Ten *et al.*
Reproductive Biology and Endocrinology (2024) 22:116
https://doi.org/10.1186/s12958-024-01285-9

Reproductive Biology and Endocrinology

RESEARCH

Open Access

Enhancing predictive models for egg donation: time to blastocyst hatching and machine learning insights

Jorge Ten[1*], Leyre Herrero[2], Ángel Linares[1], Elisa Álvarez[3], José Antonio Ortiz[4], Andrea Bernabeu[5] and Rafael Bernabéu[5]

Abstract

Background Data sciences and artificial intelligence are becoming encouraging tools in assisted reproduction, favored by time-lapse technology incubators. Our objective is to analyze, compare and identify the most predictive machine learning algorithm developed using a known implantation database of embryos transferred in our egg donation program, including morphokinetic and morphological variables, and recognize the most predictive embryo parameters in order to enhance IVF treatments clinical outcomes.

Methods Multicenter retrospective cohort study carried out in 378 egg donor recipients who performed a fresh single embryo transfer during 2021. All treatments were performed by Intracytoplasmic Sperm Injection, using fresh or frozen oocytes. The embryos were cultured in Geri® time-lapse incubators until transfer on day 5. The embryonic morphokinetic events of 378 blastocysts with known implantation and live birth were analyzed. Classical statistical analysis (binary logistic regression) and 10 machine learning algorithms were applied including Multi-Layer Perceptron, Support Vector Machines, k-Nearest Neighbor, Cart and C0.5 Classification Trees, Random Forest (RF), AdaBoost Classification Trees, Stochastic Gradient boost, Bagged CART and eXtrem Gradient Boosting. These algorithms were developed and optimized by maximizing the area under the curve.

Results The Random Forest emerged as the most predictive algorithm for implantation (area under the curve, AUC = 0.725, IC 95% [0.6232–0826]). Overall, implantation and miscarriage rates stood at 56.08% and 18.39%, respectively. Overall live birth rate was 41.26%. Significant disparities were observed regarding time to hatching out of the zona pellucida ($p = 0.039$). The Random Forest algorithm demonstrated good predictive capabilities for live birth (AUC = 0.689, IC 95% [0.5821–0.7921]), but the AdaBoost classification trees proved to be the most predictive model for live birth (AUC = 0.749, IC 95% [0.6522–0.8452]). Other important variables with substantial predictive weight for implantation and live birth were duration of visible pronuclei (DESAPPN-APPN), synchronization of cleavage patterns (T8-T5), duration of compaction (TM-TiCOM), duration of compaction until first sign of cavitation (TiCAV-TM) and time to early compaction (TiCOM).

Conclusions This study highlights Random Forest and AdaBoost as the most effective machine learning models in our Known Implantation and Live Birth Database from our egg donation program. Notably, time to blastocyst hatching out of the zona pellucida emerged as a highly reliable parameter significantly influencing our implantation

*Correspondence:
Jorge Ten
jten@institutobernabeu.com
Full list of author information is available at the end of the article

machine learning predictive models. Processes involving syngamy, genomic imprinting during embryo cleavage, and embryo compaction are also influential and could be crucial for implantation and live birth outcomes.

Keywords Artificial intelligence, Machine learning, Embryo morphokinetic, Hatching blastocyst, Implantation and live birth

Introduction

Selecting the best embryo for transfer is one of the most important tasks of every embryologist. Since the beginning of the IVF in the 70´s, visual static assessment of embryo quality under microscope led to different embryo classifications, according to embryo implantation potential, such as Gardner and Schoolcraft´s [1] which is widely used in the IVF laboratories.

Nowadays, embryo selection has changed and the final goal of every IVF cycle is the birth of a healthy baby. Data sciences and artificial intelligence (AI) are becoming encouraging tools in medicine, also in assisted reproduction, where the amount of data and images generated in the IVF laboratories has dramatically increased, favored by time-lapse technology incubators currently introduced in almost every IVF clinic. Static morphological quality assessment of the blastocysts is now supported by morphokinetic development parameters that have demonstrated statistical connection with embryo implantation and live birth [2–5].

Based on those morphokinetic parameters, either manual or automatically annotated, multiple blastocyst selection algorithms have been developed using machine learning (ML) or deep learning (DL) AI technology in the recent years to predict blastocyst implantation [6–9], blastocyst ploidy [10–15] or even live birth [16–20]. Convolutional neural networks (CNNs) are currently the most popular DL models employed in embryo assessment since they have been trained using a large amount of data sets of static images of embryos and can accurately predict the developmental potential of embryos based on morphokinetic parameters and image analysis [21].

AI algorithms have the big advantage of continuously improving its selection potential as the input data grows, being trained as long as we use them. However, when comparing the different currently available algorithms they result in different conclusions, even when they were trained on the same data, highlighting the importance of well-designed randomized control trials (RCT) to prove the real usefulness of them [22]. Furthermore, different models have not reached the same results among different laboratories [23] demonstrating that external factors, such as laboratory conditions, embryologist expertise or other human factors, influence embryo development and quality and also clinical outcomes, and must be taken into account when proposing the application of a universal algorithm [24].

Based on this information, we aim to analyze and compare different ML algorithms developed using a known implantation and live birth database (KILBD) of embryos transferred in our egg donation program, including morphokinetic and morphological variables. Hence, the aim of this investigation was to pinpoint the most prognostic algorithm and the most predictive embryo parameters delineated through this algorithm, with the aim of enhancing the clinical outcomes of our egg donation program.

Material and methods

Study population

This is a retrospective multicenter cohort study that included 378 egg donor recipients who performed single embryo transfer (SET) during 2021 in four Instituto Bernabéu clinics in Spain (Instituto Bernabéu Alicante, IBA; Instituto Bernabeu Madrid, IBM; Instituto Bernabeu Albacete, IBAB; and Instituto Bernabeu Palma de Mallorca, IBPM). All treatments were performed by Intracytoplasmic Sperm Injection (ICSI), using fresh or frozen oocytes, and fresh embryo transfers on day 5.

Exclusion criteria for the egg recipients were: uterine pathology (myomas, polyps, adenomyosis or any uterine malformation), recurrent implantation failure (three or more embryos transferred without pregnancy), recurrent miscarriage (two or more previous miscarriages on the first trimester), and body mass index (BMI) > 40. Semen samples < 1 million per milliliter (mill/ml) were also excluded.

The study conformed to the Declaration of Helsinki for Medical Research about human subjects and Spanish Data Protection Act (LO3/2018) and was approved by the Institutional Review Board on January 2021 (reference number BR20/2021). All participants provided written informed consent to participate in the study.

Oocyte donors' ovarian stimulation

Conventional antagonist short protocol was used for egg donor controlled ovarian stimulation as described previously by our team [25]. When three or more follicles of 18 mm diameter were observed, a bolus of GnRH analogues was administered to trigger ovulation in the following 36 to 38 h.

Endometrial preparation of recipients

Recipients received hormone replacement therapy for endometrial preparation through the administration of estrogens via transdermal patches or oral tablets. Vaginal progesterone was administered starting from the night of oocyte retrieval and ICSI, and continued until the 12th week of gestation. This protocol was previously described by our group [26].

Oocyte manipulation and embryo culture

After egg retrieval, oocytes were incubated in Global® Total® for fertilization medium (LifeGlobal, Coopersurgical, Ballerup, Denmark) for 4 h, until ICSIs were carried out, or 3 h post oocyte thawing in frozen oocytes cycles. In this case, metaphase II (MII) oocytes were denudated and vitrified 1 h post egg retrieval, and thawed according to Kitazato® protocol [27, 28].

After ICSI, oocytes were cultured in Global® Total® medium (LifeGlobal, Coopersurgical, Ballerup, Denmark) in Geri® time-lapse incubators (Genea Biomedx, Sydney, Australia) until embryo transfer on day 5. The culture conditions were 6% O_2, 7% CO_2, 37°C.

Oocyte fertilization, embryo development, and morphokinetics were evaluated utilizing computer software (Geri Connect®, Genea Biomedx, Sydney, Australia). In this process, all morphological characteristics, measurements, and morphokinetic parameters were meticulously annotated by a single, trained embryologist stationed at the study coordination center (IBA).

Blastocyst selection and embryo transfer

Day 5 embryo selection for transfer was based on the ASEBIR criteria [29]. Selected blastocysts were transferred to the uterus 118–122 h post ICSI on day 5, using ultrasound guidance and the Rocket catheter (Medical, Washington, USA). We employed the same transfer technique and the same catheter across all four clinics. After transfer, the remaining embryos exhibiting good quality (types A and B) were cryopreserved using the Kitazato® protocol.

