Observational Study

Medicine

OPEN

Clinical pregnancy outcomes prediction in vitro fertilization women based on random forest prediction model

A nested case-control study

Hongya Yang, MM[a], Fang Liu, MBBS[a], Yuan Ma, MD[a], Man Di, MM[a,*]

Abstract

The present study aimed to analyze the risk factors influencing the in vitro fertilization embryo transfer (IVF-ET) pregnancy and to construct a prediction model for clinical pregnancy outcome in patients receiving IVF-ET based on the predictors. In this nested case-control study, the data of 369 women receiving IVF-ET were enrolled. Univariate and multivariate Logistic regression analyses were conducted to identify the potential predictors. Ten-fold cross validation method was used to validate the random forest model for predicting the clinical pregnancy. The receiver operating characteristic curve was drawn to evaluate the prediction ability of the model. The importance of variables was shown according to Mean Decrease Gini. The data delineated that age (odds ratio [OR]= 1.093, 95% confidence interval [CI]: 1.036–1.156, $P = .0010$), body mass index (BMI) (OR = 1.094, 95%CI: 1.021–1.176, $P = .012$), 3 cycles (OR = 0.144, 95%CI: 0.028–0.534, $P = .008$), hematocrit (HCT) (OR = 0.865, 95% CI: 0.791–0.943, $P = .001$), luteinizing hormone (LH) (OR = 0.678, 95%CI: 0.549–0.823, $P < .001$), progesterone (P) (OR = 2.126, 95%CI: 1.112–4.141, $P = .024$), endometrial thickness (OR = 0.132, 95%CI: 0.034–0.496, $P = .003$) and FSH (OR = 1.151, 95%CI: 1.043–1.275, $P = .006$) were predictors associated with the clinical pregnancy outcome of patients receiving IVF-ET. The results might provide a novel method to identify patients receiving IVF-ET with a high risk of poor pregnancy outcomes and provide interventions in those patients to prevent the occurrence of poor pregnancy outcomes.

Abbreviations: BMI = body mass index, CI = confidence interval, E2 = estradiol, FSH = follicle-stimulating hormone, HCG = human chorionic gonadotropin, IVF-ET = in vitro fertilization embryo transfer, LH = luteinizing hormone, OR = odds ratio, P = progesterone.

Keywords: in vitro fertilization, pregnancy outcomes, random forest prediction model

1. Introduction

Infertility is a common clinical epidemic which has a substantial impact on the physical and mental health of patients and family happiness.[1] Nearly 15% of couples are involuntarily infertile and require fertility treatment worldwide.[2,3] With the development of clinical assisted reproduction technology (ART), in vitro fertilization embryo transfer (IVF-ET) has become an important method for the treatment of infertility.[4] Previous studies have estimated that the number of children born after IVF/intracytoplasmic sperm injection was more than 7 million, which accounts for 1.6% of births in the United States and 1.0% in mainland China.[5,6] Although the clinical pregnancy rate after IVF-ET has been greatly improved, the success rate is only 30% to 40%.[7] The pregnancy failure rate is still very high, which brought great mental stress and financial loss to patients. To improve in IVF-ET rate is of great significance in the clinic.

Researchers have provided multiple evidence that successful IVF-ET pregnancy is affected by various factors, such as the maternal age, body mass index (BMI), gestational number, duration and causes of infertility and endometrial thickness on the day of human chorionic gonadotropin (HCG) injection, the number of retrieved oocytes, the number of embryos transferred, and the hormone levels of luteinizing hormone (LH), estradiol (E2), and progesterone (P).[8–10] Some other studies revealed the clinical pregnancy rate was associated with the environment in uterine cavity and embryo quality.[11] In previous studies, several prediction models were established for predicting the clinical pregnancy rate in patients receiving IVF-ET, and they were focused on predictors including hormone levels such as LH, E2, P and maternal age or predictions using a combination of these indicators.[12] The purpose of this study was to analyze the risk factors influencing the IVF-ET pregnancy according to clinical data including not only hormone levels and maternal age, but

The authors have no funding and conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

