Journal of Ovarian Research

RESEARCH

Open Access

Preovulatory progesterone levels are the top indicator for ovulation prediction based on machine learning model evaluation: a retrospective study

Yumei Li[1*†], Hong Zeng[1†] and Jing Fu[1]

Abstract

Background Accurately predicting ovulation timing is critical for women undergoing natural cycle-frozen embryo transfer. However, the precise predicting of the ovulation timing remains challenging due to the lack of consensus among different clinics regarding the definition of this significant event.

Objective To compare the effectiveness of preovulatory serum progesterone levels (P4) versus luteinizing hormone levels (LH) in predicting ovulation time using two machine learning models.

Methods 771 patients who underwent autologous natural cycle-frozen embryo transfer between January 2015 and February 2022 were recruited. Utilizing variables including follicle diameters, preovulatory serum levels of LH, E2, and P4, two machine learning models were constructed to predict the ovulation time, the importance of the variables in predicting ovulation timing was further ranked.

Results Two machine learning models have the capability to accurately predict the timing of ovulation, specifically within 72, 48, or 24 h. The overall accuracy rates of the validation dataset, as determined by the classification trees and random forest models, were found to be 78.83% and 85.28% respectively. Notably, when predicting ovulation within 24 h, the accuracy rate of P4 ≥ 0.65ng/ml exceeded 92%. Furthermore, it was important to consider LH or E2 levels in conjunction with P4 when assessing ovulation timing in cases where P4<0.65ng/ml.

Conclusions Preovulatory serum P4 levels are better predictors of ovulation timing than LH levels and could be used as an alternative in clinical settings, and the model we developed can be used to pinpoint the day of ovulation. Ongoing research and advancements in technology are anticipated to enhance and refine the ovulation method.

Keywords Progesterone hormone, Luteinizing hormone, Natural cycle, Ovulation time prediction

†Yumei Li and Hong Zeng contributed equally to this work and are co-first authors.

*Correspondence:
Yumei Li
liyumei7511@163.com
[1]Reproductive Medicine Center, Xiangya Hospital Central South University, Changsha 410008, China

Introduction

Frozen-thawed embryo transfer (FET) has become widespread in recent years. The success of FET relies on several factors, including a receptive endometrium, an implantation-competent embryo, and synchronization between the embryo and the endometrium [1]. Recent evidence indicates that the hormone replacement FET protocol, when compared to the natural cycle (NC), may be associated with a higher incidence of bleeding and miscarriage during early pregnancy [2]. Additionally, endometrial priming with NC has yielded more favorable obstetric and perinatal outcomes [3–7]. Consequently, the natural cycle FET (NC- FET) protocol is recommended for women with regular ovulation.

Several studies have suggested that the endometrium is most receptive to embryo implantation during the window of implantation (WOI) [8–10]. Teh et al. observed that the embryo implantation rate significantly decreased when there was a discrepancy of more than ±1.5 days between the endometrium and the embryo [11]. Therefore, in order to ensure successful embryo transfer within the WOI, it is essential to accurately determine the ovulation day in NC-FET. The disappearance of the leading follicle, identified by ultrasound, serves as a reliable indicator of ovulation in many studies. A study on 271 ovulation cycles found a sensitivity of 84.3% and specificity of 89.2% for the ovulation sign [12–15]. The LH surge is another commonly utilized method for predicting the time of ovulation and embryo transfer in clinical practice. However, the LH surge exhibits considerable variability in terms of duration, amplitude, and configuration, which can manifest as a single peak, a double peak, or a plateau [9, 12], ovulation may transpire within a timeframe of 24 to 56 h subsequent to the onset of the spontaneous LH surge [16]. A clear optimal window for embryo transfer could not be ascertained due to the variation in LH surge kinetics and the time from the onset of LH surge to ovulation. Furthermore, during the reproductive cycle, there are regular fluctuations in serum P4 concentrations, which play a significant role in the periodic development and maturation of eggs prior to ovulation [17]. An update review presents compelling evidence suggesting that the authentic physiological stimulus for ovulation is an autonomous surge of preovulatory P4, with levels reaching approximately 0.5ng/ml [18]. It is worth considering whether assessing P4 levels during the preovulatory phase could serve as an alternative method for predicting the ovulation time, instead of relying solely on the LH surge.

Machine learning enables computers to acquire knowledge from data, identify patterns, and establish correlations without the need for manual feature engineering. These sophisticated algorithms exhibit enhanced predictive capabilities in comparison to statistical models [19, 20].

The aim of the study was to compare the effectiveness of preovulatory serum progesterone levels (P4) versus luteinizing hormone levels (LH) in predicting ovulation time using two machine learning models.

Methods

Study design and population

All patients who had undergone autologous NC-FET at the Reproductive Medical Center of Xiagnya Hospital of Central South University from January 2015 to February 2022 were retrospectively recruited. The inclusion criteria consisted of patients with regular menstrual cycle who were planning to undergo NC-FET. Patients who had induced ovulation using human menopausal gonadotropin, Letrozole, or clomiphene, as well as those with missing or extreme values for key variables, and patients with Luteinized Unruptured Follicle Syndrome, were excluded from the study. Data were obtained from the electronic medical record system of the in vitro fertilization (IVF) facility. The collected data encompassed basic information such as age and BMI. Additionally, key variables required for the development of predictive model were also collected, including follicle diameter, E2, P4, and LH levels of multiple time points prior to ovulation. The candidate variables such as age, BMI, follicle diameter, E2 levels, P4 levels, and LH levels were trained to classify the ovulation time (ovulation in 72 h, ovulation in 48 h, and ovulation in 24 h).

The study has been performed in accordance with the Declaration of Helsinki, and received approval from the Institutional Review Board at Central South University of Xiangya Hospital (ethics approval number: 2022012). All patients had given consent for their treatment data to be used for analysis.

NC-FET ovulation detection

In natural cycles, the follicle development was beginning to be measured with ultrasound on cycle day 8–10, with the timing dependent on the patient's menstrual cycle length. Two orthogonal diameters (d1 and d2) at the largest follicle plane were determined by transvaginal ultrasound scan, and the mean follicular diameter was calculated as (d1 + d2)/2. Scanning was subsequently repeated every 2 or 3 days until a dominant follicle diameter reached 14 mm, and then on a daily basis until evidence of ovulation was observed. Ovulation was confirmed through ultrasound scan, whereby a dominant mature follicle identified in one scan was observed to ovulate in the subsequent scan, then the day was defined as the ovulation day.

Serum E2, P4, and LH values were assessed every 24 h each morning when the dominant follicle reached or

exceeded 14 mm. Serum E2, progesterone, and LH levels were measured using electrochemiluminescence immunoassay (ECLIA) methods obtained from a commercial company (Roche Diagnostics GmbH, Germany). The assay for progesterone had a repeatability of ≤0.08 ng/ml and an intermediate precision of ≤0.12 ng/ml SD for samples≤0.5 ng/ml. For samples ranging from 0.5 to 1 ng/ml, the repeatability CV% was ≤10% and the intermediate precision CV% was ≤11.5%. For samples>1ng/ml, the repeatability CV% was ≤6% and the intermediate precision CV% was ≤7%.

