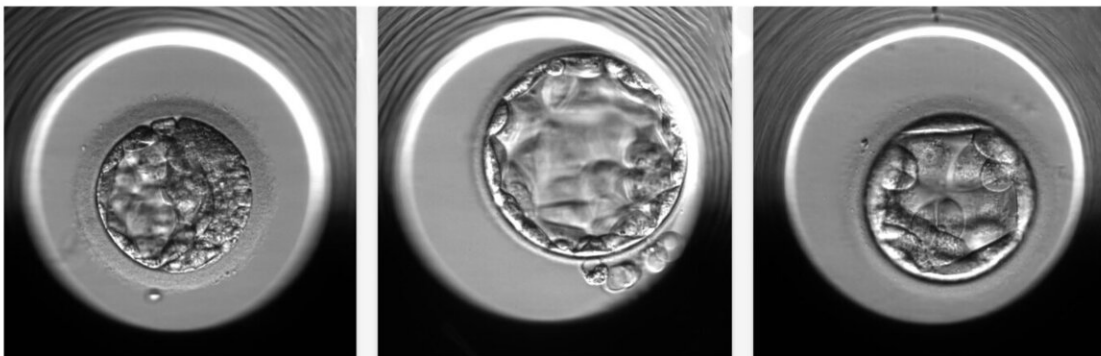# Harnessing artificial intelligence technology for IVF embryo selection

December 20 2022

Examples of embryos evaluated by the STORK-A algorithm. From left to right, an embryo predicted to have a normal chromosome count or a single chromosomal abnormality; an embryo predicted to have a normal chromosome count; an embryo predicted to have more than one chromosomal abnormality. Credit: Weill Cornell Medicine

An artificial intelligence algorithm can determine non-invasively, with about 70 percent accuracy, if an in vitro fertilized embryo has a normal or abnormal number of chromosomes, according to a new study from researchers at Weill Cornell Medicine.

Having an abnormal number of chromosomes, a condition called aneuploidy, is a major reason embryos derived from in vitro fertilization (IVF) fail to implant or result in a healthy pregnancy. One of the current methods for detecting aneuploidy involves the biopsy-like sampling and genetic testing of cells from an embryo—an approach that adds cost to the IVF process and is invasive to the embryo. The new algorithm, STORK-A, described in a paper published Dec. 19 in *Lancet Digital Health*, can help predict aneuploidy without the disadvantages of biopsy. It operates by analyzing microscope images of the embryo and incorporates information about maternal age and the IVF clinic's scoring of the embryo's appearance.

"Our hope is that we'll ultimately be able to predict aneuploidy in a completely non-invasive way, using artificial intelligence and computer vision techniques," said study senior author Dr. Iman Hajirasouliha, associate professor of computational genomics and of physiology and biophysics at Weill Cornell Medicine and a member of the Englander Institute for Precision Medicine.

The study's first author is Josue Barnes, a doctoral student in the Weill Cornell Graduate School of Medical Sciences who studies in the Hajirasouliha Laboratory. Dr. Nikica Zaninovic, associate professor of embryology in clinical obstetrics and gynecology and director of the Embryology Laboratory at the Ronald O. Perelman and Claudia Cohen Center for Reproductive Medicine at Weill Cornell Medicine and NewYork-Presbyterian/Weill Cornell Medical Center led the embryology work for the study.

According to the U.S. Centers for Disease Control and Prevention, there were more than 300,000 IVF cycles performed in the United States in 2020, resulting in about 80,000 live births. IVF experts are always looking for ways to boost that success rate, to achieve more successful pregnancies with fewer embryo transfers—which means developing better methods for identifying viable embryos.

Fertility clinic staff currently use microscopy to assess embryos for large-scale abnormalities that correlate with poor viability. To obtain information about the chromosomes, clinic staff may also use a biopsy method called preimplantation genetic testing for aneuploidy (PGT-A), predominantly in women over the age of 37.

To develop a computational approach to embryo assessment that capitalized on the Embryology Laboratory's pioneering use of time lapse photography, investigators from the Center for Reproductive Medicine teamed up with colleagues in the Englander Institute.

In a 2019 study, the teams developed an artificial intelligence (AI) algorithm, STORK, that could assess embryo quality about as well as IVF clinic staff. For the new study, they developed STORK-A as a potential replacement for PGT-A—or as a more selective way of deciding which embryos should have PGT-A testing.

The new STORK-A algorithm uses microscope images of embryos taken at five days past fertilization, clinic staff's scoring of embryo quality, maternal age, and other information that is normally gathered as part of the IVF process. Because it uses AI, the algorithm automatically "learns" to correlate certain features of the data, often too subtle for the human eye, with the chance of aneuploidy. The team trained STORK-A on a dataset of 10,378 blastocysts for which ploidy status was already known.

From its performance, they assessed the algorithm's accuracy in

predicting aneuploid versus normal-chromosome "euploid" embryos at nearly 70 percent (69.3%). In predicting aneuploidy involving more than one chromosome—complex aneuploidy—versus euploidy, STORK-A was 77.6 percent accurate. They later tested the algorithm on independent datasets, including one from an IVF clinic in Spain, and found comparable accuracy results, demonstrating the generalizability of STORK-A.

The study provides a proof of concept for an approach that is currently experimental. Standardizing the use of STORK-A in clinics would require clinical trials comparing it to PGT-A, and Food and Drug Administration approval—all years in the future. But the new algorithm represents progress on the way to making IVF embryo selection less risky, less subjective, less costly and more accurate.

"This is another great example of how AI can potentially transform medicine. The algorithm turns tens of thousands of embryo images into AI models that may ultimately be used to help improve IVF efficacy and further democratize access by reducing costs," said co-author Dr. Olivier Elemento, director of the Englander Institute for Precision Medicine and a professor of physiology and biophysics and of computational genomics in computational biomedicine at Weill Cornell Medicine.

"We believe that ultimately by using this technology we can reduce the number of embryos to be biopsied, reduce the costs, and provide a very good tool for consultation with the patient when they need to make decision whether to do PGT-A or not," said Dr. Zaninovic.

The team now plans to build on this success with algorithms trained on videos of embryo development.

"By using video classification, we can leverage both temporal and spatial information about the embryo's development, and hopefully that will

allow the detection of trends in development that distinguish aneuploidy from euploidy with even higher accuracy," Barnes said.

"This technology is being optimized with the hope that at some point its accuracy will be close to genetic testing, which is the gold standard and is more than 90 percent accurate," said co-author Dr. Zev Rosenwaks, director and physician-in-chief of the Ronald O. Perelman and Claudia Cohen Center for Reproductive Medicine at NewYork-Presbyterian/Weill Cornell Medical Center and Weill Cornell Medicine, and the Revlon Distinguished Professor of Reproductive Medicine in Obstetrics and Gynecology at Weill Cornell Medicine. "But we realize that this goal is aspirational."

More information: Josue Barnes et al, A non-invasive artificial intelligence approach for the prediction of human blastocyst ploidy: a retrospective model development and validation study, *The Lancet Digital Health* (2022). DOI: 10.1016/S2589-7500(22)00213-8

Provided by Weill Cornell Medical College