Abstract citation ID: deae108.560
P-189 Can artificial intelligence outperform traditional non-invasive assessments in prioritizing euploid blastocysts for transfer? A retrospective intra-cohort analysis of 786 PGT-A cycles

D. Cimadomo1, V. Cimadomo2, F. Innocenti1, M. Taggi1, G. Saturno1, V. Chiappetta1, V. Casciani1, L. Albricci1, G. Coticchio3, A. Ahlstrom1, S. Canosa1, A. Vaiarelli1, F. M. Ubaldi1, A. Capalbo6,7, L. Rienzi1,8
1 IVIRMA Global Research Alliance, GENERA- Clinica Valle Giulia, Roma, Italy
2 DLL Financial Solutions Partner, DLL Financial Solutions Partner, Eindhoven, The Netherlands
3 IVIRMA Global Research Alliance, 9.baby, Bologna, Italy
4 IVIRMA Global Research Alliance, Livio, Goteborg, Sweden
5 IVIRMA Global Research Alliance, Livet, Turin, Italy
6 Juno Genetics, Juno Genetics, Rome, Italy
7 “G. d’Annunzio” University of Chieti-Pescara, Unit of Molecular Genetics- Center for Advanced Studies and Technology CAST, Chieti, Italy
8 University of Urbino “Carlo Bo”, Department of Biomolecular Sciences, Urbino, Italy

Study question: Are Artificial Intelligence (AI)-powered tools more efficient than traditional non-invasive assessments to prioritize euploid blastocysts for transfer?

Summary answer: Only one third of cycles yielded ≥3 blastocysts with different diagnoses. In these cycles, AI and traditional grading prioritized euploid embryos with comparable performance.

What is known already: While embryo morphology and developmental pace to blastocyst are associated with chromosomal and reproductive competence, their evaluation remains subjective and lacks reproducibility. The introduction of Time-lapse technology (TLT) in IVF has provided valuable insights into preimplantation development but has not enhanced embryologists’ reproducibility. Recently, AI models integrated with TLT offered the potential for automating and standardizing assessments. Whole-chromosome testing for uniform aneuploidies still is the strongest predictor of embryo (in)e-competence. In fact, if transferred, >98% of aneuploid blastocysts fail to result in a live-birth. AI tools strive to non-invasively predict blastocyst (an)eu- ploidy, yet their clinical utility requires intra-cohort testing.

Study design, size, duration: Retrospective blinded analysis of 786 PGT-A cycles (maternal-age:38.9; years:2013-2020; 2184 blastocysts). For static assessment, 3 embryologists used Gardner’s grading at the time of biopsy (t-biopsy). For morphodynamic assessment, time of blastocyst-expansion (tEB)
and embryo-area (embA at tEB) were also annotated. For AI assessment, scores were generated with three commercially-available models. All approaches were compared for their effectiveness in ranking euploid blastocyst(s) as top-quality among cohorts with ≥3 blastocysts and ≥1 euploid and ≥1 aneuploid (N = 279/786, 35.5%).

**Participants/materials, setting, methods:** Embryologists ranked AA-blastocysts as top-quality (score=1), CC-blastocysts as worst-quality (score=9), prioritizing trophectoderm over ICM, with t-biopsy as third criterion. Morphodynamic assessments followed the same hierarchy but using tEB and embA in place of t-biopsy. For AI, higher scores corresponded to higher priority. For all assessments, coefficients of variation (CV=SD/mean) were calculated to assess data dispersion within cohorts with ≥3 blastocysts (N = 363/786, 46.2%). Intra-cohort CVs were tested for their association with the likelihood of prioritizing euploid blastocysts.

**Main results and the role of chance:** Intra-cohort CVs were 0.71 ± 0.35, 0.28 ± 0.14, 0.58 ± 0.26, and 0.56 ± 0.34 for Gardner’s score, AI model-1, -2, and -3, respectively. Spearman’s correlations between Gardner’s score and AI model-1, -2, and -3 Intra-cohort CVs were 0.32, 0.34, and 0.14, respectively. AI model-1 and -2 were strongly correlated (0.63), but poorly with model-3 (0.21 and 0.22).

Euploid blastocysts were prioritized in 163 cohorts by static assessments (N = 163/279, 58.4%), 179 (62.4%) by morphodynamic, 174 (62.4%), 190 (68.1%), and 145 (52%) by AI model-1, -2, and -3.

A maternal age-adjusted 0.1-larger Gardner’s grading intra-cohort CV correlated with higher likelihood of prioritizing euploid over aneuploid blastocysts (multivariate-OR:1.1, 95%CI:1.03-1.19, adjusted-p=0.005). No association was reported with AI models intra-cohort CVs.

Notably, blastocysts affected from aneuploidy compatible with implantation were prioritized over euploid in 59 cohorts (N = 59/279, 21.1%) by static assessment, 54 (19.4%) by morphodynamic, 52 (18.6%), 48 (17.2%), and 54 (19.4%) by AI model-1, -2, and -3.

A maternal age-adjusted 0.1-larger Gardner’s grading intra-cohort CV correlated with lower likelihood of prioritizing aneuploid blastocysts compatible with implantation over euploid (multivariate-OR:0.9, 95%CI:0.83-0.98, adjusted-p=0.010). No association was reported with AI models intra-cohort CVs.

**Limitations, reasons for caution:** While representing the real-life scenario of a large PGT-A program, this study is a single center retrospective analysis in advanced maternal age women with a limited number of cycles with ≥3 blastocysts showing both euploid and aneuploid. Other AI-powered models exist, some specifically trained to predict euploidy.

**Wider implications of the findings:** Trophoderm biopsy and testing aim to prevent the transfer of chromosomally-abnormal blastocysts. Although traditional/AI-based grading correlate with euploidy, their goal is prioritizing embryos for transfer, irrespective of the IVF setting (with or without PGT-A). Future research should concentrate on emphasizing live-birth prediction over misleadingly considering them as non-invasive PGT-A surrogates.

**Trial registration number:** None