Predictive model establishment and validation

In order to develop a predictive model for determining the timing of ovulation, various factors such as age, BMI, E2, P4, LH levels, and follicle diameter were taken into consideration. The data was split into a training dataset and a validation dataset using a random seed number (123456) in R software (version 4.1.3) with the "sample" function [1]. The split was done per cycle when splitting the dataset between the train and validation. And the cycles of the same patient may exist in both datasets. The training dataset comprised 80% of the total sample size, while the validation dataset accounted for 20% of the total sample size. The training dataset was utilized to construct the predictive model, whereas the validation dataset served to assess the accuracy of the aforementioned model. Two methods, namely classification trees and random forest, were employed to develop predictive models. The classification trees model was implemented using the "rpart package" (version 4.1.16) [2], The classification tree was built on the training set by "rpart" function with default parameters. We selected the CP value with the smallest average error based on 10-fold cross validation. The trees were then pruned by the selected CP value using the "prune" function. The random forest model utilized the "randomForest" package (version 4.6–14) [3]. The model was first trained on the training detaset by the "randomForest" function with default parameters. The model was tuned by the "tuneRF" function to select the optimal number of randomly drawn candidate variables (mtry). The arguments in the function were set as follows: the starting number of randomly drawn candidate variables was set to 2 (mtryStart=2); the ending number of randomly drawn candidate variables was set to 6 (mtryEnd=6); the value that mtry is inflated (or deflated) at each iteration was set to 1 (stepFactor=1); the (relative) improvement in out of bag (OOB) error for the search to continue was set to 0.01 (improve=0.01). The number of randomly drawn candidate variables was determined when the OOB error was the smallest. With the fixed value of mtry, we further determined the optimal number of trees. The number of trees was selected when the error rate was the smallest. Finally, the model was re-trained with the optimal numbers of mtry and ntree.

The importance of variables was determined by the "varImpPlot" function in the "randomForest" package. The importance of variables was ranked by Gini index. To assess the accuracy of the predictive models, a confusion matrix was employed. Besides, the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) of each model were reported.

To further compare the effectiveness of each variable in predicting ovulation time, the accuracy rate, sensitivity, specificity, PPV, and NPV of P4 were compared to those of LH, E2, and follicle diameter using the classification trees model.

Statistical analysis

All statistical analyses were conducted using R software version 4.1.3 with the "compareGroups" package [4]. The normally distributed continuous variables were reported as mean±standard deviation (SD), while non-normally distributed continuous variables were presented as median and interquartile ranges (IQR). The t-test or ANOVA analysis was employed for comparing normally distributed data, whereas the Mann-Whitney U test or Wilcoxon test was used for non-normally distributed data, depending on the distribution type. In cases where the crude p-value was less than 0.05, post hoc multiple comparisons were performed, and the p-values of pairwise comparisons were adjusted accordingly. Statistical significance was defined as a p-value less than 0.05.

Results

Patients' information and comparison

A total of 1632 records, out of 771 patients, were included in the study. These records were categorized based on the time before ovulation, with 306 records from 72 h before ovulation, 598 records from 48 h before ovulation, and 728 records from 24 h before ovulation. Various characteristics such as age, BMI, follicle diameter, E2 levels, P4 levels, and LH levels were utilized to establish a predictive model. The records were further divided into three groups, namely the 72 h group, the 48 h group, and the 24 h group, based on the time before ovulation. A comparative analysis was conducted to examine the follicle diameter and hormone levels at each point. The basic demographic characteristics, such as age and BMI, were compared among the three groups. The results presented in Table 1; Fig. 1 indicated that there were no significant differences in age and BMI between the three groups. However, there were significant differences in follicle diameter among the three groups (p-overall<0.001). Specifically, the E2 levels in the 48 h group were significantly higher than those in the 72 h and 24 h groups, no significant difference was found between the 72 h

Table 1 Baseline characteristics, follicle diameter, and hormone levels of records grouped by ovulation timing

	72 h	48 h	24 h	*p*.overall	*p*.72 h vs. 48 h	*p*.72 h vs. 24 h	*p*.48 h vs. 24 h
	N = 306	*N* = 598	*N* = 728				
Age (year)	33.2 ± 4.95	33.2 ± 4.88	33.2 ± 4.86	0.997	0.999	1.000	0.997
BMI (kg/m^2)	21.8 ± 2.81	21.7 ± 2.72	21.7 ± 2.72	0.920	0.915	0.942	0.994
Diameter (mm)	16.3 ± 1.35	17.4 ± 1.54	18.4 ± 1.71	< 0.001	0.000	0.000	0.000
E2 (pg/ml)	282 ± 107 269 [201–336]	378 ± 128 368 [283–449]	286 ± 114 273 [203–351]	< 0.001	< 0.001	0.511	< 0.001
LH (mIU/ml)	14.5 ± 5.55 13.5 [10.6–17.2]	32.2 ± 16.4 28.6 [20.5–40.0]	50.5 ± 21.2 47.3 [35.8–63.5]	< 0.001	< 0.001	< 0.001	< 0.001
P4 (ng/ml)	0.22 ± 0.16 0.20 [0.10–0.30]	0.38 ± 0.22 0.30 [0.20–0.50]	0.87 ± 0.28 0.90 [0.70-1.00]	< 0.001	< 0.001	< 0.001	< 0.001

Notes: 72 h: Ovulation within 72 h group; 48 h: Ovulation within 48 h group; 24 h: Ovulation within 24 h group; N: number of records in each group; p.overall: overall *p* value; p.72 h vs. 48 h: *p* value of comparison between the 72 h group and the 48 h group; p.72 h vs. 24 h: *p* value of comparison between the 72 h group and the 24 h group; p.48 h vs. 24 h: *p* value of comparison between the 48 h group and the 24 h group; *p* values of all post-hoc comparisons were adjusted. E2: estrogen level; LH: luteinizing hormone level; P4: progesterone level. Statistical description of hormone levels including estrogen, luteinizing hormone and progesterone were presented by both mean ± standard deviation and median [interquartile range] because of their non-normally distribution

