Application of machine learning to predict aneuploidy and mosaicism in embryos from in vitro fertilization cycles

José A. Ortiz, PhD; Ruth Morales, PhD; Belén Lledó, PhD; Juan A. Vicente, PhD; Julio González, PhD; Eva M. García-Hernández, PhD; Alba Cascales, MSc; Jorge Ten, PhD; Andrea Bernabeu, PhD, MD; Rafael Bernabeu, PhD, MD

BACKGROUND: The factors associated with embryo aneuploidy have been extensively studied. Mostly maternal age and to a lesser extent male factor and ovarian stimulation have been related to the occurrence of chromosomal alterations in the embryo. On the other hand, the main factors that may increase the incidence of embryo mosaicism have not yet been established.

OBJECTIVE: This study aimed to establish a machine learning model that would allow prediction of aneuploidies and mosaicism in embryos conceived via in vitro fertilization, and thus help to determine which variables are associated with these chromosomal alterations.

STUDY DESIGN: The study design was observational and retrospective. A total of 6989 embryos from 2476 cycles of preimplantation genetic testing for aneuploidies were included (January 2013 to December 2020). The trophoectoderm biopsies on day-5, -6, or -7 blastocysts were analyzed by preimplantation genetic testing for aneuploidies (PGT-A). The different maternal, paternal, couple, embryo, and in vitro fertilization cycle characteristics were recorded in a database (22 predictor variables) from which predictive models of embryo aneuploidy and mosaicism were developed; 16 different unsupervised classification machine learning algorithms were used to establish the predictive models.

RESULTS: Two different predictive models were performed: one for aneuploidy and the other for mosaicism. The predictor variable was of multiclass type because it included the segmental- and whole-chromosome alteration categories. The best predicting models for both aneuploidies and mosaicism were those obtained from the Random Forest algorithm. The area under ROC curve (AUC) value was 0.792 for the aneuploidy explanatory model and 0.776 for mosaicism. The most important variable in the final aneuploidy model was maternal age, followed by paternal and maternal karyotype and embryo quality. In the predictive model of mosaicism, the most important variable was the technique used in preimplantation genetic testing for aneuploidies and embryo quality, followed by maternal age and day of biopsy.

CONCLUSION: It is possible to predict embryo aneuploidy and mosaicism from certain characteristics of the patients and their embryos.

Key words: array comparative genomic hybridization, artificial intelligence, embryo aneuploidy, embryo mosaicism, machine learning, next-generation sequencing, preimplantation genetic testing for aneuploidies

Introduction

Embryo aneuploidy is the main cause of failure in in vitro fertilization (IVF) cycles.[1] These aneuploidies can cause various events, such as embryo arrest, implantation failure, first-trimester miscarriage, or congenital anomalies.[2,3] Most meiotic errors leading to aneuploid embryos originate in female gamete meiosis.[4] Several studies have shown that these alterations are associated with maternal age.[5,6]

Conversely, patients with abnormal karyotypes produce a high rate of unbalanced gametes that result in aneuploid embryos.[7] When analyzing in depth the role of the karyotype, studies conducted by our team have found an increase in embryo aneuploidy when one of the parents had karyotype polymorphism.[8]

Finally, embryo quality is an excellent predictive factor for embryo aneuploidy. Poorer-quality embryos (grades C and D) present a higher rate of chromosomic

From the Department of Molecular Biology and Genetics, Instituto Bernabeu, Alicante, Spain (Drs Ortiz, Morales, Lledó, and García-Hernández, and MSc Cascales); Departamento de Economía Aplicada y Estadística, Universidad Nacional de Educación a Distancia, Madrid, Spain (Dr Vicente); Departamento de Economía de la Empresa y Contabilidad, Universidad Nacional de Educación a Distancia, Madrid, Spain (Dr González); Departments of Reproductive Biology (Dr Ten); Reproductive Medicine (Drs A Bernabeu and R Bernabeu), Instituto Bernabeu, Alicante, Spain.

Received October 17, 2021; revised September 5, 2022; accepted September 13, 2022.

The authors report no conflict of interest.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Cite this article as: Ortiz JA, Morales R, Lledó B, et al. Application of machine learning to predict aneuploidy and mosaicism in embryos from in vitro fertilization cycles. Am J Obstet Gynecol Glob Rep 2022;XX:x.ex–x.ex.

Corresponding author: José A. Ortiz, PhD. jaortiz@institutobernabeu.com

2666-5778/$36.00

http://dx.doi.org/10.1016/j.xagr.2022.100103

AJOG Global Reports at a Glance

Why was this study conducted?

The aim of this study is to establish machine learning models to predict aneuploidy and mosaicism in embryos from IVF treatments.

Key findings

Two different predictive models were performed, one for aneuploidy (AUC: 0.792) and the other for mosaicism (AUC: 0.776). The best models were those obtained from the Random Forest algorithm.

The most important variable in the aneuploidy model was maternal age and in the predictive model for mosaicism, was the technique used in PGT-A and embryo quality.

What does this add to what is known?

Artificial intelligence can be a very useful tool in reproductive medicine, in particular the machine learning models established for embryo aneuploidy and mosaicism allow the identification of couples at risk of embryo aneuploidy/mosaicism, which could benefit from the use of PGT-A.

alterations compared with good-quality embryos (grades A and B).[9]

In addition to maternal age, karyotype and embryo quality have been considered as factors associated with embryo aneuploidy, although they are not as widely accepted because of contradictory results reported in the literature. These factors are related to the male parent (male factor), the couple's background, and ovarian stimulation.

Fluorescence in situ hybridization (FISH) analysis of spermatozoa was the main technique used to assess the paternal contribution to embryo aneuploidy. Recent studies have shown an increased embryo aneuploidy rate in males with an elevated number of aneuploid spermatozoa.[10] However, other studies consider that this paternal contribution plays a secondary role only.[11,12,13]

Previous recurrent implantation failure (RIF) or pregnancy loss (RPL) are other risk factors for embryo aneuploidy. If, in addition, the couple has a history of pregnancy with chromosomal abnormalities, the probability of chromosomal alterations in a future pregnancy is increased.[14]

Ovarian stimulation may also affect the embryo aneuploidy rate. Recently, it has been shown that embryos from patients who require more days of hormonal treatment to achieve oocyte maturation have a lower aneuploidy rate.[15]

Mosaicism is a special type of chromosomal alteration in the embryo. This genetic alteration consists of the coexistence of ≥2 populations of cells with a different chromosomal endowment.[16]

There are several established mechanisms causing mosaicism: nondisjunction of chromosomes during cell division, anaphase lagging, which prevents the correct incorporation of a single chromatid into each nucleus of the cells after mitosis, and finally, replication of DNA without cell division, called endoreplication. Anaphase lagging is the main process causing mosaicism in the preimplantation embryo.[17]

Maternal, paternal, and embryo factors that affect this phenomenon are still unclear. In some cases, these alterations can be corrected by cellular systems.[18,19] According to the studies carried out to date, maternal age does not seem to play a central role in embryo mosaicism.[20] In contrast, paternal age does seem to increase the embryo mosaicism rate.[20]

As in the case of aneuploidies, embryo quality seems to be a good predictive factor. Good-quality embryos (A and B) have lower rates of mosaicism when compared with lower-quality embryos (C and D) (unpublished data).