Clinical outcomes

The pregnancy test (β-hCG level) was assessed in blood samples taken 13 days post-ICSI. Clinical pregnancies were subsequently verified via vaginal ultrasound two weeks into gestation. At this stage, the ratio of detected sacs to the number of embryos transferred was calculated as the implantation rate. Pregnancy progress was monitored via ultrasonography, with first-trimester miscarriages, defined as intrauterine losses before reaching 12 weeks' gestational size on ultrasound, meticulously documented. Finally, the rate of live births per embryo transferred was also documented.

Study design and variables

This study was divided into two phases:

> Part one: consisting of revision of the embryo development videos of the transferred embryos and manual annotation of the morphokinetic parameters (in hours post ICSI) of each transferred embryo by one senior embryologist, designated to avoid inter-observer bias. This was done at the study coordination center. Although to mitigate observation biases ideally annotations should have been performed by two operators, with a third in case of disagreement, this was not feasible as the study was conducted by a limited number of embryologists, with the expert embryologist responsible for morphokinetic annotations at the central site performing this task.

The classic morphokinetic parameters annotated, classic development time intervals and morphological parameters included in KILBD are described in Fig. 1. The diameter of the blastocyst and the inner cell mass (ICM) were measured 110 h after ICSI and included as morphological variables in the study.

> Part two: consisted of statistical analysis and development of predictive models (algorithms) for implantation and live birth through ML, as explained below.

Data pre-processing

The database was randomly divided into a training set (80% of the database) and a test set (20% of the database). The datasets were subjected to the following stages of pre-processing before the ML algorithms development:

1. Imputation of missing values (less than 0.1% of the data). In our case, we perform a multivariate imputation of missing values, using the Multiple Imputation by Chained Equations method (mice) algorithm. We chose the imputation using the classification trees (cart) among the different options of the algorithm.
2. Outlier's analysis. These potential outliers were identified through the box-plots. These potential outliers were reviewed and checked. Finally, it was decided to keep them as they were far from the median but real values. No treatment of the outliers was necessary.
3. Balance database. Imbalanced datasets (the disparity in cases between the categories of the response variable is significant) pose a common challenge for machine learning practitioners in binary classifica-

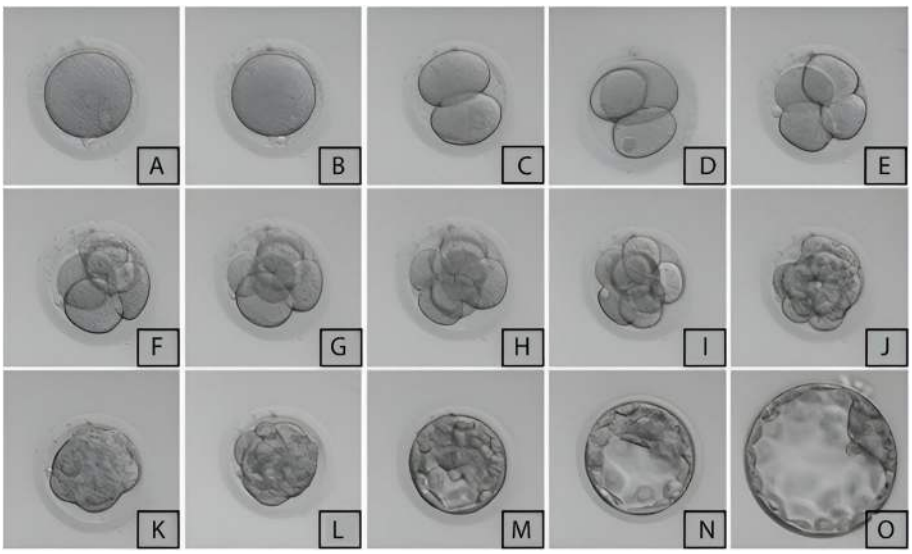

Fig. 1 Variables included in the ML algorithms: Time of appearance of the two pronuclei (APNN) (**A**), time of pronuclei fading (DESAPPN) (**B**), time to two cells (T2) (**C**), time to three cells (**D**) (T3), time to four cells (T4) (**E**), time to five cells (T5) (**F**), time to six cells (T6) (**G**), time to seven cells (T7) (**H**), time to eight cells (T8) (**I**), time to early compaction (TiCOM) (**J**), time to compacted morula formation (TM) (**K**), time to blastulation first sign (TiCAV) (**L**), time to full blastocyst (TFB) (**M**), time to expanded blastocyst (TEB) (**N**) and time to hatching out the zona pellucida (TiH) (**O**)

tion problems, as this situation can distort the model learning process by giving priority to the majority class and leading to poor performance in the minority class. To address this problem, one popular technique is the Synthetic Minority Oversampling Technique (SMOTE). SMOTE is specifically designed to tackle imbalanced datasets by generating synthetic samples for the minority class.

4. Selection of features. Given the large number of predictors, it is necessary to select the most relevant ones. This process has been carried out using different methods:

   4a. Elimination of correlated variables (cut-off point 0.8). To avoid collinearity, the correlation between the different predictors was analysed. In those cases where the correlation levels were very high, one of the two variables was eliminated. We considered as high correlation those cases in which the Pearson correlation coefficient was greater in absolute value than 0.8 and, therefore, very close to the maximum value, which is 1.

   4b. Elimination of variables with variance near to 0. Variables with variances equal to or very close to zero are variables with very constant values for all embryos in our sample and are therefore poor predictors.

   4c. Selection of predictor variables by applying different methods:

   - Boruta: It is an extension of the random forest algorithm that uses permutation tests to assess the importance of each feature and determine whether it is relevant or not.
   - Recursive feature elimination: This algorithm creates different training and test sets in which it will incorporate different sets of predictors and evaluates the reduction of the mean square error.
   - Genetic algorithm: This algorithm is inspired by the process of biological evolution for the selection of the best predictors.

After the selection, ten variables were identified as the most important, as they appeared to have the highest predictive capacity and were used to train the models. Therefore, the models were selected to maximize the AUC value. It is noteworthy that all models rely on identical predictor variables. The significance of these variables is contingent upon the target variable (implantation or live birth), dictating their importance within the predictive framework.

Hyperparameter optimization of classification models

A classical statistical analysis method such as binary logistic regression and 10 machine learning algorithms (classification) were applied including Multi-Layer Perceptron, Support Vector Machines, k-Nearest Neighbor, Cart and C0.5 Classification Trees, Random Forest

(RF), AdaBoost Classification Trees, Stochastic Gradient boost, Bagged CART and eXtrem Gradient Boosting.

The hyperparameters were fitted during the training process. These hyperparameters directly control the structure and function of each of the models. By adjusting them during the training process, model performance can be modified to achieve the best results.

To ensure independence between the data, the five-fold cross-validation technique with adjustment of the different hyperparameters was applied. During the optimization process, different performance metrics were calculated: AUC (area under the ROC curve), sensitivity, specificity, positive predictive value and negative predictive value, Accuracy and the Kappa statistic.

The selection of the best model for implantation and live birth rates was based primarily on the AUC, a parameter that measures the model's ability to discriminate the dependent variable. The final values of AUC and the other metrics were obtained from the test set.

Statistical analysis

The descriptive statistical methods used in this study depend on the type of the variable analyzed. In the case of qualitative variables, the following descriptive statistics will be obtained: frequency and percentage. Continuous variables were presented as number of cases, mean and SD or median and interquartile range (IQR), as applicable. For the univariate statistical analysis of qualitative variables, the Chi-square test or Fisher's exact test will be used. The Shapiro–Wilk tests were used to assess whether the continuous variables were normally distributed. Depending on whether the variable has a normal distribution, the comparison between means was carried out using Student's t test or Wilcoxon rank sum test.

The software used to carry out the analysis was SPSS (20.0.0, Inc., Chicago, IL, USA) and R (4.0.5). Significance was defined as $p < 0.05$.

Results

Egg recipient's average age was 41.68 ± 3.98 years (ranged 26 to 50). From the 378 embryos transferred, 203 blastocysts came from vitrified oocytes (53.7%).

The global average number of donated oocytes was 10.96 ± 1.68, with an 89.02% rate of mature oocytes and an 83.82% fertilization rate. The survival rate of frozen-thawed oocytes was 89.0%. The fertilization rate in frozen oocytes was slightly lower compared to fresh oocytes (82.94% vs. 84.74%, respectively), although this difference was not statistically significant ($p = 0.124$). To further investigate the potential effect of oocyte vitrification, an in-depth analysis is presented in Supplementary Tables 1, 2, and 3. These tables illustrate the impact of oocyte freezing/thawing by comparing embryos derived from vitrified/warmed oocytes with those from fresh oocytes. No significant differences were found in any of the parameters analyzed.