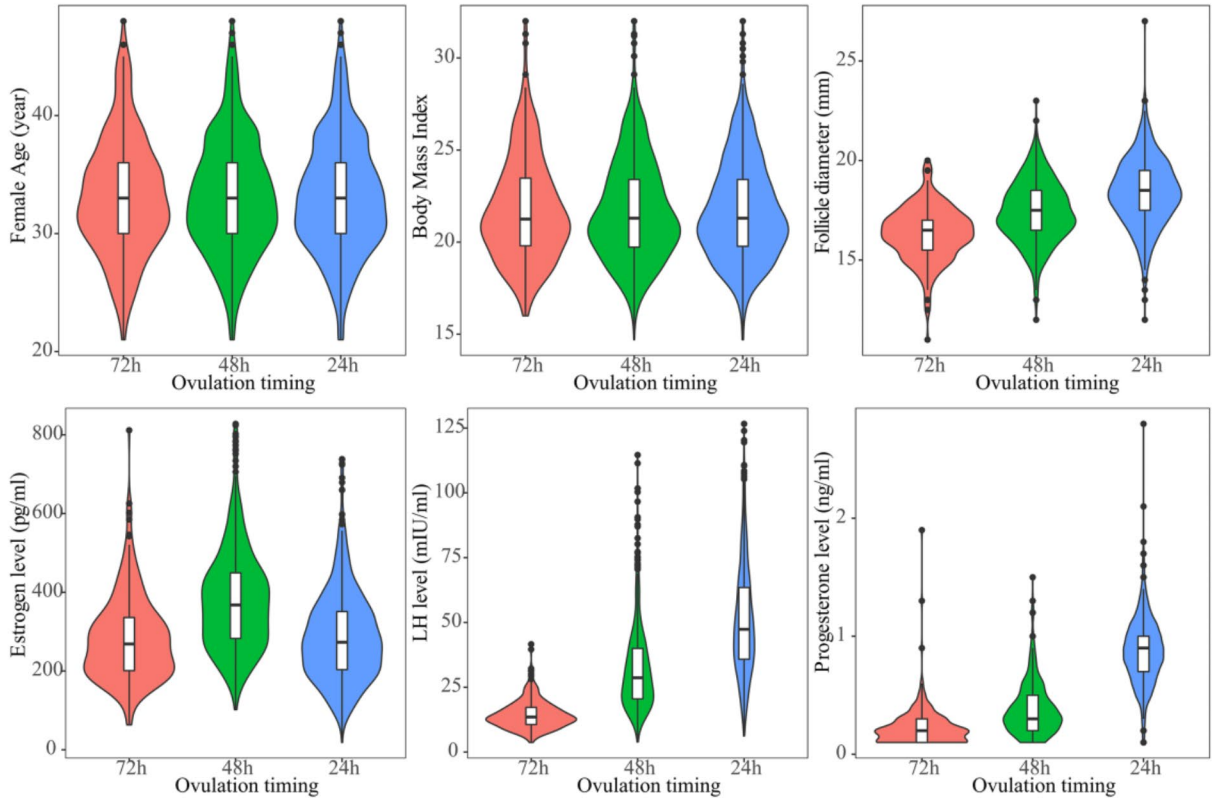

Fig. 1 Box plots and Violin plots showing the baseline characteristics, follicle diameter, and hormone levels of all records at three time points before ovulation. The bottom line, middle line, and upper line of the box represents the first quartile, the medium, and the third quartile of the variable. Outliers marked by dots that are either 1.5*IQR or more above the third quartile or 1.5*IQR or more below the first quartile. Violin plots showed the distribution of each variable. 72 h: in 72 h before ovulation; 48 h: in 48 h before ovulation; 24 h: in 24 h before ovulation

and 24 h groups (p-value = 0.511). The LH levels exhibited significant differences among the three groups (all p-values < 0.001), progressively increasing from 72 h to 24 h prior to ovulation. Similarly, the P4 levels displayed significant differences among the three groups (all p-values < 0.001), showing an increase from 72 h to 24 h before ovulation.

Characteristics of the dynamic changes in follicle diameter and hormone levels prior to and following ovulation

In order to provide a more precise depiction of the dynamic changes in follicle diameter and hormone levels prior to and following ovulation, our analysis exclusively incorporated patients who had measurements for both variables at all four time points: 72 h before ovulation,

48 h before ovulation, 24 h before ovulation, and the day of ovulation. A total of 84 patients met this criterion and were included in the study. According to the data presented in Fig. 2, the follicle diameter exhibits a continuous increasing within 72 to 24-hour timeframe preceding ovulation (Fig. 2A **and E**). Additionally, the P4 levels demonstrate an increase from 72 h prior to ovulation until the day of ovulation (Fig. 2B **and F).** Furthermore, the E2 levels display an initial increase from 72 to 48 h before ovulation, followed by a decrease within the 48 to 24-hour period preceding ovulation, and a continued decrease post-ovulation. Notably, there is no significant difference in E2 levels between the 72 h and 24 h time points before ovulation (Fig. 2C **and G**). The LH levels exhibited increase from 72 to 24 h prior to ovulation, followed by a significant decrease on the day after ovulation. There was no significant difference in LH levels between the 72 h before ovulation and the day of ovulation (Fig. 2D **and H**). To account for potential variations in hormone levels and follicle diameter among individuals, we additionally presented paired line plots depicting the trends of follicle diameter and hormone levels at four specific time points for each individual (Fig. 2I **and L**), each line represents the trend of an individual's follicle diameter or hormone levels at four time points. According to Fig. 2I, the follicle diameter of the majority of patients exhibits increase from 72 h to 24 h prior to ovulation, reaching its peak at 24 h before ovulation. Figure 2J demonstrates that the P4 level of most patients shows increase from 72 h before ovulation until the day of ovulation day, with minimal variability. Figure 2K reveals a significant variability in the timing of the E2 peak, as some patients reach their peak at 48 h while others reach it at 24 h. Figure 2K also demonstrates a notable variability in the timing of the LH peak, with some patients reaches their peak at 48 h while others reach it at 24 h.

Classification trees model

A total of 1306 records were utilized in the training dataset to train a categorical regression model for predicting ovulation timing. The dataset consisted of 252 records for a 72-hour timeframe, 477 records for a 48-hour timeframe, and 577 records for a 24-hour timeframe. Additionally, a validation dataset comprising 326 records was employed, with 54 records for a 72-hour time-frame, 121 records for a 48-hour timeframe, and 151 records for a 24-hour timeframe. The classification trees analysis conducted on the training model (Fig. 3) revealed that a preovulatory P4 level of ≥0.65 ng/ml indicates a high probability of ovulation occurring within 24 h. However,