Other factors associated with mosaicism are those related to the embryo culture conditions during the IVF treatment.[21]

The most common chromosomal alterations in embryos involve whole chromosomes, but there is a percentage of aberrations that affect only a chromosome segment. The origin of segmental alterations is diverse[22] and can be produced either by errors during meiosis in the gametogenesis or by a de novo phenomenon that occurs after chromosomal breaks that take place in mitosis at the early stages of embryo development. A third possibility involves one of the parents carrying a balanced translocation in their karyotype.

Machine learning algorithms are being used in medicine increasingly frequently. These models are helping clinicians to diagnose and optimize treatments in a wide range of medical fields. Human fertility is not an exception to this phenomenon, and in recent years, there has been a growing interest in this analysis methodology.[23,24] Assisted reproductive technology (ART) generates a large amount of data, which makes it a perfect target for the application of different artificial intelligence algorithms.

This study emerged in this context with the aim of being able to predict the personalized probability of having chromosomally normal embryos by taking into account data from patients' medical records.

Materials and Methods

Study design

The study design was observational and retrospective. The data were derived from the results of preimplantation genetic testing for aneuploidies (PGT-A) of embryos from couples with fertility problems who attended Instituto Bernabeu. These data covered the period from January 2013 to December 2020, corresponding to 6989 embryos from 2476 IVF cycles.

Ethical approval

All work was conducted with previous formal approval of the Instituto Bernabeu Institutional Review Board and followed the principles of the Declaration of Helsinki.

Preimplantation genetic testing for aneuploidies

The medical indications for PGT-A were advanced maternal age, abnormal karyotype of one of the parents, high rate of chromosome aneuploidies in sperm samples, history of chromosomal abnormalities, and history of repeated miscarriages and embryo implantation failure (defined as the transfer of at least 2 embryos in the uterine cavity without implantation).

All couples were informed and signed the corresponding informed consent for the PGT-A procedure.

The embryo biopsies were performed at the blastocyst stage on days 5, 6, or 7 of embryo development. Fragments between 3-6 cells were removed from the trophoectoderm.

After lysing the biopsied cells, embryo genome amplification was performed using the PicoPLEX kit (Rubicon Genomics, Ann Arbor, MI) following the manufacturer's instructions.

All biopsies were processed for chromosomal analysis by Agilent SurePrint G3 8 × 60K CGH microarrays (n=2335) (Agilent Technologies, Santa Clara, CA) or VeriSeq next-generation sequencing (NGS) (Illumina, San Diego, CA) (n=4654).

Embryos with a percentage of aneuploid cell line <25% were classified as euploid and were suitable to be transferred to the maternal uterus. Embryos were classified as mosaic if the percentage of the aneuploid cell line was between 25% and 50%. These mosaic embryos were transferred according to the recommendations of the Preimplantation Genetic Diagnosis International Society (PGDIS).[25] Despite the lack of scientific evidence, these embryos are discarded in many fertility clinics, which they justify with avoiding the risk of the child being born with chromosomal alterations.[16]

Finally, if the proportion of aneuploid cells was >50%, the embryo was classified as aneuploid and considered unsuitable for transfer.

Univariate analysis

Comparison between the study groups (euploid vs aneuploid and mosaic vs nonmosaic) for categorical variables was performed using the Pearson chi-square test. The normal distribution of the variables was analyzed using the Shapiro−Wilk test. If the distribution was normal, the comparison between the different groups was carried out using the Student *t* test, otherwise the Mann−Whitney U-test was used. Differences were considered statistically significant when $P<.05$.

Data preprocessing

The dataset was analyzed using different multiclass classification algorithms. These were all supervised algorithms. Analysis was done on the basis of a data frame after anonymization of the results obtained after PGT-A. Noninformative embryos were excluded from the analysis.

Firstly, a preprocessing of the database was carried out. In the factors, levels with few observations were grouped together. Missing values (0.15%) were imputed and there were no outliers detected.

The starting database contained 29 predictor variables, which could be classified into 6 groups: general, maternal, paternal, couple-related, IVF cycle −related, and embryo-related (Table 1). Nonrelevant variables and those showing a strong correlation were eliminated. Because Bayesian algorithms were to be used, the numeric variables were discretized according to quartiles. For the specific case of the XGBoost model, the variables were transformed into numeric variables.

Two different multiclass models were implemented: one to predict aneuploidies (euploid, whole-chromosome aneuploid, segmental-chromosome aneuploid, or both types of aneuploidy) and one to predict mosaicism (nonmosaic, whole-chromosome mosaic, segmental-chromosome mosaic, and both types of mosaicism).

Before training the models, class balancing for the variables to be predicted was carried out.

Hyperparameter optimization of multiclass models

A total of 16 machine learning algorithms (classification) were applied (Supplemental Table 1), ranging from multinomial regression, artificial neural networks, support vector machines, neighborhood-based methods, classification trees, gradient boosting, ensemble methods, Bayesian methods, and discriminant analysis−based methods.

The best model was selected on the basis of the area under the curve (AUC) or area under the receiver operating characteristic (ROC) curve, a parameter that reflects the model's ability to discriminate the dependent variable. AUC values were obtained using a cross-validation (5-fold) procedure after randomly dividing the data frame into training (80%) and validation data (20%). During the optimization process, different performance metrics were calculated: logLoss (negative of the multinomial log-likelihood based on the class probabilities), accuracy, AUC, sensitivity, and specificity.

Once each of the algorithms had been optimized, the best models were compared pairwise by applying a paired *t* test with Bonferroni correction.

Final predictive model

The final Random Forest model was obtained using the complete database. To guarantee the independence between the data and to be able to properly evaluate the models, the cross-validation technique was applied with 10 folds with adjustment of the different hyperparameters. This time the validation error was estimated from the out-of-bag (OOB) error. This error is determined from the predictions of observations not included in the model generation.