The inclusion of both fresh and frozen/thawed oocytes was due to the logistics of our oocyte donation program. At our main site (IBA), treatments with fresh oocytes and vitrification for distribution to other sites (IBM, IBAB, IBPM) are performed. Out of the 378 embryos evaluated, 168 embryos were transferred in IBA, 159 in IBM, 12 in IBAB, and 39 in IBPM. There were no differences in embryonic yield and clinical outcomes among the sites (data not shown).

The implantation rate was 56.08% (212 gestational sacs) and the miscarriage rate was 18.39%. The overall live birth rate was 41.26% (156 babies).

Statistical description of population baseline characteristics was performed and summarized in Table 1 for implantation and in Table 2 for live birth. The age of the recipient was the only significant variable between groups, which was significantly lower in the group with positive implantation (Table 1), so it was included as confounding factor when comparing the rest of the parameters in the study.

Average values of continuous and categorized morphokinetic and morphological parameters, and comparison between study groups, are detailed in the Supplementary tables 4 and 5 for implantation and Supplementary tables 6 and 7 for live birth. Statistical differences were only found for implantation analysis between time to hatching out of the zona pellucida ($p = 0.039$) (Supplementary Table 5). No significant differences were found regarding morphokinetic parameters between embryos that achieve live birth (Supplementary tables 6 and 7).

From the blastocysts population that implanted, 57.07% (121/212) were initiating hatching, while from those embryos that did not, only 43.37% (72/166) began the hatching process (p 0.093).

After data pre-processing, ten variables were selected and used to develop the algorithms: duration of visible pronuclei, synchronization of cell division, synchronization of cleavage patterns, time to early compaction, time to blastulation first sign appearance, duration of compaction until first sign of cavitation, duration of compactation, time to hatching out the zona pellucida, duration of blastocyst expansion until hatching and ICM diameter.

Eleven algorithms were developed and optimized by maximizing the AUC for the KILBD (Figs. 2 and 3). The eleven models for both implantation and

Table 1 Baseline characteristics for oocyte recipients and donors, and comparison between embryos that implanted vs those that not:

Parameters	Overall $N=378$[1]	Non-implanted embryos $N=166$[1]	Implanted embryos $N=212$[1]	*p*-value[2]
Years of seeking pregnancy	3.84 (3.21)	3.70 (3.11)	3.95 (3.29)	0.576
Previous pregnancies	0.72 (0.97)	0.71 (1.01)	0.72 (0.94)	0.490
Previous miscarriages	0.48 (0.70)	0.49 (0.74)	0.47 (0.68)	0.921
Number of previous cycles	1.61 (1.74)	1.57 (1.78)	1.63 (1.71)	0.607
Number of previously transferred embryos	1.87 (2.35)	1.74 (2.26)	1.95 (2.41)	0.440
Recipient age	41.68 (3.98)	42.38 (3.64)	41.17 (4.15)	0.021
Paternal age	42.56 (5.63)	43.20 (5.08)	42.10 (5.96)	0.064
Egg donor age	25.68 (4.17)	25.65 (4.10)	25.71 (4.23)	0.953
Recipient BMI	23.40 (4.14)	23.48 (3.92)	23.34 (4.30)	0.504
Egg donor BMI	22.49 (2.91)	22.16 (2.78)	22.74 (2.98)	0.073
Sterility cause: Low responder	12.9%	9.9%	15.2%	0.062
Sterility cause: advanced maternal age	79.8%	86.3%	74.8%	
Sterility cause: unexplained infertility	3.5%	1.9%	4.8%	
Sterility cause: early ovarian failure	3.2%	1.9%	4.3%	
Smoking recipient	13.0%	14.8%	11.7%	0.279
Smoking egg donor	37.9%	42.1%	34.8%	0.254
Endometrial thickness	8.52 (1.49)	8.53 (1.50)	8.51 (1.48)	0.671
Vitrified oocyte	53.7%	57.9%	50.5%	0.150

[1] Mean (SD); %

[2] Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

Table 2 Baseline characteristics for oocyte recipients and donors, and comparison between embryos that achieved a live birth vs those that not

Parameters	Overall $N=378$[1]	Non-Live birth embryos $N=222$[1]	Live birth embryos $N=156$[1]	*p*-value[2]
Years of seeking pregnancy	3.84 (3.21)	3.90 (3.09)	3.76 (3.39)	0.167
Previous pregnancies	0.72 (0.97)	0.71 (0.99)	0.73 (0.94)	0.541
Previous miscarriages	0.48 (0.70)	0.48 (0.69)	0.49 (0.72)	>0.999
Number of previous cycles	1.61 (1.74)	1.49 (1.67)	1.78 (1.82)	0.091
Number of previously transferred embryos	1.87 (2.35)	1.75 (2.25)	2.03 (2.47)	0.265
Recipient age	41.68 (3.98)	41.64 (4.23)	41.74 (3.63)	0.950
Paternal age	42.56 (5.63)	42.62 (5.78)	42.49 (5.42)	0.946
Egg donor age	25.68 (4.17)	25.59 (4.28)	25.81 (4.01)	0.536
Recipient BMI	23.40 (4.14)	23.47 (4.21)	23.29 (4.06)	0.831
Egg donor BMI	22.49 (2.91)	10.6%)	16.2%)	0.141
Sterility cause: Low responder	12.9%	83.4%	74.7%	0.145
Sterility cause: advanced maternal age	79.8%	62.8%	4.5%	
Sterility cause: unexplained infertility	3.5%	3.2%	3.2%	
Sterility cause: early ovarian failure	3.2%	10.6%	16.2%	
Smoking recipient	13.0%	13.8%	11.9%	0.906
Smoking egg donor	37.9%	36.8%	39.5%	0.681
Endometrial thickness	8.52 (1.49)	8.54 (1.47)	8.48 (1.51)	0.660
Vitrified oocyte	53.7%	54.5%	52.6%	0.704

[1] Mean (SD); %

[2] Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test

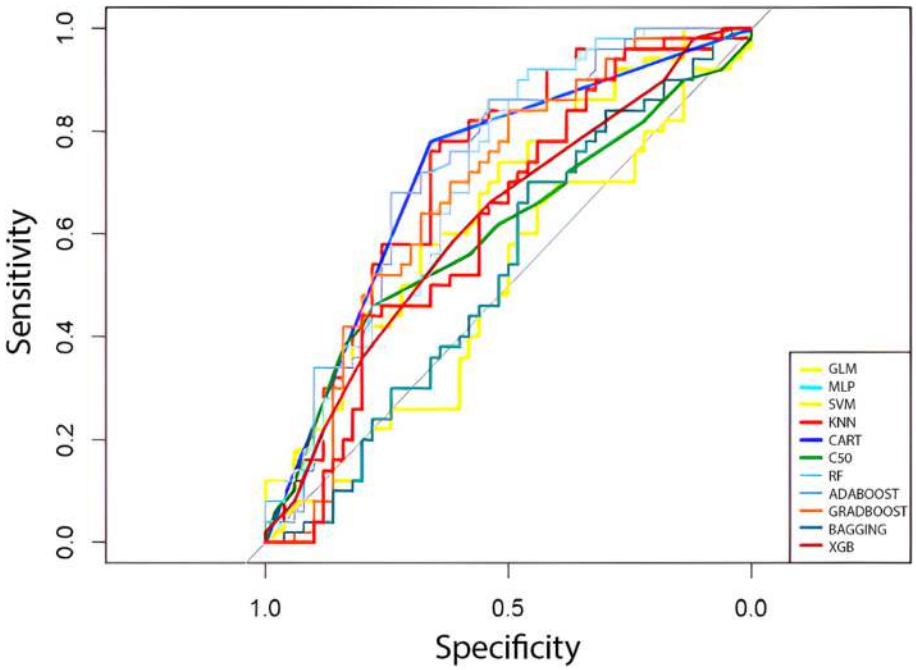

Fig. 2 Receiver operating characteristic curves for implantation prediction and its areas under the curve (AUC) corresponding with the eleven ML models developed: Models: Generalized Linear Model (GLM), Multi-Layer Perceptron (MLP), Support Vector Machines (SVM), k-Nearest Neighbor (KNN), Cart (CART), C0.5 (C0.5), Random Forest (RF), AdaBoost Classification Trees (ADABOOST), Stochastic Gradient boost (GRADBOOOST), Bagged CART (BAGGINS) and eXtrem Gradient Boosting (XGB)