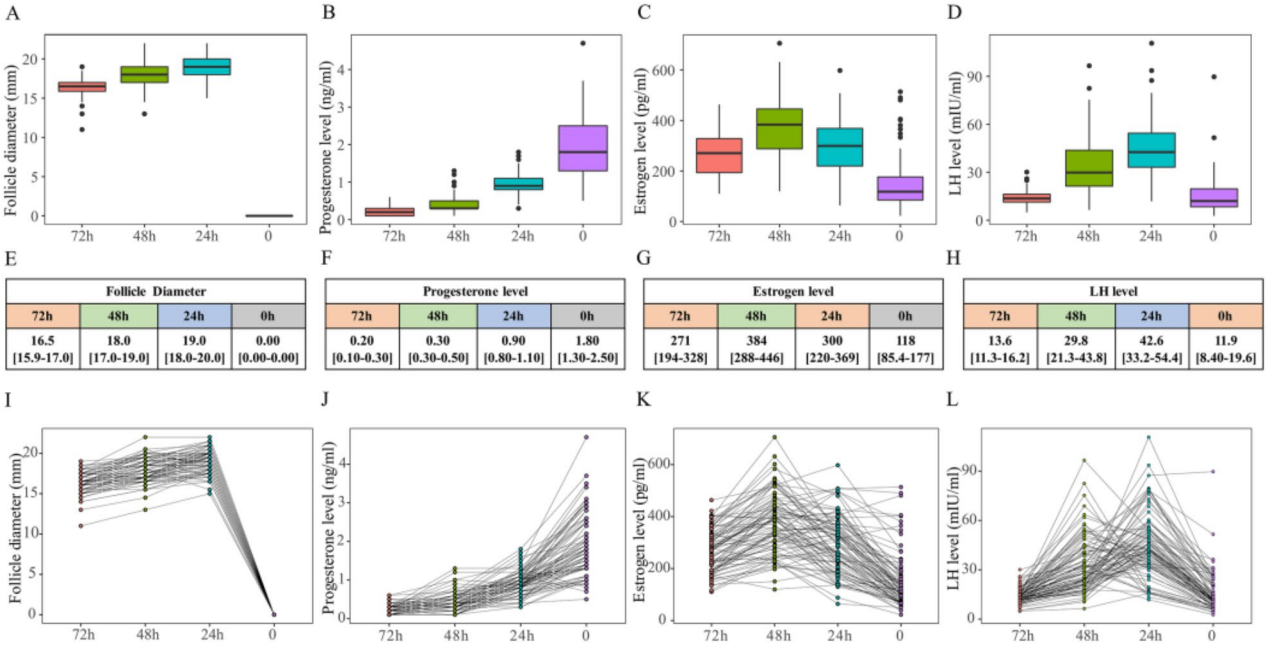

Fig. 2 Dynamic changes of follicle diameter and hormone levels over time before 72 h within ovulation to ovulation day in selected patients with complete data at all four time point. **A**-**D** Box plots describe the follicle diameter and hormone levels at four time points. The bottom line, middle line, and upper line of the box represents the first quartile, the medium, and the third quartile of the variable. Outliers marked by dots that are either 1.5*IQR or more above the third quartile or 1.5*IQR or more below the first quartile. **E-H** Statistical tables describe the diameter and hormone levels at four time points. The different color in the table indicates that the *p* value of the two group comparison was statistical significant, while the same color color in the table indicates that the *p* value of the two group comparison was not statistical significant. The statistical description were presented by median [interquartile range]. **I-L** Paired line plots describe per individual's follicle diameter and hormone levels over four time points.72 h: in 72 h before ovulation; 48 h: in 48 h before ovulation; 24 h: in 24 h before ovulation; 0 h: the day of ovulation

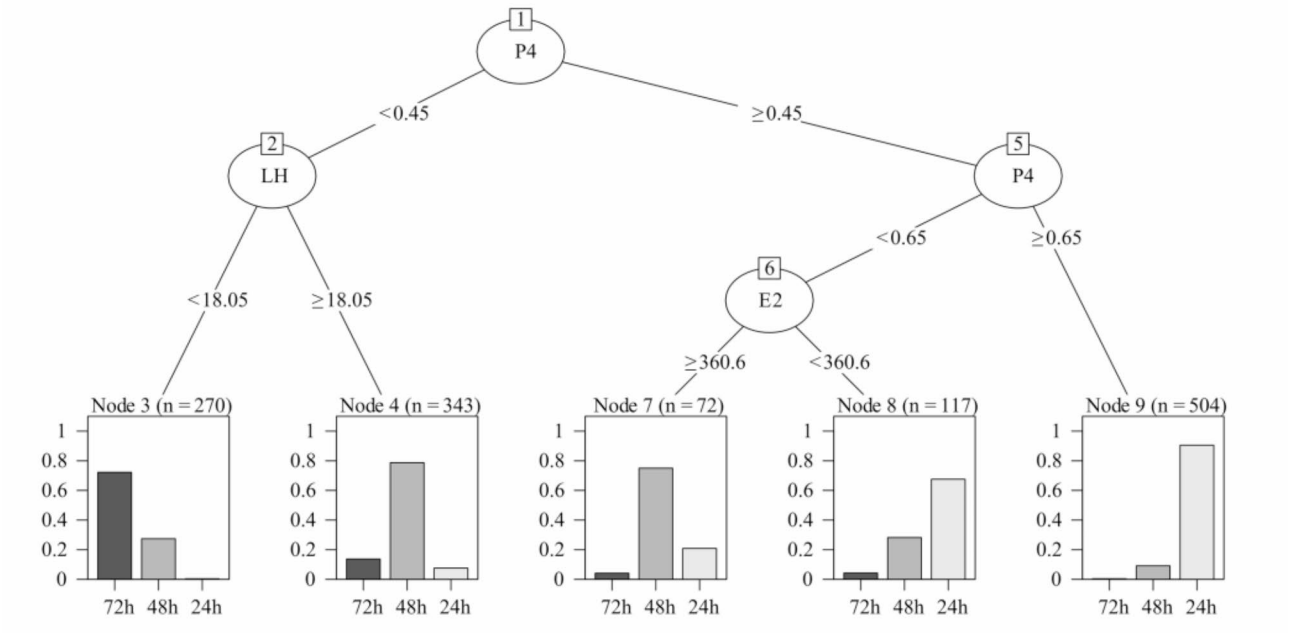

Fig. 3 Regression tree plot of the classification trees model. P4: progesterone (ng/ml), LH: luteinizing hormone (mIU/ml); E2: estrogen (pg/ml)

Table 2 Confusion matrix of classification trees model

Training data (*N* = 1306)				
	72 h (Real)	**48 h (Real)**	**24 h (Real)**	**Total**
72 h (Predictive)	195	74	1	
48 h (Predictive)	50	324	41	
24 h (Predictive)	7	79	535	
Total	252	477	577	1306
Accuracy rate	0.7738	0.6792	0.9272	0.8070
Sensitivity	0.7222	0.7807	0.8615	
Specificity	0.9450	0.8283	0.9387	
PPV	0.7738	0.6792	0.9272	
NPV	0.9288	0.8902	0.8820	
Validation data (*N* = *326)*				
	72 h (Real)	**48 h (Real)**	**24 h (Real)**	**Total**
72 h (Predictive)	40	16	1	
48 h (Predictive)	10	77	10	
24 h (Predictive)	4	28	140	
Total	54	121	151	326
Accuracy rate	0.7407	0.6364	0.9272	0.7883
Sensitivity	0.7018	0.7938	0.8140	
Specificity	0.9480	0.8079	0.9286	
PPV	0.7407	0.6364	0.9272	
NPV	0.9375	0.9024	0.8171	