[a] *Reproductive Medical Center, Department of Gynecology and Obstetrics, Tangdu Hospital, Air Force Medical University, Xi'an, Shaanxi Province, China.*

* *Correspondence: Man Di, Reproductive Medical Center, Department of Gynecology and Obstetrics, Tangdu Hospital, Air Force Medical University, No. 569, Xinsi Road, Baqiao District, Xi'an, Shaanxi Province 710038 China (e-mail: mandyDii2022@163.com).*

How to cite this article: Yang H, Liu F, Ma Y, Di M. Clinical pregnancy outcomes prediction in vitro fertilization women based on random forest prediction model: A nested case-control study. Medicine 2022;101:49(e32232).

Received: 21 January 2022 / Received in final form: 16 November 2022 / Accepted: 17 November 2022

http://dx.doi.org/10.1097/MD.0000000000032232

also some other demographic and laboratory indicators and to construct a prediction model for clinical pregnancy rate in patients receiving IVF-ET based on the predictors. The findings of our study might provide a clinical guidance for the treatment in infertility patients receiving IVF-ET and improve the success rate of clinical pregnancy in those patients.

2. Materials and Methods

2.1. Study design and population

This study was a nested case-control study, the data of 391 patients were collected in Tangdu Hospital, Air Force Medical University. Patients included were all women of reproductive age and receiving IVF-ET. Those with polycystic ovarin syndrome, low ovarian reserve function, people used hyperhormone drugs and ovulation induction drugs used in the past 3 months, participants with uterine malformations, endometrial abnormal echo, endometriosis, adenomyosis, uterine tuberculosis or uterine adhesions and the infertility caused by men were excluded. Finally, 369 subjects were included. The case group was defined as patients with pregnancy (n = 210) and the control group was defined as patients without pregnancy (n = 159) after IVF-ET according to their clinical pregnant status. The informed consents were obtained from the participates and this study was approved by the ethical committee of Second Affiliated Hospital of Air Force Military Medical University (No. TDLL-KY-202106-05).

2.2. IVF treatment

Patients received a long-acting gonadotropin-releasing hormone agonist follicular protocol for ovarian hyperstimulation with a daily injection of recombinant follicle-stimulating hormone (GONAL-f, Merck, Serono, Italy; Puregon, Organon, the Netherlands) and/or human menopausal gonadotropins (Menopur, Ferring, Germany). When the leading follicles reached a diameter of 17 mm, HCG was administered. After 34 to 36 hour oocytes were retrieved followed by conventional IVF. The presence of 2 pronuclei (2PN) was observed 16 to 18 hours later, which were the indicator of fertilization. Embryos were incubated in G5-medium (Vitrolife, Sweden) and 1 or 2 high-quality embryos were transferred into the uterus 3 to 5 days after retrieving the oocytes. Patients were followed up, and biochemical HCG was performed 14 days after ET. Luteal support was continued if serum HCG > 50 IU/L. Transvaginal ultrasound was conducted 5 weeks after ET.

2.3. Outcome measurement

Pregnancies were diagnosed according to the serum B-hCG levels 14 days after embryo transfer and the clinical pregnancy was confirmed by visualization of a gestational sac on ultrasonographic examination 28 to 35 days after the embryo transfer.[13] The biochemical pregnancy was defined as the serum β-HCG > 25U/L 14 days after embryo transfer. Live birth was considered when a living fetus was born after 28 weeks of pregnancy.

2.4. Data collection

Clinical data of patients were collected including age (years), BMI (kg/m^2), duration of infertility (years), number of pregnancy, number of cycle, infertility reasons (female causes or both causes), red blood cell (RBC, 10^9/L), hemoglobin (g/L), hematocrit (HCT, L/L), alanine aminotransferase (ALT, U/L), glutamic-oxalacetic transaminase (AST, U/L), total bilirubin (TBI, μmol/L), serum albumin (ALB, g/L), LH (mU/mL), E2 (pmol/L), P (nmol/L), endometrial thickness (mm), retrieved oocytes, number of high-quality embryos, number of transplanted embryos, follicle-stimulating hormone (FSH) and anti-Mullerian hormone.

