, 0.58±0.26, and 0.56±0.34, respectively. Spearman's correlation analysis indicated moderate associations between Gardner's grading CV and AI models 1 and 2 (0.32 and 0.34, respectively), whereas model 3 displayed a weaker correlation (0.14). Notably, AI models 1 and 2 were strongly correlated (0.63) but exhibited weak associations with model 3 (0.21 and 0.22). Euploid blastocysts were prioritized in 58.4% of cases using static assessment (N=163/279), 64.2% using morphodynamic assessment (N=179/279), and in 62.4%, 68.1%, and 52.0% of cases using AI models 1, 2, and 3, respectively. Multivariate analysis revealed that a 0.1-unit increase in intra-cohort CV for Gardner's grading was significantly associated with a higher likelihood of prioritizing euploid over aneuploid blastocysts (OR: 1.1, 95% CI: 1.03-1.19, adjusted p=0.005). In contrast, no such association was observed for AI models. Additionally, aneuploid blastocysts compatible with implantation were prioritized over euploid ones in 21.1% of cases using static assessment (N=59/279), 19.4% with morphodynamic assessment (N=54/279), and in 18.6%, 17.2%, and 19.4% of cases using AI models 1, 2, and 3, respectively. A 0.1-unit increase in Gardner's grading intra-cohort CV was associated with a lower likelihood of prioritizing such aneuploid blastocysts over euploid ones (OR: 0.9, 95% CI: 0.83-0.98, adjusted p=0.010), while AI models showed no significant associations.

Conclusions

While this study reflects real-world clinical scenarios in a large PGT-A program, it is limited by its retrospective, single-center nature and its focus on AMA patients with relatively few cycles meeting inclusion criteria. Although both traditional and AI-based grading methods correlate with euploidy, their primary goal remains embryo prioritization for transfer rather than serving as non-invasive alternatives to PGT-A. Future research should focus on improving live birth prediction rather than solely relying on non-invasive AI models as surrogates for genetic testing.

Recommended reading

- *Cimadomo D, Trio S, Canosi T, et al. Quantitative standardized expansion assay: an artificial intelligence-powered morphometric description of blastocyst expansion and zona thinning dynamics. Life (Basel). 2024;14(11):1396.*
- *Khosravi P et al. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. NPJ Digit Med. 2019;2:21.*
- *Kragh MF, Karstoft H. Embryo selection with artificial intelligence: How to evaluate and compare methods? J Assist Reprod Genet. 2021;38(7):1675-89.*