The ROC curve was determined for each of the response variable clases, and the overall mean was calculated by 2 different methods: macro, calculated by averaging the results of all groups (one vs the rest) by linear interpolation between the points of the ROC curves; and micro, obtained by aggregating all the groups, thus converting the multiclass classification into a binary classification.

The most important predictor variables in the final model using the Random Forest algorithm were determined from the relationship between 3

TABLE 1
Descriptive of patients, in vitro fertilization cycle and preimplantation genetic testing for aneuploidies

Variable	Variable name	Total	Aneuploidy			Mosaicism		
			−	+	P value	−	+	P value
PGT-A technique		n=6989	n=3731	n=3258		n=5836	n=1153	
aCGH (%)	TECHNIQUE	33.4	38.6	27.4	<.001[a]	35.2	24.3	<.001[a]
NGS (%)		66.6	61.4	72.6		64.8	75.7	
Patients								
Maternal age (mean±SD)	MATERNAL_AGE	33.82±6.82	31.65±6.61	36.27±6.21	<.001[b]	33.95±6.79	33.19±6.95	.001[b]
Paternal age (mean±SD)	PATERNAL_AGE	39.29±6.99	39.35±6.87	39.22±7.13	.423[b]	39.21±6.98	39.70±7.00	.008[b]
Male karyotype	MALE_KARYOTYPE							
Altered (%)		2.4	1.8	3.1	.004[a]	2.4	2.4	.988[a]
Polymorphism (%)		8.6	8.9	8.2		8.6	8.4	
Female karyotype	FEMALE_KARYOTYPE							
Altered (%)		1.4	0.9	2.0	<.001[a]	1.5	1.0	.547[a]
Polymorphism (%)		7.5	8.8	5.9		7.5	7.2	
Sperm Count	SPERM_COUNT							
Normozoospermia (%)		78.2	77.8	78.7	.644[a]	78.0	79.3	.260[a]
Oligozoospermia (%)		18.6	19.0	18.2		18.7	18.2	
Cryptozoospermia (%)		2.6	2.8	2.5		2.8	1.8	
Azoospermia		0.5	0.5	0.6		0.5	0.6	
Astenozoospermia (%)	ASTENOZOOSPERMIA	20.2	20.1	20.3	.824[a]	20.1	20.4	.848[a]
Teratozospermia (%)	TERATOZOSPERMIA	14.3	14.1	14.5	.666[a]	14.8	11.6	.005[a]
Altered sperm aneuploidy test -FISH- (%)	FISH	11.2	13.2	8.8	<.001[a]	11.1	11.6	.159[a]
Pathologic sperm DNA fragmentation-TUNEL-(%)	TUNEL	3.9	4.1	3.7	<.001[a]	3.9	3.6	.412[a]
RIF (%)	RIF	20.9	20.5	16.3	<.001[a]	20.3	24.3	.002[a]
RPL (%)	RPL	21.0	22.6	19.2	.001[a]	21.0	21.1	.980[a]
History of chromosomopathies (%)	CHROMOSOMOPATHIES	21.9	20.6	23.3	.007[a]	21.7	22.8	.409[a]
IVF Cycle								
Cycles with donated oocyte (%)	OWN_OOCYTE	34.8	47.1	20.8	<.001[a]	34.0	38.9	.002[a]

Ortiz. Prediction of aneuploid and mosaic embryos. Am J Obstet Gynecol Glob Rep 2022.

(continued)

TABLE 1
Descriptive of patients, in vitro fertilization cycle and preimplantation genetic testing for aneuploidies *(continued)*

Variable	Variable name	Total	Aneuploidy			Mosaicism		
			−	+	P value	−	+	P value
Stimulation in luteal phase (%)	LUTEAL_PHASE_OOCYTE	4.7	2.8	6.8	<.001[a]	5.0	3.3	.014[a]
Vitrified oocyte (%)	VITRIFIED_OOCYTE	10.0	10.0	10.1	.829[a]	9.9	10.8	.366[a]
Cycles with donated sperm (%)	DONATED_SPERM	8.4	7.0	10.2	<.001[a]	8.6	7.5	.227[a]
Frozen sperm cycles (%)	FROZEN_SPERM	10.7	12.4	8.8	<.001[a]	10.3	13.1	.005[a]
Recovered oocytes (mean±SD)	RECOVERED_OOCYTES	12.73±6.25	13.01±6.04	12.40±6.47	<.001[b]	12.74±6.26	12.64±6.23	.740[b]
Mature oocytes (MII) (mean±SD)	MII	10.63±5.01	10.97±4.85	10.24±5.15	<.001[b]	10.64±4.99	10.55±5.06	.806[b]
Number of embryos biopsied (mean±SD)	EMBRYOS_BIOPSIED	4.01±2.18	4.06±2.10	3.96±2.27	.002[b]	4.01±2.19	3.99±2.15	.870[b]
Embryo								—
Aneuploidy (%)	ANEUPLOIDY	46.6	—	—	—	47.2	43.7	.031[a]
Mosaicism (%)	MOSAICISM_	16.5	17.4	15.5	.031[a]	—	—	—
Segmental alterations (%)	SEGMENTAL_ALT	13.4	7.2	20.6	<.001[a]	8.1	40.7	<.001[a]
Whole chromosome alterations (%)	WHOLE_CHROMOSOME_ALT	48.5	12.4	89.9	<.001[a]	42.0	81.5	<.001[a]
Embryo quality (%)	EMBRYO_QUALITY							
A		45.1	54.0	35.0	<.001[a]	46.5	38.2	<.001[a]
B		48.1	42.5	54.6		47.0	53.9	
C		5.3	2.9	7.9		5.1	6.1	
D		1.5	0.6	2.5		1.4	1.8	
Biopsy day (%)	BIOPSY_DAY							
D+5	D_5	60.3	66.5	53.3	<.001[a]	60.7	58.2	.252[a]
D+6	D_6	38.2	32.6	44.5		37.8	40.1	
D+7	D_7	1.5	0.9	2.2		1.5	1.7	

aCGH, array comparative genomic hybridization; *IVF*, in vitro fertilization; *NGS*, next-generation sequencing; *PGT-A*, preimplantation genetic testing for aneuploidies; *SD*, standard deviation.

Comparison of values corresponding to euploid vs. aneuploid and mosaic vs. non-mosaic embryos. Numerical variables were compared using the Mann Whitney U test (a) and categorical variables using Pearson's chi-square (b). Differences were considered statistically significant when $p<0.05$.