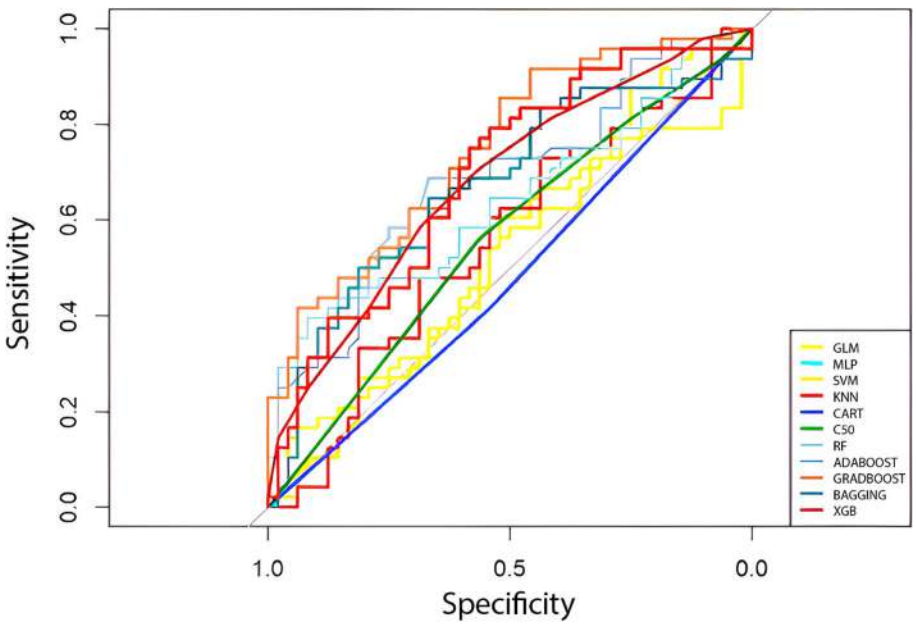

Fig. 3 Receiver operating characteristic curves for prediction of live birth and its areas under the curve (AUC) corresponding with the eleven ML models developed: Models: Generalized Linear Model (GLM), Multi-Layer Perceptron (MLP), Support Vector Machines (SVM), k-Nearest Neighbor (KNN), Cart (CART), C0.5 (C0.5), Random Forest (RF), AdaBoost Classification Trees (ADABOOST), Stochastic Gradient boost (GRADBOOOST), Bagged CART (BAGGINS) and eXtrem Gradient Boosting (XGB)

live birth prediction, and its statistics, are summarized in Tables 3 and 4, respectively. The most predictive algorithms for implantation were: Random Forest (AUC = 0.725, IC 95% [0.6232–0826]), K-nearest Neighbors (AUC = 0.720, IC 95% [0.6319–0.8081]) and Support Vector Machines with Radial Basis Function Kernel (AUC = 0.717, IC 95% [0.6139–08197]) as is described in Fig. 4. Regarding live birth, AdaBoost Classification Trees (AUC = 0.749, IC 95% [0.6522–0.8452]) was the most predictive algorithm, followed by eXtreme Gradient Boosting (AUC = 0691, IC 95% [0.585–0.7978]) and Random Forest (AUC = 0.689, IC 95% [0.5821–0.7921]) (Fig. 5). The variables with the highest predictive weight in RF for implantation were duration of visible pronuclei (DESAPPN-APPN), duration of compaction until the first sign of cavitation (TiCAV-TM), and time to early compaction (TiCOM) (Fig. 4). Conversely, the variables with the highest predictive weight in AdaBoost for live birth were duration of compaction (TM-TiCOM), time to expanded blastocyst (TEB), and duration of blastulation until hatching (TiH-TiCAV) (Fig. 5).

Discussion

The incorporation of artificial intelligence (AI) into embryology represents a promising advancement for enhancing assisted reproduction [21]. Despite its current experimental phase [19, 24], the burgeoning interest is palpable in the scientific literature pertaining to IVF laboratories. Multiple algorithms have been introduced and rigorously assessed, addressing various aspects of the IVF cycle [21, 22, 30–32].

In this research we have developed easy predictive ML algorithms for prediction of implantation and live birth in egg donation programs, that are ideal as models to perform such kinds of studies, as they have consistent population and clinical outcomes. In fact, as in our study, important selection models have been developed in egg donation cycles.

In our KILBD study, the implantation and live birth rates of the donation cycles were 56.08% and 41.26%, respectively, consistent with previous publications in our egg donation program [26, 33].

Baseline characteristics of our patients were similar between groups with the exception of the age of the recipient, which was significant lower in the group with positive implantation (Table 1). However, no significant differences were found after statistical analysis considering the age of the recipient as a confounding factor. Moreover, attending to the similar donor age in both groups, we could consider that this finding has no impact in the clinical results and could be an occasional finding of our study population.

Given that 53.7% of the embryos analyzed in our study originated from vitrified/warmed oocytes, we aimed to investigate the potential impact of the vitrification/warming process on clinical outcomes. As previously mentioned in the results section, the logistics of our donation program prevented us from including cycles exclusively with either fresh or vitrified oocytes. Our examination, as detailed in Supplementary Table 1, revealed that patients from both vitrified and fresh oocyte groups exhibited comparable baseline characteristics. Furthermore, our analysis of morphokinetic and morphological variables, as well as the calculated time intervals incorporated into our KILBD, demonstrated similar values between fresh and frozen oocytes (Supplementary Table 2), ultimately yielding consistent clinical outcomes (Supplementary Table 3). In a recent publication, Montgomery et al., found a significant delay of 2–3 h across all early cleavage divisions (2- through to 8-cell) and time to start of compaction in the vitrified oocyte group versus fresh oocyte controls [34]. However, they did not find differences in the time of reaching the blastocyst stage or in the derived clinical outcomes [34], in accordance with our results. Murria et al., found that embryo scores provided by AI algorithms were lower for embryos originated from vitrified/warmed oocytes than for those that came from fresh oocytes [35]. However, the potential impact on clinical outcomes remains to be fully assessed pending evaluation with a larger sample size of analysed embryos.

Our findings highlight the significant impact of natural hatching on implantation, with 57.01% of implanted blastocysts initiating hatching compared to 43.37% among non-hatching embryos. The statistical analysis revealed a significant difference between implanted and non-implanted embryos regarding the timing of hatching out of the zona pellucida (TiH, $p = 0.039$).

The significance of hatching is essential. At the blastocyst stage, the embryo undergoes two crucial processes: hatching and implantation. These processes are essential for initiating post-implantation development and are influenced by cleavage-stage development, ultimately determining pregnancy outcomes. Any defects in blastocyst hatching and implantation can result in early embryo loss and infertility [36]. More than 30% of embryos are estimated to be lost during implantation, with approximately 55% of blastocysts failing to hatch [36]. In fact, in cases of implantation failure during in vitro fertilization procedures, the proportion of unhatched embryos ranged from 50 to 70% [37].

However, previous studies have not placed significant emphasis on hatching observation in ML models. Key predictors in algorithms for predicting blastocyst

Table 3 Implantation predictive models developed, and its statistics, obtained through machine learning

ML Models	AUC	Accuracy	Positive predictive value	Negative predictive value	95% IC
Random Forest	0.725	0.68	0.661	0.705	0.6232–0826
k-Nearest Neighbors	0.720	0.72	0.696	0.750	0.6319–0.8081
Support Vector Machines with Radial Basis Function Kernel	0.717	0.66	0.660	0.660	0.6139–08197
C5.0	0.705	0.66	0.633	0.700	0.6024–0.808
AdaBoost Classification Trees	0.683	0.64	0.635	0.646	0.5752–0.7904
Multi-Layer Perceptron	0.658	0.67	0.649	0.698	0.5517–0.7651
Bagged CART	0.626	0.58	0.577	0.583	0.5169–0.7347
eXtreme Gradient Boosting	0.620	0.59	0.585	0.596	0.5084–0.7324
CART	0.605	0.57	0.564	0.578	0.493–0.7166
Stochastic Gradient Boosting	0.529	0.51	0.509	0.511	0.4135–06441
Generalized Linear Model	0.492	0.51	0.510	0.510	0.3765–0.6067

Table 4 Live birth predictive models developed, and its statistics, obtained through machine learning

ML Models	AUC	Accuracy	Positive predictive value	Negative predictive value	95% IC
AdaBoost Classification Trees	0.749	0.667	0.654	0.682	0.6522–0.8452
eXtreme Gradient Boosting	0.691	0.646	0.630	0.667	0.585–0.7978
Random Forest	0.689	0.646	0.635	0.659	0.5821–0.7921
Bagged CART	0.681	0.615	0.612	0.617	0.5756–0.7869
Stochastic Gradient Boosting	0.670	0.635	0.644	0.627	0.5596–0.7807
C5.0	0.637	0.583	0.574	0.595	0.5239–0.7495
CART	0.562	0.562	0.562	0.562	0.4558–0.6692
Support Vector Machines with Radial Basis Function Kernel	0.556	0.562	0.560	0.565	0.4386–0.6725
Multi-Layer Perceptron	0.540	0.531	0.531	0.532	0.4238–0.6569
Generalized Linear Model	0.506	0.542	0.540	0.543	0.3881–0.624
k-Nearest Neighbors	0.479	0.521	0.519	0.524	0.379–0.5794

implantation include time to 5 cells, length of the second cell cycle, and second-to-third division synchrony [2, 38, 39]. Additionally, time to morula and blastocyst formation have shown high predictive value [2, 40]. A review has identified the most critical parameters as the length of the second cell cycle, time to five cells, and second synchrony [5].