2.5. Statistical analysis

All statistical analyses were performed using R 4.0.3. Statistical tests were performed using 2-side test and $\alpha = 0.05$ was used to determine the statistically significance. The Shapiro test was applied to test the normality of measurement data, and the Mean ± standard deviation (Mean ± SD) was used to analyze the continuous variables of normal distribution and t test was used for comparison between groups. The measurement data of skewness distribution were expressed by M (Q_1, Q_3). The number and frequency of each category were described by Mann–Whitney U test for classification variables between groups. Chi-square test (χ^2) was used for comparison between groups. The occurrence of clinical pregnancy was defined as the dependent variable, and Logistic stepwise regression model was conducted to identify the potential predictors. Those with $P < .1$ was excluded by R and Logistic stepwise regression was used to identify those with $P > .05$. Variables with $P < .05$ were included as predictors. Ten-fold cross validation method was used to validate the random forest prediction model for predicting the clinical pregnancy. The receiver operating characteristic curve was drawn to evaluate the prediction ability of the model. The importance of variables was shown according to Mean Decrease Gini. Higher Mean Decrease Gini value indicated more importance of the variable.

3. Results

3.1. Baseline data of the characteristics in all subjects

In total, 391 patients at reproductive age and receiving IVF-ET in the Second Affiliated Hospital of Air Force Military Medical University were enrolled in our study. Those with polycystic ovarin syndrome (n = 8), low ovarian reserve function (n = 3), people used hyperhormone drugs and ovulation induction drugs used in the past 3 months (n = 5), participants with uterine malformations, endometrial abnormal echo, endometriosis, adenomyosis, uterine tuberculosis or uterine adhesions (n = 6) and the infertility caused by men (n = 0) were excluded. Finally, 369 subjects were included. The detailed screen process was shown in Figure 1. For all participants, the average age of all participants was 31.53 ± 4.39 years old. The average BMI was 22.61 ± 3.41 kg/m^2 and the average duration time of infertility was 3.49 ± 2.86 years. Among all patients, 303 (82.11%) had 1 IVF cycle, 47 (12.73%) had 2 IVF cycles, 15 (4.07%) had 3 IVF cycles and 4 (1.09%) had 4 IVF cycles. 254 (68.83%) patients with the pregnancy times ≤ 1. In terms for the reasons of infertility, 262 (71.00%) patients were unable to get pregnant because of the woman, and 107 (29.00%) were unable to get pregnant because of both the women and men. The follow-up time was 40 days, and no cases were lost to follow-up. Finally, 210 (56.91%) patients were pregnant after the IVF-ET, and 159 (43.09%) patients were not pregnant (Table 1). There were 169 participants had live birth.

3.2. Comparisons of characteristics in different clinical pregnancy outcomes

As depicted in Table 2, the mean age of women in the control group was older than women in the case group (32.45 years vs 30.83 years, t = −3.49, $P < .001$), and the proportion of women in the control group who had > 1 pregnancy was higher than women in the case group (16.26% vs 14.91%, $\chi^2 = 5.10$, $P = .023$). The levels of RBC (4.47 10^9/L vs 4.36 10^9/L, Z = 2.48, $P = .013$), hemoglobin (132.26 g/L vs 129.67 g/L, t = 2.27, $P = .024$),