Ortiz. Prediction of aneuploid and mosaic embryos. Am J Obstet Gynecol Glob Rep 2022.

measures of importance: mean depth, number of trees, and total number of nodes.

Results

A descriptive analysis of the set of predictor variables is shown in Table 1. The method of analysis of embryo aneuploidy was mainly NGS.

The mean age of female participants included in the study was 33.82±6.82 years, whereas male participants were older (39.29±6.99 years). In the case of gamete (oocyte and/or sperm) donation cycles, the age of the donor was recorded as maternal or paternal age. Overall, male participants had good semen quality, with 78.2% having normal sperm counts (normozoospermia) according to the World Health Organization classification. Only 11.2% of patients showed elevated levels of sperm aneuploidy.

Conversely, 20.9% of couples had suffered RIF and 21.0% had suffered RPL.

Donated oocytes and sperm were used in 34.8% and 8.4% of the cycles, respectively. The mean number of oocytes retrieved after ovarian stimulation was 12.73±6.25 oocytes, out of which 10.63±5.01 were mature.

The rate of embryo aneuploidy detected was 46.6% and the rate of mosaicism was 16.5%. The segmental alterations rate, including both mosaicism and aneuploidy, was 13.4%. Most of the embryos analyzed were of very good quality (A: 45.1% and B: 48.1%), and biopsies were mainly performed on days 5 (60.3%) or 6 (38.2%).

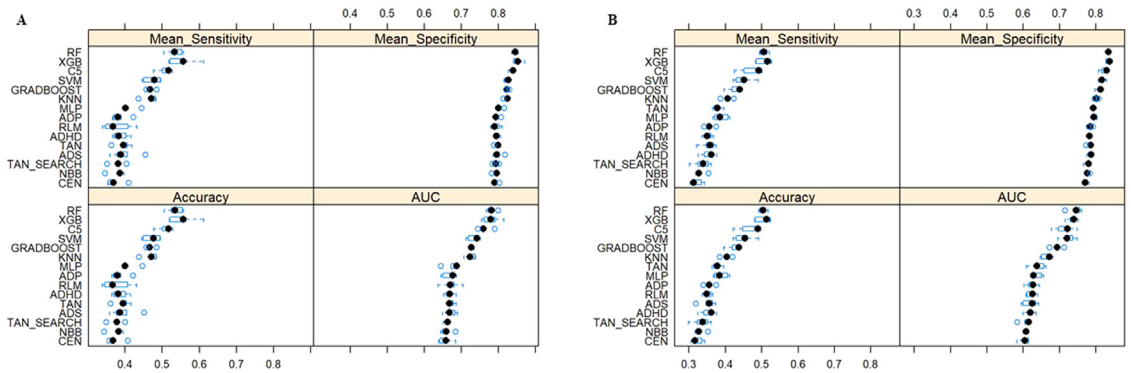
When a univariate analysis of the different characteristics of the patient, the IVF cycle, and the embryos was performed, a large number of variables showed statistically significant differences between the euploid and aneuploid embryo group, and between embryos with and those without mosaicism (Table 1). As expected, maternal age was higher in aneuploid embryos (36.27±6.21 vs 31.65±6.61 years in euploid embryos). This was the opposite in the case of mosaicism, with mosaic embryos having slightly lower maternal age (33.19±6.95 years) compared with embryos without mosaicism (33.95±6.79 years). The pattern of paternal age was different. No differences were observed in the case of aneuploidies (39.35±6.87 vs 39.22±7.13 years), but in contrast, mosaic embryos showed slightly higher paternal ages (39.70±7.00 years) than embryos without mosaicism (39.21±6.98 years). When one of the members of the couple had an abnormal karyotype, the proportion of aneuploid embryos increased, both in the case of paternal (3.1% vs 1.8%) and maternal (2.0% vs 0.9%) origin. No differences were observed in the case of mosaicism. Semen quality did not seem to be modified by either embryo aneuploidy or mosaicism. The only exception was teratozoospermia, which was higher in embryos without mosaicism (14.8% vs 11.6%).

Surprisingly, FISH analysis in spermatozoa (euploid: 13.2% vs aneuploid: 8.8%), sperm DNA fragmentation (euploid: 4.1% vs aneuploid: 3.7%), history of RIF (euploid: 20.5% vs aneuploid: 16.3%), and RPL (euploid: 22.6% vs aneuploid: 19.2%) showed inverse tendencies relative to those previously described in the literature.

Conversely, in both euploid and non-mosaic embryos the proportion of good-quality embryos was higher than that observed in aneuploid or mosaic embryos (54.0% vs 35.0% and 46.5% vs 38.2%, respectively).

Two machine learning models were developed: one to predict aneuploidies

FIGURE 1
Performance metrics of different predictive models.



Results obtained in terms of AUC, mean sensitivity, mean specificity, and accuracy for each of the machine learning algorithms used in the multiclass predictive model (**A,** aneuploidy; **B,** mosaicism). The results obtained with *Bagged CART* are not shown in the figure but are inferior to the best models.

ADHD, high-dimensional discriminant analysis; *ADP*, penalized discriminant analysis; ADS, shrinkage discriminant analysis; *AUC*, area under the receiver operating characteristic curve; C5, C5.0; CEN, nearest shrunken centroids; *GRADBOOST*, stochastic gradient boosting; *KNN*, k-nearest neighbors; *MLP*, multilayer perceptron; *NBB*, naive Bayes; *RF*, Random Forest; *RLM*, penalized multinomial regression; *SVM*, support vector machines with radial basis function kernel; *TAN*, tree-augmented naive Bayes classifier; *TAN_SEARCH*, tree-augmented naive Bayes classifier structure learner wrapper; *XGB*, extreme gradient boosting.

and the other to predict mosaicism. In both cases, multiclass classification models were chosen, in which the variable to be predicted had 4 categories depending on whether the identified alteration was present or absent and whether it affected a complete chromosome or only a part (segmental).

The metric to optimize was the AUC, which quantifies the predictive power of the different models. Figure 1 shows the results of the best models for each of the trained algorithms. The results are detailed for the different model quality metrics: AUC (macro), mean sensitivity, mean specificity, and accuracy.