We considered including blastocyst diameter and inner cell mass (ICM) size before transfer (at 110 h post-injection) as potential variables, given their presumed

(See figure on next page.)

Fig. 4 Receiver operating characteristic curve for prediction of implantation and its AUC of the most predictive models: Random Forest (**A**), K-nearest neighbors (**B**), and Support Vector Machines with Radial Basis Function Kernel (**C**), and list of most important parameters in the most predictive variables of the algorithms developed: Variables included in the models: Time of appearance of the two pronuclei (APPN), time of pronuclei fading (DESAPPN), time to two cells (T2), time to three cells (T3), time to four cells (T4), time to five cells (T5), time to six cells (T6), time to seven cells (T7), division time to eight cells (T8), time to early compactation (TiCOM), time to compacted morula formation (TM), time to blastulation starting (TiCAV), time to full blastocyst (TFB), time to expanded blastocyst (TEB), time to hatching out the zona pellucida (TiH), duration of visible pronuclei (DESAPPN-APPN), duration of the second cell cycle (T4-T2), duration of the third cell cycle (T8-T4), synchronization of cells division (T4-T3), synchronization of cleavage patterns (T8-T5), duration of compaction until first sign of cavitation (TiCAV-TM), duration of compaction (TM-TiCOM), duration of blastulation until hatching (TiH-TiCAV) and duration of blastocyst expansion until hatching (TiH-TFB), diameter of the blastocyst (diametroblasto) and the inner cell mass (diametromasa) before the transfer (at 110 h post-injection). Random Forest (A) (AUC = 0.725, IC 95% [0.6232-0826]), K-nearest neighbors (B) (AUC = 0.720, IC 95% [0.6319-0.8081]), and Support Vector Machines with Radial Basis Function Kernel (C) (AUC = 0.717, IC 95% [0.6139-08197]).

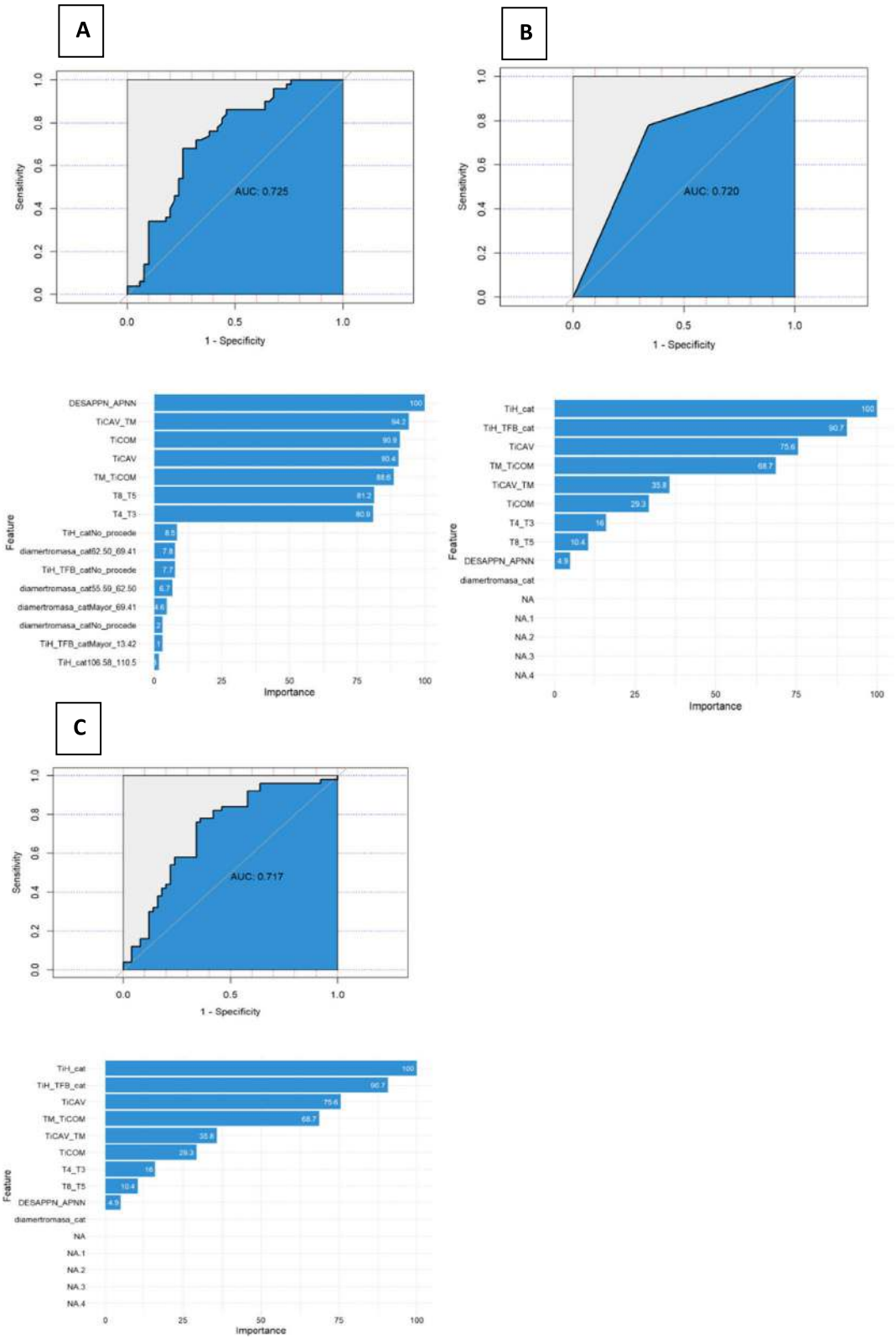

Fig. 4 (See legend on previous page.)

relevance. Euploid blastocysts, associated with higher implantation rates, tend to expand earlier than aneuploid ones [41]. Both diameters were measured twice, and the KILBD analysis utilized their mean values. However, these parameters did not exhibit strong predictive capacity in the implantation and live birth models developed. The blastocyst diameter of implanting and live birth blastocysts averaged 158.03 µm and 157.88 µm, respectively, while non-implanting and non-live birth blastocysts averaged 154.67 µm and 155.62 µm, respectively (Supplementary tables 3 and 5). Furthermore, ICM size did not emerge as an influential variable in any model. In contrast, Almagor et al. [42] reported a correlation between blastocyst and ICM diameter and clinical outcomes.

Previous studies, such as that by Bori et al. [13], emphasized the importance of maximizing relevant information to enhance predictive capability. Bori and colleagues proposed novel markers, including pronuclear kinetics, ICM, and blastocyst measurements, for inclusion in AI models to predict implantation in egg donation cycles.

In our study, after data preprocessing to identify key variables, blastocyst diameter was not selected. Surprisingly, subsequent ML models remained highly predictive. This contrasts with the notion that "all characteristics together were more predictive than individually" [13]. However, we agree with Bori et al. on the potential for improved predictive value through the inclusion of patient-related parameters. Recently, H. Liu et al. demonstrated significant advancements in predictive modeling for live birth based on blastocyst evaluation and clinical features, achieving an AUC of 0.77 [20]. In our study, utilizing exclusively kinetic and morphological embryo parameters, our AdaBoost ML model achieved an AUC of 0.749 for live birth prediction, indicating promising performance. This raises the question of integrating additional clinical parameters to potentially surpass an AUC of 0.8, a benchmark that has yet to be reached by any model.