Notes: N: number of records; 72 h (Real): number of records in 72 h group in real situation; 48 h (Real): number of records in 48 h group in real situation; 24 h (Real): number of records in 24 h group in real situation; 72 h (Predictive): number of records predicted in the 72 h group according to the classification trees model; 48 h (Predictive): number of records predicted in the 48 h group according to the classification trees model; 24 h (Predictive): number of records predicted in the 24 h group according to the classification trees model. PPV: positive predictive value; NPV: negative predictive value

when the preovulatory P4 level falls between 0.45 and 0.65 ng/ml, it is recommended to combined with E2 levels for accurate prediction of ovulation timing. A preovulatory P4 level ranging from 0.45 to 0.65 ng/ml, in conjunction with an estradiol (E2) level of ≥360.6 pg/ml, serves as a strong indicator that ovulation will take place within 48 h. However, in cases where a reduction in E2 levels following an E2 surge is observed, the presence of a preovulatory P4 level between 0.45 and 0.65 ng/ml, in conjunction with an E2 level<360.6 pg/ml, suggests a high probability of ovulation occurring on the subsequent day. When the P4 level<0.45 ng/ml, it can be combined with LH levels to obtain dependable outcomes. If the LH level≥18.05 mIU/ml, there is a high probability of ovulation occurring within 48 h. Conversely, if both LH<18.05 mIU/ml and P4<0.45 ng/ml, it indicates a high probability of ovulation will not taking place within 48 h, an ultrasound scan can be arranged two days later.

The confusion matrix revealed that the classification trees model achieved an overall predictive accuracy of 80.70% on the training dataset. Specifically, the accuracy rates for predicting ovulation within 24 h, 48 h, and 72 h were 92.72%, 67.92%, and 77.38% respectively. Similarly, on the validation dataset, the classification trees model demonstrated an overall predictive accuracy of 78.83%. The accuracy rates for predicting ovulation within 24 h, 48 h, and 72 h were 92.72%, 63.64%, and 74.07% respectively (Table 2). According to the findings of the classification trees model, the preovulatory P4 levels exhibit increase leading up to ovulation and emerge as the most significant parameter in predicting the timing of ovulation.

Table 3 Confusion matrix of random forest model

Training data (*N* = 1306)				
	72 h (Real)	**48 h (Real)**	**24 h (Real)**	**Total**
72 h (Predictive)	252	0	0	
48 h (Predictive)	0	477	0	
24 h (Predictive)	0	0	577	
Total	252	477	577	1306
Accuracy rate	1.0	1.0	1.0	1.0
Sensitivity	1.0	1.0	1.0	1.0
Specificity	1.0	1.0	1.0	1.0
PPV	1.0	1.0	1.0	1.0
NPV	1.0	1.0	1.0	1.0
Validation data (*N* = 326)				
	72 h (Real)	**48 h (Real)**	**24 h (Real)**	**Total**
72 h (Predictive)	42	9	0	
48 h (Predictive)	11	90	5	
24 h (Predictive)	1	22	146	
Total	54	121	151	326
Accuracy rate	0.7778	0.7438	0.9669	0.8528
Sensitivity	0.8235	0.8491	0.8639	
Specificity	0.9564	0.8591	0.9682	
PPV	0.7778	0.7438	0.9669	
NPV	0.9669	0.9220	0.8686	

Note: N: number of records; 72 h (Real): number of records in 72 h group in real situation; 48 h (Real): number of records in 48 h group in real situation; 24 h (Real): number of records in 24 h group in real situation; 72 h (Predictive): number of records predicted in the 72 h group according to the random forest model; 48 h (Predictive): number of records predicted in the 48 h group according to the random forest model; 24 h (Predictive): number of records predicted in the 24 h group according to the random forest model. PPV: positive predictive value; NPV: negative predictive value

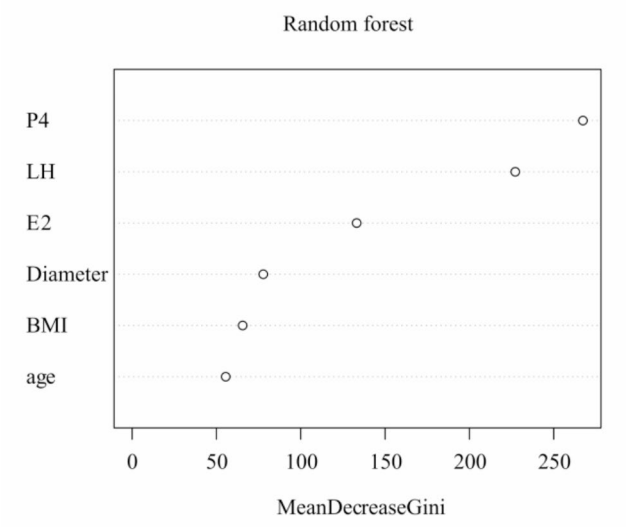

Fig. 4 Dot plot showing the importance of variables in predicting ovulation timing by Gini index in random forest model. P4: progesterone, LH: luteinizing hormone; E2: estrogen; BMI: body mass index

Random forest model

We further employed the random forest method to develop an alternative predictive model. The evaluation of this model using confusion matrix revealed a 100% accuracy rate for the training dataset. Moreover, the validation dataset exhibited an accuracy rate of 85.8%. Additionally, the accuracy rates for predicting ovulation within 24 h, 48 h, and 72 h were found to be 96.69%, 74.38%, 77.78% respectively (Table 3). The random forest model has demonstrated a notable enhancement in the accuracy rate compared to the classification trees model. Figure 4 illustrates the ranking of variable importance using the Gini index, revealing that hormone levels such as P4, LH, and E2 are the top three influential variables in predicting ovulation timing. Conversely, variables such as follicle diameter, BMI, and age do not significant important compared to hormone levels. This finding aligns with the classification trees model, which also identifies P4 as the most crucial variable for predicting ovulation time.

Comparison of effectiveness of each variable in predicting ovulation time

To compare the effectiveness of each individual variable in predicting ovulation time, the overall accuracy rate, sensitivity, specificity, PPV, and NPV of P4 were compared to those of LH, E2, and follicle diameter using the classification trees model.

When training predictive model using single P4 levels, the cutoff values are 0.25ng/ml and 0.45ng/ml (Fig. 5A). In the validation dataset, the overall accuracy is 69.33%. The accuracy for predicting ovulation within 24, 48, and 72 h is 95.36%, 38.84%, and 64.81% respectively. Sensitivity values are 76.60%, 70.15%, 49.30%, while specificity values are 94.93%,71.43%, and 92.55% respectively. PPV values are 95.36%, 38.84%, 64.81%, and NPV values are 74.86%, 90.24%, and 86.76% respectively (Table 4). When training predictive model using single LH levels, the cutoff values of LH are 16.75mIU/mL and 30.04 mIU/mL (Fig. 5B). In the validation dataset, the overall accuracy is 69.94%. The accuracy for predicting ovulation within 24, 48, and 72 h is 88.08%, 45.45%, and 74.07% respectively. Sensitivity values are 70.47%, 65.48%, 74.07%, while specificity values are 86.96%,72.73%, 94.85% respectively. PPV values are 88.08%, 45.45%, 74.07%, and NPV values are 68.57%, 85.85%, and 94.84% respectively (Table 4). While P4 and LH had similar overall accuracy, P4 was more effective than LH in predicting ovulation within 24 h. Models using single E2 levels or follicle diameter had low accuracy (<60%, Table 4). The cutoff values for E2 and diameter were 359.9 pg/ml and 17.75 mm, respectively (Fig. 5C and D).

