American Society for Reproductive Medicine 2017 Scientific Congress & Expo
October 28 to November 1, 2017 • San Antonio, TX, USA

Title

LEVERAGING ARTIFICIAL INTELLIGENCE FOR MORE DATA-DRIVEN PATIENT COUNSELING AFTER FAILED IVF CYCLES

Authors

K. Hunter Cohn,[1] Q. Zhang,[1] A. B. Copperman,[2] P. Yurttas Beim[1]

Affiliations

1. Celmatix Inc, New York, NY
2. Obstetrics and Gynecology, RMANY-Mount Sinai, New York, NY

Objective

Baseline metrics such as age and hormone levels are used to counsel patients undergoing IVF. For patients who fail multiple cycles, counseling can become more challenging and cycle-level metrics become an important consideration. As datasets grow, artificial intelligence (AI) systems are augmenting the diagnostic and prognostic capabilities of the human brain. The limitation is that these systems are only as smart as the input data they are trained on. However, the advantage is that they are unbiased, bring greater standardization to counseling, and often yield non-obvious insights. Here, we applied AI methods as a proof-of-concept and to better understand trends after failed IVF cycles.

Design

Retrospective study of 21,832 autologous IVF cycles from 13 centers.

Materials and Methods

We built a predictive model for ongoing pregnancy after failed IVF cycles using Extreme Gradient Boosting methodology. Features included patient baseline score (multivariate model for first cycle probability of pregnancy using age, BMI, diagnosis, basal antral follicle count, and levels of AMH, FSH, LH, and estradiol), metrics from the most recent retrieval cycle (eggs retrieved (ER), total usable embryos (TUE), embryo stage), and IVF cycle history (number of times a patient had undergone an embryo transfer (ET), biochemical and clinical losses, cancelled cycles). The algorithm also considered ET parameters (number of embryos to be transferred and whether embryos were PGS screened) on the cycle being predicted. A feature importance (FI) matrix provided a score for each feature's importance in the model. P-value for a feature was determined by whether exclusion significantly decreased AUC.

Results

AI models revealed that, while the most important prognostic feature for ongoing pregnancy after a failed cycle was baseline score (FI 61%, $p<0.0001$), additional significant features included: ET parameters (10%); clinic (7%); retrieval metrics, including TUE (6%) and number of ER (6%); number of failed transfers (3%); and prior clinical (2%) or biochemical (1%) losses.

Interestingly, the predictive importance of cycle-level metrics such as number of ER and TUE varied significantly across failed cycles, whereas the importance of baseline metrics stayed constant. For example, patients with poor response (5 ER,1 TUE) are predicted to have twice as rapid a decline in probabilities with each failed cycle as patients with good response (26 ER, 8 TUE).

Conclusions

AI models demonstrate that, while baseline patient metrics reflect overall prognosis, they do not provide additional information over multiple failed cycles. In contrast, retrieval metrics become increasingly important for refining counseling after failed cycles. AI systems hold the promise of augmenting and bringing greater standardization to patient counseling, especially around difficult decision points that require the balancing of several variables, such as whether to persist in treatment after failed cycles.

Support

Celmatix Inc.