Another compelling aspect to explore would be evaluating the predictive performance of these ML models for implantation and live birth outcomes, and juxtaposing them with the decision-making proficiency of expert embryologists in embryo selection. In this context, Fordham and colleagues found that the AUC for the deep neural network (DNN) in predicting embryo implantation was higher compared to that achieved by embryologists overall (0.70 for DNN vs 0.61 for embryologists) [43].

Some authors suggest that integrating demographic parameters, as proposed by Petersen et al. [39], may pose data acquisition challenges. Conversely, Cai et al. [44] argue that demographic characteristics are inherently implicit within embryo morphokinetics. Furthermore, d´Estaing et al. [45] caution that excessive parameter inclusion during algorithm construction could undermine predictive accuracy.

Kovacic et al. [46] emphasized that various confounding factors, including laboratory conditions and manual parameter annotation, may undermine the reliability of embryo selection algorithms. In fact, one of the limitations of our study is the susceptibility to biases due to the manual annotation of morphokinetic parameters. Although it was conducted by a single trained embryologist at the study coordination center, ideally, annotations should have been performed by two operators, with a third in case of disagreement. While some algorithms advocate for manual annotation, its subjectivity could be mitigated through automation. However, current automated methods face challenges in recognizing direct divisions, cell fusions, or abnormal cell nucleus division [22]. Nonetheless, recent publications have demonstrated the efficacy and reliability of automated annotation [47–49].

Another limitation of our study is its retrospective nature. Ideally, algorithms require external validation to assess its robustness and accuracy in different conditions from those where it was developed [31]. For that purpose, setting interfaces that integrate the algorithms, facilitating its use to the embryologists would be highly recommendable [30]. Unfortunately, despite our ongoing work in this area, we currently lack external validation

(See figure on next page.)

Fig. 5 Receiver operating characteristic curve for live birth prediction and its AUC of the most predictive models: AdaBoost Classification Trees (**A**), eXtreme Gradient Boosting (**B**) and Random forest (**C**), and list of most important parameters in the most predictive. Variables included in the models: Time of appearance of the two pronuclei (APPN), time of pronuclei fading (DESAPPN), time to two cells (T2), time to three cells (T3), time to four cells (T4), time to five cells (T5), time to six cells (T6), time to seven cells (T7), division time to eight cells (T8), time to early compaction (TiCOM), time to compacted morula formation (TM), time to blastulation starting (TiCAV), time to full blastocyst (TFB), time to expanded blastocyst (TEB), time to hatching out the zona pellucida (TiH), duration of visible pronuclei (DESAPPN-APPN), duration of the second cell cycle (T4-T2), duration of the third cell cycle (T8-T4), synchronization of cells division (T4-T3), synchronization of cleavage patterns (T8-T5), duration of compaction until first sign of cavitation (TiCAV-TM), duration of compaction (TM-TiCOM), duration of blastulation until hatching (TiH-TiCAV) and duration of blastocyst expansion until hatching (TiH-TFB), diameter of the blastocyst (diametroblasto) and the inner cell mass (diametromasa) before the transfer (at 110 h post-injection). AdaBoost Classification Trees (A) (AUC = 0.749, IC 95% [0.6522-0.8452]), eXtreme Gradient Boosting (B) (AUC = 0.691, IC 95% [0.585-0.7978]) and Random forest (C) (AUC = 0.689, IC 95% [0.5821-0.7921])

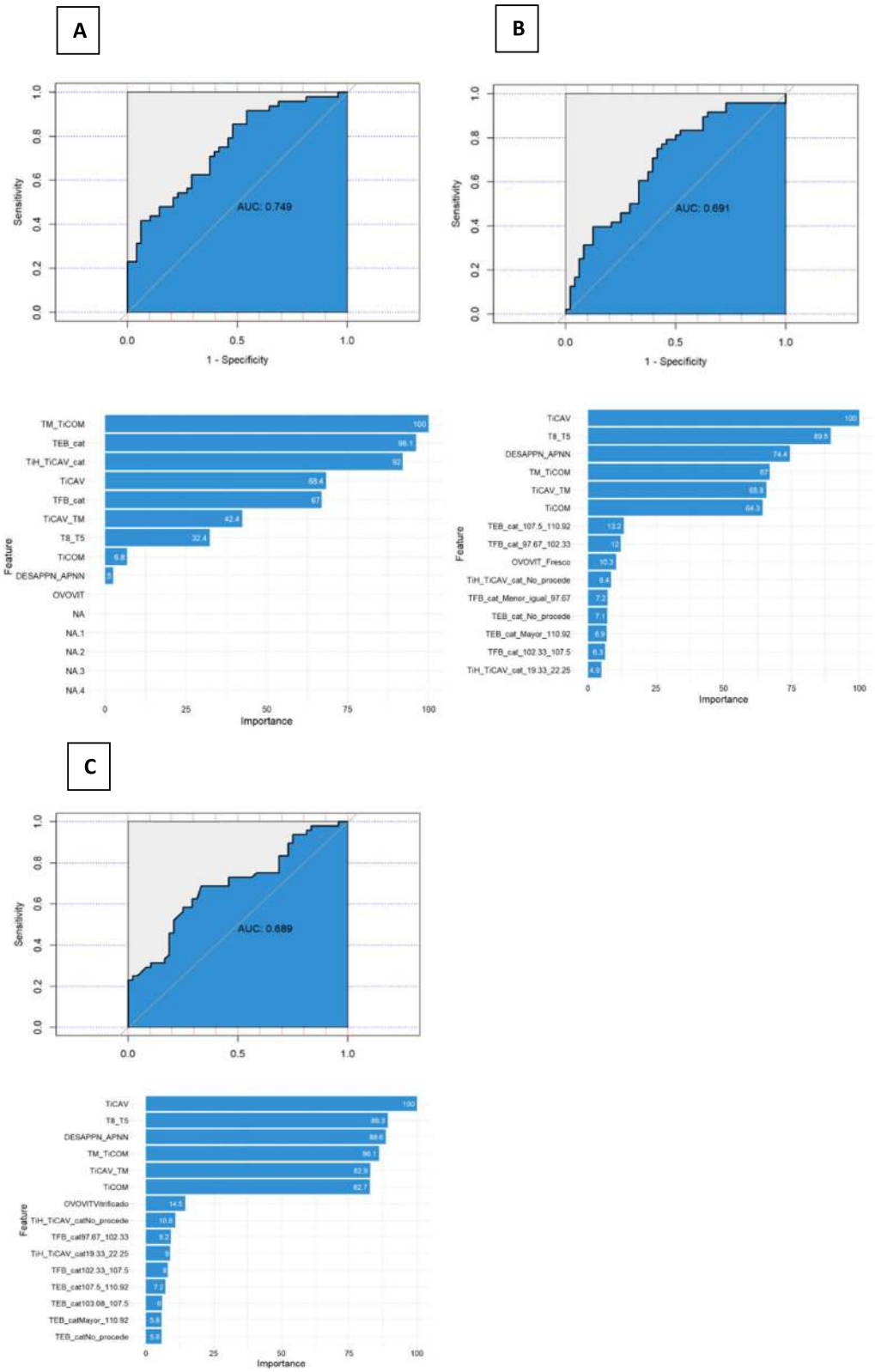

Fig. 5 (See legend on previous page.)

data that would be crucial for generalizing the findings. The execution of a prospective study is also pending. However, it is important to consider the significance of the hatching related variables, which we believe should be incorporated into other algorithms to enhance their predictions.

On the contrary, a key advantage in our study lies in the homogeneity of the egg recipient population and the multicentric approach, which was mitigated by uniform laboratory protocols and procedures, resulting in consistent clinical outcomes across centers. Moreover, the exclusion of low-quality seminal samples was also considered to avoid potential biases they could exert on embryo quality/kinetics.

In our study, we have developed ML algorithms based on morphological and morphokinetic features, but we have made a previous variable selection, with the purpose of getting better predictive results [45]. Four of the eleven implantation ML models had an AUC > 0.70 (Table 3). In general, the kinetic variables related to the blastocyst expansion and hatching processes were the most important variables associated with implantation and also with live birth for AdaBoost algorithm. This led us to regard the observation of the hatching process prior to transfer as a crucial factor in predicting implantation potential, as previously discussed. However, ML models for live birth had a different behavior with less predictive power as shown in Table 4. This could be due to the contribution of unknown maternal clinical features to live birth prediction and endometrium status-related features, such as endometrium preparation, thickness and pattern, that are also critical factors impacting live birth outcomes [20]. It is important to mention that other important variables with high predictive weight for implantation and live birth were those related to syngamy (duration of visible pronuclei, DESAPPN-APPN), embryo cleavage during genomic imprinting (synchronization of cleavage patterns, T8-T5), as well as embryo compaction (duration of compaction, TM-TiCOM; duration of compaction until first sign of cavitation, TiCAV-TM; and time to early compaction, TiCOM) (Figs. 4 and 5). These processes may significantly contribute to achieving successful implantation and live birth. Furthermore, RF has consistently demonstrated superior performance in both implantation and live birth rates. This ML algorithm has been used by other authors demonstrating its high performance [50] even to predict the first trimester miscarriage [51].