For both aneuploidies (Figure 1, A) and mosaicism (Figure 1, B), the model with the highest AUC value was the Random Forest (aneuploidies AUC, 0.780; mosaicism AUC, 0.744), although statistically equivalent (t test) to other models. In the case of aneuploidy, the equivalent models to Random Forest were XGBoost (AUC, 0.780), C5 (AUC, 0.762), and K-Nearest-Neighbor (AUC, 0.725). For mosaicism the best models were, in addition to the above-mentioned Random Forest, XGBoost (AUC, 0.736), Support Vector Machine (AUC, 0.723), C5 (AUC, 0.715), and Gradient Boosting (AUC, 0.692).

When analyzing the other metrics such as accuracy, mean sensitivity, and mean specificity, the final Random Forest model was also the one that presented the highest value, although statistically equivalent to the models mentioned above, and therefore, it was chosen as the final model.

To improve the results, the final model for both aneuploidy and mosaicism was achieved using the Random Forest algorithm, which allows the entire database to be used without the need to split it between training and validation data. It also allows the validation error to be calculated from the observations that have not been used when training the model (OOB) (Supplementary Figure S2).

The AUC values for the final aneuploidy and mosaicism models were 0.792 and 0.776, respectively, which were better than those initially obtained (0.780 and 0.744). The remaining metrics of both models are summarized in Table 2, B and D. The confusion matrix showing the high accuracy of the final predictive models is shown in Table 2, A and C.

Figure 2 shows the ROC curves for the *Random Forest* model for each of the classes of the variable to be predicted, and the global mean (macro and micro).

Figure 3 shows the relative importance of the variables (in terms of number of trees and minimum depth) in the *Random Forest* models predicting embryo aneuploidy and mosaicism, respectively. The size of the points is a measure of the number of nodes into which the corresponding variable is split. The top 10 variables are highlighted in blue.

For aneuploidies (Figure 3, A), the most important variable was maternal age, followed by embryo quality. Male and female karyotype were also 2 important variables in the final model. Surprisingly, the number of oocytes retrieved after ovarian stimulation was a variable that also achieved a high relative importance, whereas history of

TABLE 2
Confusion matrix and performance metrics of final random forest models for aneuploidy and mosaicism Ortiz. Prediction of aneuploid and mosaic embryos. Am J Obstet Gynecol Glob Rep 2022.

A

	Reference			
Prediction	E	A_CC	A_CS	A_CC_CS
E	725	33	12	13
A_CC	24	731	8	13
A_CS	43	22	784	15
A_CC_CS	48	34	16	819

B

logLoss	AUC	Mean Sensitivity	Mean Specificity	Accuracy
1.119	0.792	0.561	0.854	0.562

C

	Reference			
Prediction	NM	M_CC	M_CS	M_CC_CS
NM	578	23	7	8
M_CC	23	614	9	11
M_CS	19	26	615	9
M_CC_CS	20	21	8	612

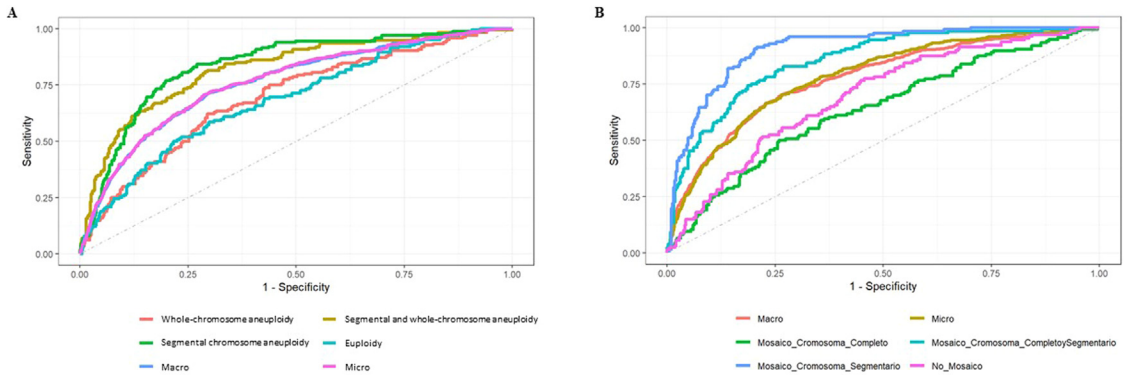
D

logLoss	AUC	Mean Sensitivity	Mean Specificity	Accuracy
1.112	0.776	0.545	0.847	0.542

A: Confusion matrix for the final Random Forest model for the prediction of embryo aneuploidy (multi-class). E: Euploid; A_CC: Aneuploide Cromosoma Completo (Whole-chromosome aneupolidy); A_CS: Aneuploide Cromosoma Segmentario (Segmental chromosome aneuploidy); A_CC_CS: Aneploide Cromosoma Completo y Segmentario (Segmental and whole-crhomosme aneuploidy). **C:** Confusion matrix for the final Random Forest model for the prediction of embryo mosaicism (multi-class). NM: Non-mosaicism; M_CC: Mosaicismo Cromosoma Completo (Whole-chomosome mosaicism); M_CS: Mosaicismo Cromosoma Segmentario (Segmental Chromosome Mosaicism): M_CC_CS: Mosaicismo Cromosoma Completo y Segmentario (Segmental and whole-chromosome mosaicism) **B y D:** Results obtained in terms of logLoss (negative of the multinomial log-likelihood based on the class probabilities), AUC (area under the ROC curve), mean sensitivity, mean specificity and accuracy for each of the machine learning algorithms used in the multiclass predictive model (IIB: aneuploidy; IID: mosaicism).

Ortiz. Prediction of aneuploid and mosaic embryos. Am J Obstet Gynecol Glob Rep 2022.

FIGURE 2
Receiver operating characteristic (ROC) curves for each of the classes of the multiclass response variable



(**A,** aneuploidy; **B,** mosaicism). The multi_ROC function of the multiROC library was used. The overall mean was calculated by 2 different methods (macro and micro).

Ortiz. Prediction of aneuploid and mosaic embryos. Am J Obstet Gynecol Glob Rep 2022.

chromosomal alterations was not among the 10 most important variables.

The PGT-A diagnostic technique was the variable with the highest relative importance in the predictive model of embryo mosaicism, whereas this was not as important in the case of embryo aneuploidy. The 2 variables associated with embryo mosaicism were embryo quality and day of biopsy. In this case, maternal age did not play such a relevant role as in aneuploidies.