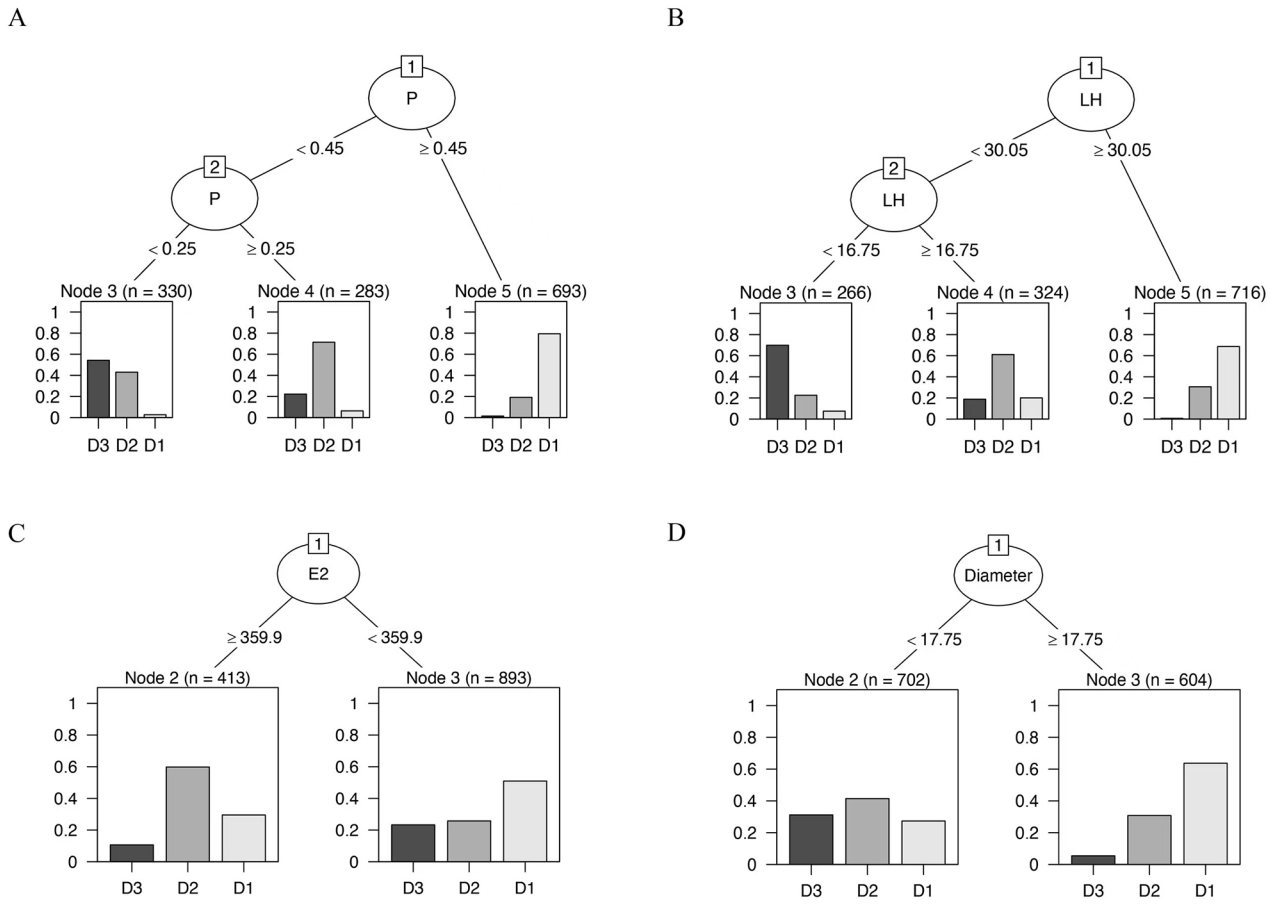

Fig. 5 Regression tree plot of classification trees model with each single variable. **A** classifications trees model with progesterone level (P, ng/ml). **B** classifications trees model with luteinizing hormone level (LH, mIU/ml). **C** classifications trees model with Estradiol level (E2, pg/ml). **D** classifications trees model with follicle diameter (mm)

Discussion

Accurately predicting the ovulation time is critical for women undergoing NC-FET. Our two predictive models both indicate that preovulatory serum P4 levels are better predictors of ovulation timing than LH levels. Furthermore, we have devised a convenient and practical model for predicting ovulation by combining the relationship of P levels with LH and E2 levels, which can provide fertility physicians with more standardized and precise tools for identifying ovulation time based on the P levels.

Accurately predicting the ovulation timing remains challenging due to the lack of consensus among different clinics regarding the definition of this significant event [12, 13]. Current literatures lack extensive documentations on using preovulatory serum P4 for ovulation prediction. This knowledge gap primarily stems from the prevailing belief that the rise of E2 initiates the LH surge, which is subsequently leading to the ovulation. Additionally, the pulsatile nature of P4 secretion further complicates its role in ovulation prediction. Many studies had revealed that serum P4 concentrations are low and relatively constant throughout the follicular phase and begin to increase as ovulation approaches [21, 22]. Hoff and his colleagues had evaluated the dynamics of ovarian and pituitary hormone changes during the midcycle period and demonstrated that a rapid rise of P4 occurred approximately 12 h precedes both the E2 peak and LH flare and continues at a variable rate throughout and beyond the LH surge, the P4 surge level was small, only approximately 0.5 ng/ml [22]. Maman et al. found that P4 levels>2 nmol/L had a high sensitivity of 91.5% but low specificity of 62.7% in predicting ovulation the following day [23]. Another study also found the rise in P4>4 nmol/l gives a sufficiently accurate prediction of ovulation within 24 h [24]. Our study found that preovulatory P4 levels were superior to LH levels as an ovulation predictor. Both models indicate that P4 levels are the most significant variable in predicting the ovulation time. If the P4 level is ≥0.65 ng/ml, ovulation will occur within 24 h in more than 90% of patients. According to the validation dataset, the classification trees achieved an accuracy rate of 92.72%, while the random forest model achieved

Table 4 Effectiveness of each variable in predicting ovulation time in the validation dataset using classification trees model