However, available scientific evidence that supports the routinely use of these techniques for selecting the best embryo that led to a live birth, is still not enough. In fact, when comparing different algorithms currently available, they result in different conclusions, even when they were trained on the same data, underlining the essential importance of a well-designed mathematical and computational approach [6, 22]. Furthermore, AI models used for the blastocyst selection need expert embryologist supervision to validate the results before performing the embryo transfer and even some authors state that every lab must create its own selection algorithm [52].

Conclusions

To summarize, this is an important embryo sample study, with highly predictive ML algorithms developed (AUC 0.725 for implantation and 0.749 for live birth) and 10 well selected predictive variables. Time to blastocyst hatching out of the zona pellucida appears to have a significant impact in our implantation ML predictive models.

RF was the best ML model for implantation and Adaboost for live birth. Processes related to syngamy, embryo cleavage during genomic imprinting, as well as embryo compaction also can play a relevant role in achieving implantation and live birth. Finally, oocyte vitrification/warming appears to have no impact on clinical outcomes.

It is undeniable, despite numerous variables influencing IVF outcome (intrinsic and extrinsic to embryo development), that AI approaches may improve the prioritization of the most viable embryo favoring single embryo transfer. Predictive models and automatization of the IVF lab is the near future, and it will allow the embryologist to accomplish new tasks in the daily clinic, delegating routine technique work to the AI machines.

Abbreviations

AI	Artificial Intelligence
ASEBIR	Spanish association for the study of reproductive biology
AUC	Area under the ROC curve
DL	Deep learning
Gnrh	Gonadotropin-releasing hormone
IBA	Instituto Bernabéu Alicante
IBAB	Instituto Bernabéu Albacete
IBM	Instituto Bernabéu Madrid
IBPM	Instituto Bernabéu Palma De Mallorca
ICM	Inner Cell Mass
ICSI	Intra-Cytoplasmic sperm injection
IVF	In Vitro fertilization
KILBD	Known implantation and live birth database
LB	Live birth
ML	Machine Learning
RF	Random Forest
GLM	Generalized linear model
MLP	Multi-layer perceptron
SVM	Support vector machines
KNN	K-Nearest neighbor
CART	Cart
ADABOOST	AdaBoost classification trees
GRADBOOOST	Stochastic Gradient boost
BAGGINS	Bagged CART
XGB	EXtrem gradient boosting
ROC	Receiver operating characteristics

SET	Single Embryo Transfer
DESAPPN-APPN	Duration of visible pronuclei
T4-T2	Duration of the second cell cycle
T8-T4	Duration of the third cell cycle
T4-T3	Synchronization of cells division
T8-T5	Synchronization of cleavage patterns
TiCAV-TM	Duration of compaction until first sign of cavitation
TM-TiCOM	Duration of compaction
TiH-TiCAV	Duration of blastulation until hatching
TiH-TFB	Duration of blastocyst expansion until hatching
Diametroblasto	Diameter of the blastocyst

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12958-024-01285-9.

Supplementary Material 1.

Acknowledgements
We would like to express our gratitude to the entire team of the Embryology Unit at Instituto Bernabeu for their support and understanding, which enabled us to dedicate sufficient time to carry out this study.

Authors' contributions
J.T. wrote the main manuscript text. L.H. and E.A. collected the clinical information of the patients and the characteristics of the treatments. A.L. performed manual annotations of embryonic kinetics and assisted in the final formatting of the article. J.A.O. performed data pre-processing, hyperparameter optimization of classification models, final predictive modeling and statistical analysis. L.H., E.A., A.L., A.B. and R.B. reviewed the final version of the paper..

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

Availability of data and materials
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
This study was approved by the Institutional Review Board of Instituto Bernabeu in January 2021 (reference BR20).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Instituto Bernabéu Alicante, Avda. Albufereta, 31, 03016 Alicante, Spain. [2]Instituto Bernabéu Madrid, Madrid, Spain. [3]Instituto Bernabéu Palma, Palma, Spain. [4]Molecular Biology and Genetics, Instituto Bernabéu Biotech, Alicante, Spain. [5]Reproductive Medicine, Instituto Bernabéu, Alicante, Spain.

Received: 8 May 2024 Accepted: 23 August 2024
Published online: 11 September 2024

References

1. Gardner DK, Schoolcraft WB. Culture and transfer of human blastocysts. Curr Opin Obstet Gynecol. 1999;11:307–11.
2. Basile N, Vime P, Florensa M, Aparicio Ruiz B, García Velasco JA, Remohí J, et al. The use of morphokinetics as a predictor of implantation: a multicentric study to define and validate an algorithm for embryo selection. Hum Reprod. 2015;30:276–83.
3. Milewski R, Kuczyńska A, Stankiewicz B, Kuczyński W. How much information about embryo implantation potential is included in morphokinetic data? A prediction model based on artificial neural networks and principal component analysis. Adv Med Sci. 2017;62(1):202–6.
4. Carrasco B, Arroyo G, Gil Y, Gómez MJ, Rodríguez I, Barri PN, et al. Selecting embryos with the highest implantation potential using data mining and decision tree based on classical embryo morphology and morphokinetics. J Assist Reprod Genet. 2017;34:983–90.
5. Gallego RD, Remohí J, Meseguer M. Time-lapse imaging: the state of the art. Biol Reprod. 2019;101(6):1146–54.
6. Khosravi P, Kazemi E, Zhan Q, Malmsten JE, Toschi M, Zisimopoulos P, et al. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. NPJ Digit Med. 2019;4(2):21.
7. Bormann CL, Kanakasabapathy MK, Thirumalaraju P, Gupta R, Pooniwala R, Kandula H, et al. Performance of a deep learning based neural network in the selection of human blastocysts for implantation. Elife. 2020;15(9):e55301.
8. Fitz VW, Kanakasabapathy MK, Thirumalaraju P, Kandula H, Ramirez LB, Boehnlein L, et al. Should there be an "AI" in TEAM? Embryologists selection of high implantation potential embryos improves with the aid of an artificial intelligence algorithm. J Assist Reprod Genet. 2021;38(10):2663–70.
9. Bori L, Dominguez F, Fernandez EI, Del Gallego R, Alegre L, Hickman C, et al. An artificial intelligence model based on the proteomic profile of euploid embryos and blastocyst morphology: a preliminary study. Reprod Biomed Online. 2021;42:340–50.
10. Chavez-Badiola A, Flores-Saiffe-Farías A, Mendizabal-Ruiz G, Drakeley AJ, Cohen J. Embryo Ranking Intelligent Classification Algorithm (ERICA): artificial intelligence clinical assistant predicting embryo ploidy and implantation. Reprod Biomed Online. 2020;41:585–93.
11. Meyer A, Dickinson J, Kelly N, Kandula H, Kanakasabapathy M, Thirumalaraju P, et al. Can deep convolutional neural network (CNN) be used as a non-invasive method to replace preimplantation genetic testing for aneuploidy (PGT-A)? Hum Reprod. 2020;35:I238–I238.
12. Pennetta F, Lagalla C, Borini A. Embryo morphokinetic characteristics and euploidy. Curr Opin Obstet Gynecol. 2018;30:185–96.
13. Bori L, Paya E, Alegre L, Viloria TA, Remohi JA, Naranjo V, et al. Novel and conventional embryo parameters as input data for artificial neural networks: an artificial intelligence model applied for prediction of the implantation potential. Fertil Steril. 2020;114:1232–41.
14. Ortiz JA, Morales R, Lledó B, Vicente JA, González J, García-Hernández EM, et al. Application of machine learning to predict aneuploidy and mosaicism in embryos from in vitro fertilization cycles. AJOG Glob Rep. 2022;2(4):100103.
15. Jiang VS, Bormann CL. Noninvasive genetic screening: current advances in artificial intelligence for embryo ploidy prediction. Fertil Steril. 2023;120:228–34.
16. Miyagi Y, Habara T, Hirata R, Hayashi N. Feasibility of predicting live birth by combining conventional embryo evaluation with artificial intelligence applied to a blastocyst image in patients classified by age. Reprod Med Biol. 2019;18(4):344–56.
17. Miyagi Y, Habara T, Hirata R, Hayashi N. Feasibility of artificial intelligence for predicting live birth without aneuploidy from a blastocyst image. Reprod Med Biol. 2019;18(2):204–11.
18. Huang B, Zheng S, Ma B, Yang Y, Zhang S, Jin L. Using deep learning to predict the outcome of live birth from more than 10,000 embryo data. BMC Pregnancy Childbirth. 2022;22(1):36.
19. Yang L, Peavey M, Kaskar K, Chappell N, Zhu L, Devlin D, et al. Development of a dynamic machine learning algorithm to predict clinical pregnancy and live birth rate with embryo morphokinetics. F S Rep. 2022;3(2):116–23.
20. Liu H, Zhang Z, Gu Y, Dai C, Shan G, Song H, et al. Development and evaluation of a live birth prediction model for evaluating human blastocysts from a retrospective study. eLife. 2023;12:e83662.
21. Jiang VS, Bormann CL. Artificial intelligence in the in vitro fertilization laboratory: a review of advancements over the last decade. Fertil Steril. 2023;120:17–23.