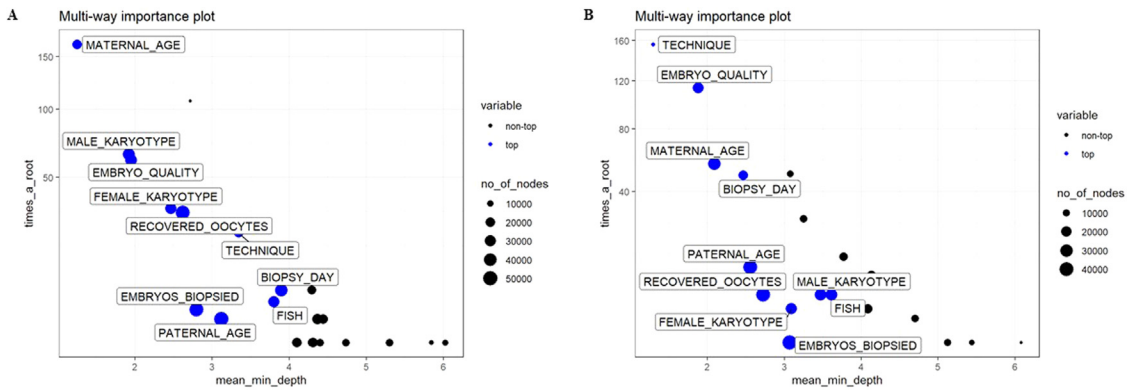
In both models, paternal age was one of the 10 variables with the greatest relative importance, having a greater weight in mosaicism than in aneuploidies. Among the variables associated with the male factor, only sperm aneuploidy (FISH) appeared in both the aneuploidy and mosaicism models. Sperm quality parameters were not among the 10 most important variables in any of the *Random Forest* predictive models.

Variables related to the type of IVF cycle (donated oocyte or sperm, vitrified or fresh oocyte, frozen or fresh sperm, and follicular or luteal phase stimulation) played a minor role in the explanatory models.

Discussion

There has been spectacular progress in reproductive medicine and technology since the first child conceived via IVF was born in 1978. However, presently, the pregnancy rates of couples undergoing reproductive techniques seem to have reached a limit that cannot be surpassed. This is why the scientific community has doubled the efforts to make incremental progress in a variety of stages of the IVF treatment, such as embryo culture media and conditions, protocols and hormones used in ovarian stimulation,

FIGURE 3
Relative importance of the variables in terms of the number of trees and the minimum depth of the Random Forest predictive model



Relative importance of the variables in terms of the number of trees and the minimum depth of the Random Forest predictive model (**A,** aneuploidy; **B,** mosaicism). The size of the point is proportional to the total number of nodes into which it is finally divided in that variable.

Ortiz. Prediction of aneuploid and mosaic embryos. Am J Obstet Gynecol Glob Rep 2022.

endometrial preparation methods, and embryo selection.

Artificial intelligence applied to other fields of medicine, supported by big data, is allowing analysis of data that goes beyond basic statistical approaches and is able to make predictions that help clinicians increase their diagnostic and therapeutic abilities.[26−28]

Machine learning applied to reproductive medicine will provide specialized gynecologists with data analysis methods that will allow them to maximize the potential of ART.[29] These data analysis methods represent an advance in personalized medicine because they facilitate knowledge about the contribution of each of the potential predictor variables to the outcome of a fertility treatment, and enable individually tailoring protocols for each patient.

Different machine learning and deep learning algorithms are being applied in various processes of reproductive medicine, mainly in embryo classification, supporting the work of embryologists by helping them to select embryos with the greatest implantation potential.[30−32]

A priori knowledge of a couple's chances of achieving a euploid embryo in ART is of utmost importance in the search for the best therapy for couples attending fertility clinics. Retrospective data analysis using machine learning can be a powerful tool for establishing explanatory models of embryo aneuploidy and mosaicism. There is a wide range of machine learning algorithms with different theoretical foundations that can provide a great variety of these models.

The data that we used in the optimization of the models were obtained from approximately 7000 embryos analyzed by PGT-A (array comparative genomic hybridization [aCGH] or NGS) in the last 8 years. A total of 29 predictor variables were used, which could be classified into 6 groups: general, maternal, paternal, couple-related, IVF cycle−related, and embryo-related.

When elaborating the predictive models for embryo aneuploidy and mosaicism, we used different strategies. At first, they were used as binary variables for prediction (euploid vs aneuploid; mosaic vs nonmosaic), but the explanatory capacity of the obtained models was very low (unpublished data). The final multiclass models had better quality compared with those obtained by strategies analyzing aneuploidy and mosacism individually. The multiclass analysis results in AUC were close to 80%, which were much higher values than those obtained with the individual strategies (AUC, 65%) and of objectively good quality. A biological explanation for this result is that the origins of segmental and whole-chromosome aneuploidies and mosaicism are not identical, and therefore grouping them in the same class is not an optimal solution in the search for a predictive model. The different types of aneuploidy and mosaicism require separate categories.

For both aneuploidy and mosaicism, the best model was achieved with the Random Forest algorithm, which is based on the combination of classification trees generated from a random sampling of the starting data and a random selection of variables before evaluating each split. Random Forest has been widely used in predictive models in the field of medicine.[33−35] It can reduce the variance in predictions by averaging the set of models generated. The number of randomly sampled variables in each split is a hyperparameter (mtry) to optimize in the training of predictive models. Specifically, the model is evaluated on different combinations of parameters included in a grid (mtry values ranging from 1−100, at intervals of 1 unit). In the predictive model for embryo aneuploidy, the optimal value of mtry that maximized the AUC was 8, and in the case of the mosaicism model it was 7 (Supplemental Fig. S1).

In the Random Forest algorithm, it is possible to indirectly estimate the validation error without the need for cross-validation or an independent test sample. This is why the initial sample has not been split into a training and a validation sample. The response of the excluded observations (OOB) can be predicted when generating and averaging the classification trees. If the number of trees is sufficiently high, the OOB error is comparable to the validation error. In this way, the model is adjusted and validated while different models are trained. Supplemental Figure S1 shows the values of the OOB error according to the hyperparameter mtry (number of predictors used).

The relative importance of the variables used to predict embryo aneuploidy and mosaicism was assessed by the number of trees in which the variable was included and the minimum depth of the variable in the tree. The minimum depth allows to determine the importance of the variable by its position in the classification tree, thus variables that tend to split near the root node should have more importance in the prediction. The size of the point is proportional to the total number of nodes into which it eventually splits on that variable (Figure 3).

Although it is important to highlight that what is quantified is the influence of the predictors on the model and not their relationship with the response variable, it seems to be of interest to analyze which variables have greater importance in the 2 predictive models obtained, and the comparison between them.