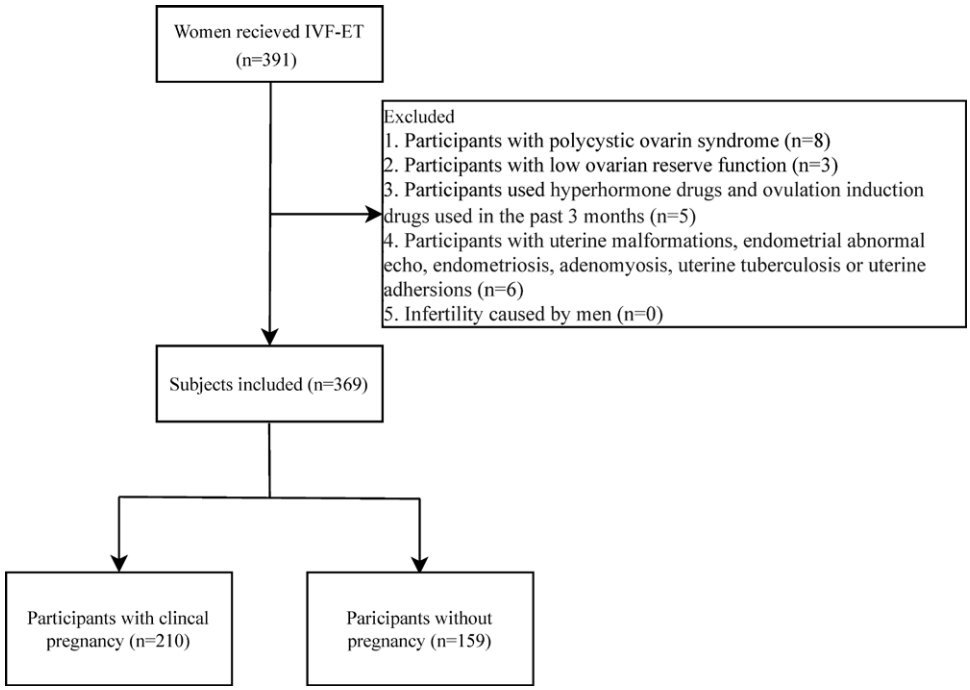

Figure 1. The screen process of the participants in this study.

Table 1

Baseline data of the characteristics in all subjects.

Variables	Description (n = 369)
Age (Mean ± SD)	31.53 ± 4.39
BMI (Mean ± SD)	22.61 ± 3.41
Duration of infertility (Mean ± SD)	3.49 ± 2.86
Number of cycle, n (%)	
1	303 (82.11)
2	47 (12.73)
3	15 (4.07)
4	4 (1.09)
Number of pregnancies, n (%)	
≤1	254 (68.83)
>1	115 (31.17)
Infertility reasons, n (%)	
Male	0 (0)
Female	262 (71.00)
Both	107 (29.00)
Clinical pregnancy, n (%)	
Yes	210 (56.91)
No	159 (43.09)

BMI = body mass index.

HCT (40.01 L/L vs 39.26 L/L, t = 2.58, $P = .010$), TBI (11.08 μmol/L vs 9.65 μmol/L, Z = –2.75, $P = .006$), ALB (44.80 g/L vs 43.73 g/L, Z = 3.01, $P = .003$) and LH (2.28 mU/mL vs 1.9 mU/mL, Z = 3.03, $P = .002$) in the case group were higher than in the control group, and endometrium (1.03 mm vs 0.98 mm, t = 2.70, $P = .007$) was thicker in the case group than in the control group. The levels of AST (22.00 U/L vs 24.00 U/L, Z = –2.60, $P = .009$) and P (0.64 nmol/L vs 0.73 nmol/L, Z = –2.67, $P = .008$) in the case group were lower than in the control group.

3.3. Multivariable logistic stepwise regression of factors associated with clinical pregnancy

According to the results from multivariable Logistic stepwise regression analysis, age (OR = 1.093, 95%CI: 1.036–1.156, $P = .001$), BMI (OR = 1.094, 95%CI: 1.021–1.176, $P = .012$), 3 cycles (OR = 0.144, 95%CI: 0.028–0.534, $P = .008$), HCT (OR = 0.865, 95% CI: 0.791–0.943, $P = .001$), LH (OR = 0.678, 95%CI: 0.549–0.823, $P < .001$), P (OR = 2.126, 95%CI: 1.112–4.141, $P = .024$), endometrial thickness (OR = 0.132, 95%CI: 0.034–0.496, $P = .003$) and FSH (OR = 1.151, 95%CI: 1.043–1.275, $P = .006$) were predictors associated with the clinical pregnancy outcome of patients receiving IVF-ET (Fig. 2).