22. Dimitriadis I, Zaninovic N, Badiola AC, Bormann CL. Artificial intelligence in the embryology laboratory: a review. Reprod Biomed Online. 2022;44(3):435–48.
23. Bori L, Meseguer M. Will the introduction of automated ART laboratory systems render most embryologists redundant? Reprod Biomed Online. 2021;43(6):979–81.
24. Zaninovic N, Rosenwaks Z. Artificial intelligence in human in vitro fertilization and embryology. Fertil Steril. 2020;114:914–20.
25. Guerrero J, Castillo JC, Ten J, Ortiz JA, Lledó B, Orozco D, et al. Random-start ovarian stimulation in an oocyte donation programme: a large, single-centre, experience. Reprod Biomed Online. 2024;48(1):103572.
26. Llácer J, Díaz N, Serrano-Revuelta E, Álvarez E, Castillo JC, Ortíz JA, et al. Vitrification does not affect birth weight: lessons from the oocyte donation model. Reprod Biomed Online. 2022;45(2):355–63.
27. Kuwayama M, Vajta G, Kato O, Leibo SP. Highly efficient vitrification method for cryopreservation of human oocytes. Reprod BioMed Online. 2005;11:300–8.
28. Cobo A, Meseguer M, Remohí J, Pellicer A. Use of cryo-banked oocytes in an ovum donation programme: a prospective, randomized, controlled, clinical trial. Hum Reprod. 2010;25:2239–46.
29. Cuevas-Saiz I, Carme Pons Gatell M, Vargas MC, Delgado Mendive A, Rives Enedáguila N, Moragas Solanes M, et al. The Embryology Interest Group: updating ASEBIR's morphological scoring system for early embryos., morulae and blastocysts. Med Reprod y Embriol Clinica. 2018;5(1):42–54.
30. Ratna MB, Bhattacharya S, Abdulrahim B, McLernon DJ. A systematic review of the quality of clinical prediction models in in vitro fertilisation. Hum Reprod. 2020;35(1):100–16.
31. Fernandez EI, Ferreira AS, Cecílio MHM, Chéles DS, de Souza RCM, Nogueira MFG, et al. Artificial intelligence in the IVF laboratory: overview through the application of different types of algorithms for the classification of reproductive data. J Assist Reprod Genet. 2020;37(10):2359–76.
32. Abdullah KAL, Atazhanova T, Chavez-Badiola A, Shivhare SB. Automation in ART: Paving the way for the future of infertility treatment. Reprod Sci. 2022;30(4):1006–16.
33. Ten J, Peinado P, Guerrero J, Bernabeu A, Llácer J, Orozco-Beltran D, et al. Comparison of the assisted reproductive technology outcomes between conventional IVF and ICSI with donor oocytes in normozoospermic patients. Hum Fertil (Camb). 2022;25(1):56–62.
34. Montgomery K, Montgomery S, Campbell A, Nash DM. A comparison of the morphokinetic profiles of embryos developed from vitrified versus fresh oocytes. Reprod BioMed Online. 2023;47(1):51–60.
35. Murria L, Meseguer M, Coello A, Campos P, Bori L, Cobo A. Artificial intelligence algorithms in assisted reproduction: differences in the evaluation of embryos from fresh or vitrified donor oocytes. Hum Reprod. 2023;38:i147–8.
36. Meixiang M, Liang Z, Zihan L, Yadi T, Miaolong L, Xinrong P, et al. Effect of blastocyst development on hatching and embryo implantation. Theriogenology. 2024;214:66–72.
37. Lu X, Liu Y, Cao X, Liu SY, Dong X. Laser-assisted hatching and clinical outcomes in frozen-thawed cleavage-embryo transfers of patients with previous repeated failure. Laser Med Sci. 2019;34:1137–45.
38. Liu Y, Chapple V, Feenan K, Roberts P, Matson P. Time-lapse deselection model for human day 3 in vitro fertilization embryos: the combination of qualitative and quantitative measures of embryo growth. Fertil Steril. 2016;105:656–62.
39. Petersen BM, Boel M, Montag M, Gardner DK. Development of a generally applicable morphokinetic algorithm capable of predicting the implantation potential of embryos transferred on day 3. Hum Reprod. 2016;31:2231–44.
40. Chamayou S, Patrizio P, Storaci G, Tomaselli V, Alecci C, Ragolia C, et al. The use of morphokinetic parameters to select all embryos with full capacity to implant. J Assist Reprod Genet. 2013;30:703–10.
41. Huang TT, Huang DH, Ahn HJ, Arnett C, Huang CT. Early blastocyst expansion in euploid and aneuploid human embryos: evidence for a non-invasive and quantitative marker for embryo selection. Reprod Biomed Online. 2019;39:27–39.
42. Almagor M, Harir Y, Fieldust S, Or Y, Shoham Z. Ratio between inner cell mass diameter and blastocyst diameter is correlated with successful pregnancy outcomes of single blastocyst transfers. Fertil Steril. 2016;106:1386–91.
43. Fordham DE, Rosentraub D, Polsky AL, Aviram T, Wolf Y, Perl O, et al. Embryologist agreement when assessing blastocyst implantation probability: is data-driven prediction the solution to embryo assessment subjectivity? Hum Reprod. 2022;37:2275–90.
44. Cai J, Luo J, Wang S, Yang S. Feature selection in machine learning: A new perspective. Neurocomputing. 2018;300:70–9.
45. Giscard d'Estaing S, Labrune E, Forcellini M, Edel C, Salle B, Lornage J, et al. A machine learning system with reinforcement capacity for predicting the fate of an ART embryo. Syst Biol Reprod Med. 2021;67(1):64–78.
46. Kovacic B, Le Clef N, Marques M, Meseguer M, Montjean D, Sfontouris I, et al. Good practice recommendations for the use of time-lapse technology. ESHRE Working group on Time-lapse technology. Hum Reprod Open. 2020;2020(2):hoaa008.
47. Malmsten J, Zaninovic N, Zhan Q, Rosenwaks Z, Shan J. Automated cell division classification in early mouse and human embryos using convolutional neural networks. Neural Comput Applic. 2020;33:2217–28.
48. Leahy BD, Jang WD, Yang HY, Struyven R, Wei D, Sun Z, et al. Automated measurements of key morphological features of human embryos for IVF. Med Image Comput Assist Interv. 2020;12265:25–35.
49. Bormann CL, Thirumalaraju P, Kanakasabapathy MK, Kandula H, Souter I, Dimitriadis I, et al. Consistency and objectivity of automated embryo assessments using deep neural networks. Fertil Steril. 2020;113(4):781-787.e1.
50. Liu X, Chen Z, Ji Y. Construction of the machine learning-based live birth prediction models for the first in vitro fertilization pregnant women. BMC Pregnancy Childbirth. 2023;23(1):476.
51. Amitai T, Kan-Tor Y, Or Y, Shoham Z, Shofaro Y, Richter D, et al. Embryo classification beyond pregnancy: early prediction of first trimester miscarriage using machine learning. J Assist Reprod Genet. 2023;40(2):309–22.
52. Barrie A, Homburg R, McDowell G, Brown J, Kingsland C, Troup S. Examining the efficacy of six published time-lapse imaging embryo selection algorithms to predict implantation to demonstrate the need for the development of specific, in-house morphokinetic selection algorithms. Fertil Steril. 2017;107:613–21.

Publisher's Note