Many of the most influential variables in the aneuploidy model had been previously described as being related to an increased risk of chromosomal alterations in the embryo. The most important variable in this predictive model is maternal age. It is well known that the rate of embryonic aneuploidy increases with maternal age. As ovarian aging progresses, incorrect meiosis during ovogenesis starts occurring, with improperly segregated chromosomes, resulting in unbalanced oocytes that produce aneuploid embryos.[5,6]

Among the embryo-related variables, embryo quality and the day of the biopsy were among the most relevant ones in the predictive model. It has been described that embryos of better quality (A and B) have a lower aneuploidy rate[9]; therefore, embryo quality, at the morphologic level, is a good prognostic factor for correct chromosomal endowment. Likewise, embryos

that have been biopsied on day 5 of development and that have consequently had a normal evolution have lower aneuploidy rates compared with those with slower growth that have been biopsied on day 6 or 7 of embryo development.

The number of embryos biopsied also appears among the important variables, and is to a certain extent a measure of good-quality embryos at days 5, 6, and 7 and therefore susceptible to PGT-A analysis.

Another group of important variables in the explanatory model of embryo aneuploidy is the karyotype of both the male and the female. It is widely acknowledged in the literature that chromosomal alterations in the karyotype increase the proportion of aneuploid embryos.[7]

Regarding the male factor, among the important variables we found sperm aneuploidy (FISH) and, surprisingly, paternal age. Different articles have proposed the implication of sperm aneuploidy with an increased risk of embryo aneuploidy, although other studies ruled it out.[10–13] Two recent publications have analyzed the possible correlation between paternal age and chromosomal aberrations in the embryo. Carraquillo et al[36] observed no association between paternal age and chromosomal alterations in blastocyst biopsies from egg donor cycles, whereas Dviri et al[37] observed that segmental alterations are increase with paternal age.

It is surprising that sperm quality variables do not appear among the 10 most important variables in the predictive model of aneuploidy. Previous studies have pointed to an important role of different sperm parameters in embryo aneuploidy, especially teratozoospermia.[38–41]

The most important variable in the machine learning predictive model of embryo mosaicism was the type of diagnostic technique used for PGT-A. NGS has a higher diagnostic sensitivity than aCGH for the detection of embryo mosaicism.[42] However, diagnostic sensitivity does not seem to be as technique-dependent in the case of aneuploidies.

There is scarce literature on the type of factors associated with mosaicism, and therefore there is no consensus. In regard to the most important variables in the final models, there was quite a lot of similarity with aneuploidies, although the relative position varied in some cases. Embryo quality and the day of the biopsy continue to be important variables, but paternal and maternal karyotypes were moved to the last positions in the ranking of the top 10 variables. These 2 variables had not been previously proposed as predictors of embryo mosaicism.

For mosaicism, paternal and maternal variables had nearly equal importance in the final selected model, with paternal and maternal age being the most important variables. Studies of our own group had already noted the association of these variables with embryo mosaicism (unpublished data).

Among the important variables we also found the number of biopsied embryos and the number of retrieved oocytes, which are related to the results of ovarian stimulation. Different studies have analyzed the possible association between the different parameters of ovarian stimulation and aneuploidy and mosaicism, with contradictory conclusions in some cases.[15]

With this study, we have described predictive models for embryo aneuploidy and mosaicism that can be of great relevance in establishing the best protocol in personalized assisted reproduction treatment for maximizing the chances of achieving an euploid embryo.

The knowledge of predictive variables for aneuploidy and mosaicism can be of great importance in the field of reproductive medicine. It can also be a very useful tool for clinicians to offer patients different therapeutic alternatives (eg, own or donated oocytes or sperm, stimulation in the luteal or follicular phase) to achieve euploid embryos, thus minimizing the possibility of miscarriage and increasing the probability of ongoing pregnancy.

This study is not intended to be an endpoint; we intend to carry out a prospective study in which we can confirm the goodness of our model and assess if there is need to increase the number of embryos analyzed or to incorporate additional variables in the model.

Highlights

There are paternal, maternal, embryo, and in vitro fertilization–related factors associated with embryonic chromosomal status that can be used as predictors in machine learning models.

The variable for predicting aneuploidies and mosaicisms used in the machine learning models was multiclass, with 4 categories based on whether the identified alteration was present or absent and whether it affected the whole or only a part of the chromosome (segmental).

The best prediction model for both aneuploidy and mosaicism was obtained with the Random Forest algorithm. ■

ACKNOWLEDGMENTS

We are grateful to Virginia Alarcón, MS, Anastasia Utkina, MS, and Anna Pitas, PhD for their valuable assistance in the preparation and review of the English manuscript.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.xagr.2022.100103.

REFERENCES

1. Fragouli E. Next generation sequencing for preimplantation genetic testing for aneuploidy: friend or foe? Fertil Steril 2018;109:606–7.
2. Alfarawati S, Fragouli E, Colls P, et al. The relationship between blastocyst morphology, chromosomal abnormality, and embryo gender. Fertil Steril 2011;95:520–4.
3. Hodes-Wertz B, Grifo J, Ghadir S, et al. Idiopathic recurrent miscarriage is caused mostly by aneuploid embryos. Fertil Steril 2012;98:675–80.
4. MacLennan M, Crichton JH, Playfoot CJ, Adams IR. Oocyte development, meiosis and aneuploidy. Semin Cell Dev Biol 2015;45:68–76.
5. Baart EB, Martini E, van den Berg I, et al. Preimplantation genetic screening reveals a high incidence of aneuploidy and mosaicism in embryos from young women undergoing IVF. Hum Reprod 2006;21:223–33.