P4				
	72 h (Real)	**48 h (Real)**	**24 h (Real)**	**Total**
72 h (Predictive)	35	34	2	
48 h (Predictive)	15	47	5	
24 h (Predictive)	4	40	144	
Total	54	121	151	326
Accuracy rate	0.6481	0.3884	0.9536	0.6933
Sensitivity	0.4930	0.7015	0.7660	
Specificity	0.9255	0.7143	0.9493	
PPV	0.6481	0.3884	0.9536	
NPV	0.8676	0.9024	0.7486	
LH				
	72 h (Real)	**48 h (Real)**	**24 h (Real)**	**Total**
72 h (Predictive)	40	12	2	
48 h (Predictive)	13	55	16	
24 h (Predictive)	1	54	133	
Total	54	121	151	326
Accuracy rate	0.7407	0.4545	0.8808	0.6994
Sensitivity	0.7407	0.6548	0.7074	
Specificity	0.9485	0.7273	0.8696	
PPV	0.7407	0.4545	0.8808	
NPV	0.9485	0.8585	0.6857	
E2				
	72 h (Real)	**48 h (Real)**	**24 h (Real)**	**Total**
72 h (Predictive)	0	0	0	
48 h (Predictive)	14	70	41	
24 h (Predictive)	40	51	110	
Total	54	121	151	326
Accuracy rate	NA	0.5785	0.7285	0.5521
Sensitivity	NA	0.5600	0.5473	
Specificity	0.8344	0.7463	0.6720	
PPV	NA	0.5785	0.7285	
NPV	NA	0.7317	0.4800	
Diameter				
	72 h (Real)	**48 h (Real)**	**24 h (Real)**	**Total**
72 h (Predictive)	0	0	0	
48 h (Predictive)	47	64	52	
24 h (Predictive)	7	52	99	
Total	54	121	151	326
Accuracy rate	NA	0.5289	0.6556	0.50
Sensitivity	NA	0.3926	0.6074	
Specificity	0.8344	0.6503	0.6810	
PPV	NA	0.5289	0.6556	
NPV	NA	0.5171	0.6343	

Notes: 72 h (Real): number of records in 72 h group in real situation; 48 h (Real): number of records in 48 h group in real situation; 24 h (Real): number of records in 24 h group in real situation; 72 h (Predictive): number of records predicted in the 72 h group according to the classification trees model; 48 h (Predictive): number of records predicted in the 48 h group according to the classification trees model; 24 h (Predictive): number of records predicted in the 24 h group according to the classification trees model. PPV: positive predictive value; NPV: negative predictive value

an accuracy rate of 96.69%. Although the P4 secretion is thought to be pulsatile and values measured in short periods may fluctuate, our study reveals a steady and progressive rise in P4 levels from three days prior to ovulation until the day of ovulation, with minimal fluctuations, consistent with the existing literature [25]. These findings indicate that preovulatory P4 is a reliable marker for ovulation prediction.

Furthermore, the efficient prediction of ovulation can be achieved when the preovulatory P4 level falls within the range of 0.45 to 0.65 ng/ml, in combination with E2 levels. Numerous studies have consistently shown that during the follicular development, E2 levels gradually rise to a peak then rapidly decline until ovulation occurs. Therefore, the decrease in preovulatory E2 level serves as a significant indicator that ovulation is likely to occur the following day [26, 27]. Maman E., et al. found that in approximately 19% of cases, the decline in E2 levels only happened on the day of ovulation. They proposed that in instances where no decline is observed, LH and P4 levels should be considered for ovulation prediction. Our findings revealed that the E2 levels in the 48 h group are significantly higher compared to other groups, while the E2 levels between the 72 h and 24 h groups exhibit no significant differences (p-value = 0.842). However, the E2 peak levels varied significantly among different individuals, suggesting limitations in using E2 alone for predicting ovulation. Nevertheless, our models demonstrated that a preovulatory P4 level ranging from 0.45 to 0.65 ng/ml, in combination with an E2 level ≥ 360.6 pg/ml, strongly predicts the occurrence of ovulation within 48 h. If the E2 levels begin to decrease following an E2 surge, the presence of a preovulatory P4 level between 0.45 and 0.65 ng/ml, combined with an E2 level < 360 pg/ml, indicates a high likelihood of ovulation occurring the following day. The predictive accuracy of our machine learning models, which utilize preovulatory P4 and E2 levels, reaches a certain rate. Consequently, these models enable the precise determination of the optimal timing for embryo transfer.

In the literature, LH surge has been extensively studied and is widely used as a predictor for identifying ovulation. However, a consensus on the standard criteria for defining the LH surge is still pending [12, 16, 26, 28]. Correlatively, in our clinic setting, ultrasound scan combined with hormone measurement are employed to determine the optimal timing for embryo thawed transfer in the NC-FET protocol, ovulation is identified by ultrasound, as illustrated in Fig. 2, LH values at 72 h, 48 h, 24 h prior to ovulation were significantly different (all p-values < 0.0001), the data further demonstrated the presence of substantial variation in LH levels among different patients during the preovulatory period. Our predictive models have revealed that relying solely on

absolute LH levels or relative changes in LH levels does not yield satisfactory predictive values. However, when the LH level is ≥18.5 IU/L combined with P4 <0.45 ng/ml, ovulation is highly likely to occur within 48 h. Further studies are needed to determine if an LH level of 18.5 IU/L can be reliably used as the threshold for identifying the LH surge.

The retrospective design is a main limitation for this study. Additional prospective studies are needed to validate the conclusions of the model. Besides, the exact threshold may vary among laboratories based on the dynamics of the assays being used. More studies are needed to validate the predictive model.

Conclusions

Preovulatory serum P4 levels are better predictors of ovulation timing than LH levels and could be used as an alternative in clinical settings, and the model we developed can be used to pinpoint the day of ovulation. Ongoing research and advancements in technology are anticipated to enhance and refine the ovulation method.

Abbreviations

BMI	Body mass index
E2	Estradiol
ECLIA	Electrochemiluminescence immunoassay
FET	Frozen-thawed embryo transfer
IVF	In vitro fertilization
LH	Luteinizing hormone
NC-FET	Natural cycle - frozen embryo transfer
NC	Natural cycle
P4	Progesterone
WOI	The window of implantation

Author contributions

Yumei Li conceptualized the study and wrote the main manuscript, Hong Zeng constructed predictive models and made statistical analysis in the work and prepared Figs. 1, 2, 3, 4 and 5; Tables 1, 2, 3 and 4, Jing Fu made data extraction. All authors have reviewed the manuscript and approved the submitted version, and agreed to be personally accountable for the author's own contributions.

Funding

The study did not receive any funding.

Data availability

The data that support the findings of this study are available from the authors but restrictions apply to the availability of these data, which were used under license from the Xiangya Hospital Central South University for the current study, and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission from the Xiangya Hospital Central South University .

Declarations

Ethics approval and consent to participate

The study has been performed in accordance with the Declaration of Helsinki, and received approval from the Institutional Review Board at Central South University of Xiangya Hospital (ethics approval number: 2022012). All patients had given consent for their treatment data to be used for analysis.