6. Babariya D, Fragouli E, Alfarawati S, Spath K, Wells D. The incidence and origin of segmental aneuploidy in human oocytes and preimplantation embryos. Hum Reprod 2017; 32:2549–60.
7. Lledó B, Ortiz JA, Morales R, et al. The paternal effect of chromosome translocation carriers observed from meiotic segregation in embryos. Hum Reprod 2010;25:1843–8.
8. Morales R, Lledó B, Ortiz JA, Ten J, Llácer J, Bernabeu R. Chromosomal polymorphic variants increase aneuploidies in male gametes and embryos. Syst Biol Reprod Med 2016; 62:317–24.
9. Minasi MG, Colasante A, Riccio T, et al. Correlation between aneuploidy, standard morphology evaluation and morphokinetic development in 1730 biopsied blastocysts: a consecutive case series study. Hum Reprod 2016;31:2245–54.
10. Rodrigo L, Meseguer M, Mateu E, et al. Sperm chromosomal abnormalities and their contribution to human embryo aneuploidy. Biol Reprod 2019;101:1091–101.
11. Lu S, Zong C, Fan W, et al. Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. Science 2012;338:1627–30.
12. Wang J, Fan HC, Behr B, Quake SR. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. Cell 2012;150:402–12.
13. Asoglu MR, Celik C, Serefoglu EC, Findikli N, Bahceci M. Preimplantation genetic testing for aneuploidy in severe male factor infertility. Reprod Biomed Online 2020;41:595–603.
14. Munné S, Sandalinas M, Magli C, Gianaroli L, Cohen J, Warburton D. Increased rate of aneuploid embryos in young women with previous aneuploid conceptions. Prenat Diagn 2004;24:638–43.
15. Cascales A, Lledó B, Ortiz JA, et al. Effect of ovarian stimulation on embryo aneuploidy and mosaicism rate. Syst Biol Reprod Med 2021;67:42–9.
16. Kemper JM, Vollenhoven BJ, Talmor AJ. Preimplantation genetic testing for aneuploidy: a review. Obstet Gynecol Surv 2019;74:727–37.
17. Taylor TH, Gitlin SA, Patrick JL, Crain JL, Wilson JM, Griffin DK. The origin, mechanisms, incidence and clinical consequences of chromosomal mosaicism in humans. Hum Reprod Update 2014;20:571–81.
18. Bolton H, Graham SJL, Van der Aa N, et al. Mouse model of chromosome mosaicism reveals lineage-specific depletion of aneuploid cells and normal developmental potential. Nat Commun 2016;7:11165.
19. Munné S, Velilla E, Colls P, et al. Self-correction of chromosomally abnormal embryos in culture and implications for stem cell production. Fertil Steril 2005;84:1328–34.
20. Coll L, Parriego M, Mateo S, et al. Prevalence, types and possible factors influencing mosaicism in IVF blastocysts: results from a single setting. Reprod Biomed Online 2021; 42:55–65.
21. Swain JE. Controversies in ART: can the IVF laboratory influence preimplantation embryo aneuploidy? Reprod Biomed Online 2019; 39:599–607.
22. Vera-Rodríguez M, Michel CE, Mercader A, et al. Distribution patterns of segmental aneuploidies in human blastocysts identified by next-generation sequencing. Fertil Steril 2016;105. 1047–55.e2.
23. Jenkins J, van der Poel S, Krüssel J, et al. Empathetic application of machine learning may address appropriate utilization of ART. Reprod Biomed Online 2020;41:573–7.
24. Wang R, Pan W, Jin L, et al. Artificial intelligence in reproductive medicine. Reproduction 2019;158:R139–54.
25. Cram DS, Leigh D, Handyside A, et al. PGDIS position statement on the transfer of mosaic embryos 2019. Reprod Biomed Online 2019;39. e1–4.
26. Zheng C, Johnson TV, Garg A, Boland MV. Artificial intelligence in glaucoma. Curr Opin Ophthalmol 2019;30:97–103.
27. Nensa F, Demircioglu A, Rischpler C. Artificial intelligence in nuclear medicine. J Nucl Med 2019;60. 29S–37.
28. Dias R, Torkamani A. Artificial intelligence in clinical and genomic diagnostics. Genome Med 2019;11:70.
29. Raef B, Ferdousi R. A review of machine learning approaches in assisted reproductive technologies. Acta Inform Med 2019;27:205–11.
30. Raef B, Maleki M, Ferdousi R. Computational prediction of implantation outcome after embryo transfer. Health Informatics J 2020; 26:1810–26.
31. Chavez-Badiola A, Flores-Saiffe Farias A, Mendizabal-Ruiz G, Garcia-Sanchez R, Drakeley AJ, Garcia-Sandoval JP. Predicting pregnancy test results after embryo transfer by image feature extraction and analysis using machine learning. Sci Rep 2020;10: 4394.
32. Chavez-Badiola A, Mendizabal-Ruiz G, Flores-Saiffe Farias A, Garcia-Sanchez R, Drakeley AJ. Deep learning as a predictive tool for fetal heart pregnancy following time-lapse incubation and blastocyst transfer. Hum Reprod 2020;35:482.
33. Zukotynski K, Gaudet V, Kuo PH, et al. The use of random forests to identify brain regions on amyloid and FDG PET associated with MoCA score. Clin Nucl Med 2020;45:427–33.
34. Esmaily H, Tayefi M, Doosti H, Ghayour-Mobarhan M, Nezami H, Amirabadizadeh A. A comparison between decision tree and random Forest in determining the risk factors associated with type 2 diabetes. J Res Health Sci 2018;18:e00412.
35. Rasti R, Mehridehnavi A, Rabbani H, Hajizadeh F. Automatic diagnosis of abnormal macula in retinal optical coherence tomography images using wavelet-based convolutional neural network features and random forests classifier. J Biomed Opt 2018;23:1–10.
36. Carrasquillo RJ, Kohn TP, Cinnioglu C, et al. Advanced paternal age does not affect embryo aneuploidy following blastocyst biopsy in egg donor cycles. J Assist Reprod Genet 2019;36:2039–45.
37. Dviri M, Madjunkova S, Koziarz A, et al. Is there a correlation between paternal age and aneuploidy rate? An analysis of 3,118 embryos derived from young egg donors. Fertil Steril 2020;114:293–300.
38. Vendrell JM, García F, Veiga A, et al. Meiotic abnormalities and spermatogenic parameters in severe oligoasthenozoospermia. Hum Reprod 1999;14:375–8.
39. Gole LA, Wong PF, Ng PL, Wang XQ, Ng SC, Bongso A. Does sperm morphology play a significant role in increased sex chromosomal disomy? A comparison between patients with teratozoospermia and OAT by FISH. J Androl 2001;22:759–63.
40. Burrello N, Arcidiacono G, Vicari E, et al. Morphologically normal spermatozoa of patients with secretory oligo-astheno-teratozoospermia have an increased aneuploidy rate. Hum Reprod 2004;19:2298–302.
41. Brahem S, Mehdi M, Elghezal H, Saad A. Study of aneuploidy rate and sperm DNA fragmentation in large-headed, multiple-tailed spermatozoa. Andrologia 2012;44:130–5.
42. Goodrich D, Xing T, Tao X, et al. Evaluation of comprehensive chromosome screening platforms for the detection of mosaic segmental aneuploidy. J Assist Reprod Genet 2017; 34:975–81.