Consent for publication

This manuscript was original research that has not been published previously, and not under consideration for publication elsewhere, in whole or in part.

All authors on this submission have adhered to all editorial policies for submission as described in the information for Authors, attest to having met all authorship criteria.

Competing interests

The authors declare no competing interests.

Received: 26 February 2024 / Accepted: 14 August 2024
Published online: 21 August 2024

References

1. Wilcox JA, Donna DB, Claricer RW. <time of implantation of the conceptus.pdf>. N Engl J Med. 1999;340:1796–9.
2. Pape J, Levy J, von Wolff M. Early pregnancy complications after frozen-thawed embryo transfer in different cycle regimens: a retrospective cohort study. Eur J Obstet Gynecol Reprod Biol. 2022;279:102–6.
3. Makhijani R, Bartels C, Godiwala P, Bartolucci A, Nulsen J, Grow D, et al. Maternal and perinatal outcomes in programmed versus natural vitrified-warmed blastocyst transfer cycles. Reprod Biomed Online. 2020;41(2):300–8.
4. Hu KL, Zhang D, Li R. Endometrium preparation and perinatal outcomes in women undergoing single-blastocyst transfer in frozen cycles. Fertil Steril. 2021;115(6):1487–94.
5. Asserhøj LL, Spangmose AL, Aaris Henningsen AK, Clausen TD, Ziebe S, Jensen RB, Pinborg A. Adverse obstetric and perinatal outcomes in 1,136 singleton pregnancies conceived after programmed frozen embryo transfer (FET) compared with natural cycle FET. Fertil Steril. 2021;115(4):947–56.
6. Lawrenz B, Coughlan C, Melado L, Fatemi HM. The ART of frozen embryo transfer: back to nature! Gynecol Endocrinol. 2020;36(6):479–83.
7. Roelens C, Blockeel C. Impact of different endometrial preparation protocols before frozen embryo transfer on pregnancy outcomes: a review. Fertil Steril. 2022;118(5):820–7.
8. Diaz-Gimeno P, Ruiz-Alonso M, Blesa D, Bosch N, Martinez-Conejero JA, Alama P, et al. The accuracy and reproducibility of the endometrial receptivity array is superior to histology as a diagnostic method for endometrial receptivity. Fertil Steril. 2013;99(2):508–17.
9. Direito A, Bailly S, Mariani A, Ecochard R. Relationships between the luteinizing hormone surge and other characteristics of the menstrual cycle in normally ovulating women. Fertil Steril. 2013;99(1):279–85 e3.
10. Hertig AT, Rock J, Adams EC. A description of 34 human ova within the first 17 days of development. Am J Anat. 1956;98(3):435–93.
11. Teh WT, McBain J, Rogers P. What is the contribution of embryo-endometrial asynchrony to implantation failure? J Assist Reprod Genet. 2016;33(11):1419–30.
12. Erden M, Mumusoglu S, Polat M, Yarali Ozbek I, Esteves SC, Humaidan P, Yarali H. The LH surge and ovulation re-visited: a systematic review and meta-analysis and implications for true natural cycle frozen thawed embryo transfer. Hum Reprod Update. 2022;28(5):717–32.
13. Wetzels LC, Hoogland HJ. Relation between ultrasonographic evidence of ovulation and hormonal parameters: luteinizing hormone surge and initial progesterone rise. Fertil Steril. 1982;37(3):336–41.
14. Marinho AO, Sallam HN, Goessens LK, Collins WP, Rodeck CH, Campbell S. Real time pelvic ultrasonography during the periovulatory period of patients attending an artificial insemination clinic. Fertil Steril. 1982;37(5):633–8.
15. Shoham Z, Schacter M, Loumaye E, Weissman A, MacNamee M, Insler V. The luteinizing hormone surge–the final stage in ovulation induction: modern aspects of ovulation triggering. Fertil Steril. 1995;64(2):237–51.
16. Godbert S, Miro F, Shreeves C, Gnoth C, Johnson S. Comparison between the different methods developed for determining the onset of the LH surge in urine during the human menstrual cycle. Arch Gynecol Obstet. 2015;292(5):1153–61.
17. Roos J, Johnson S, Weddell S, Godehardt E, Schiffner J, Freundl G, Gnoth C. Monitoring the menstrual cycle: comparison of urinary and serum reproductive hormones referenced to true ovulation. Eur J Contracept Reprod Health Care. 2015;20(6):438–50.
18. Dozortsev DI, Diamond MP. Luteinizing hormone-independent rise of progesterone as the physiological trigger of the ovulatory gonadotropins surge in the human. Fertil Steril. 2020;114(2):191–9.
19. Bartek MA, Saxena RC, Solomon S, Fong CT, Behara LD, Venigandla R, et al. Improving operating room efficiency: machine learning approach to predict case-time duration. J Am Coll Surg. 2019;229(4):346–54. e3.

20. Tuwatananurak JP, Zadeh S, Xu X, Vacanti JA, Fulton WR, Ehrenfeld JM, Urman RD. Machine learning can improve estimation of surgical case duration: a pilot study. J Med Syst. 2019;43(3):44.
21. Laborde N, Carril M, Cheviakoff S, Croxatto HD, Pedroza E, Rosner JM. The secretion of progesterone during the periovulatory period in women with certified ovulation. J Clin Endocrinol Metab. 1976;43(5):1157–63.
22. Hoff JD, Quigley ME, Yen SS. Hormonal dynamics at midcycle: a reevaluation. J Clin Endocrinol Metab. 1983;57(4):792–6.
23. Maman E, Adashi EY, Baum M, Hourvitz A. Prediction of ovulation: new insight into an old challenge. Sci Rep. 2023;13(1):20003.
24. Djahanbakhch O, Swanton IA, Corrie JE, McNeilly AS. Prediction of ovulation by progesterone. Lancet. 1981;2(8256):1164–5.
25. Aydin T, Koroglu N, Albayrak N, Insel MA. Serial progesterone levels more accurately predict the time of ovulation in subfertile women: a prospective cohort study. J Assist Reprod Genet. 2023;40(8):1897–903.
26. Irani M, Robles A, Gunnala V, Reichman D, Rosenwaks Z. Optimal parameters for determining the LH surge in natural cycle frozen-thawed embryo transfers. J Ovarian Res. 2017;10(1):70.
27. Lu X, Khor S, Zhu Q, Sun L, Wang Y, Chen Q, et al. Decrease in preovulatory serum estradiol is a valuable marker for predicting premature ovulation in natural/unstimulated in vitro fertilization cycle. J Ovarian Res. 2018;11(1):96.
28. Youngster M, Luz A, Baum M, Hourvitz R, Reuvenny S, Maman E, Hourvitz A. Artificial intelligence in the service of intrauterine insemination and timed intercourse in spontaneous cycles. Fertil Steril. 2023;120(5):1004–12.

Publisher's note