3.4. The construction of random forest model for predicting the clinical pregnancy outcomes

All patients were randomly divided into the training set and the testing set in a ratio of 8:2. The results of the equilibrium test of training set and testing set were exhibited in Table 3, showing that no significant difference was observed in the characteristics between patients in the training set and testing set.

Based on the results of multivariable Logistic stepwise regression analysis, factors including age, BMI, 3 cycles, HCT, LH, P, endometrial thickness, and FSH were involved in the random forest model. The importance of variables was evaluated via Mean Decrease Gini and the data revealed that HCT was the most important variable affecting the clinical pregnancy outcomes followed by the level of and endometrial thickness (Fig. 3). Ten-fold cross-validation was used to validate the results of the model. As depicted in Figure 4, the AUC value of random forest model was 0.940 (95%CI: 0.914–0.966), the sensitivity was 0.797 (95%CI: 0.724–0.869) and the specificity was 0.932 (95%CI: 0.891–0.973). The positive predictive value (PPV) and negative predictive value (NPV) were 0.904 (95%CI: 0.847–0.961) and 0.851 (95%CI: 0.796–0.906), respectively. In the testing set, the AUC value was 0.817 (95%CI: 0.716–0.918), the sensitivity was 786 (95%CI: 0.634–0.938), the specificity was 0.717(95%CI: 0.587–0.848), the PPV was 0.629 (95%CI: 0.468–0.789) and the NPV was 0.846 (95%CI: 0.733–0.959) (Table 4).

4. Discussion

In the current study, clinical data of 369 patients at reproductive age and receiving IVF-ET were collected to analyze the factors

Table 2

Comparisons of characteristics in different clinical pregnancy outcomes.

Variable	Total (n = 369)	Group		Statistical magnitude	*P*
		Patients with pregnancy (n = 210)	Patients without pregnancy (n = 159)		
Baseline characteristics					
Age, M (Q_1,Q_3)	31.53 ± 4.39	30.83 ± 4.01	32.45 ± 4.70	t = −3.49	<.001
BMI, M (Q_1,Q_3)	22.61 ± 3.41	22.43 ± 3.34	22.84 ± 3.49	t = −1.14	.254
Duration of infertility, M (Q_1,Q_3)	3.00 (1.00, 5.00)	3.00 (2.00, 5.00)	2.00 (1.00, 5.00)	*Z* = 0.39	.696
Number of pregnancies, n (%)				χ^2 = 5.10	.023
≤1	254 (68.83)	155 (42.01)	99 (26.82)		
>1	115 (31.17)	55 (14.91)	60 (16.26)		
Number of cycle, n (%)				-	.261
1	303 (82.11)	169 (45.80)	134 (36.31)		
2	47 (12.73)	26 (7.04)	21 (5.69)		
3	15 (4.07)	12 (3.25)	3 (0.82)		
4	4 (1.09)	3 (0.82)	1 (0.27)		
Infertility reasons, n (%)				χ^2 = 0.001	1.000
Female	262 (71.00)	149 (40.38)	113 (30.62)		
Both	107 (29.00)	61 (16.53)	46 (12.47)		
Blood routine index					
Red blood cell, M (Q_1,Q_3)	4.43 (4.20, 4.62)	4.47 (4.25, 4.65)	4.36 (4.17, 4.57)	Z = 2.48	.013
Hemoglobin, (Mean ± SD)	131.15 ± 10.98	132.26 ± 11.17	129.67 ± 10.59	t = 2.27	.024
HCT, (Mean ± SD)	39.69 ± 2.76	40.01 ± 2.73	39.26 ± 2.76	t = 2.58	.010
Liver function index					
ALT, M (Q_1,Q_3)	23.00 (17.00, 29.00)	21.00 (16.00, 30.00)	24.00 (19.00, 29.00)	Z = −1.73	.084
AST, M (Q_1,Q_3)	23.00 (20.00, 27.00)	22.00 (20.00, 26.00)	24.00 (21.00, 28.00)	Z = −2.60	.009
TBI, M (Q_1,Q_3)	10.61 (8.09, 13.83)	11.08 (8.53, 14.25)	9.65 (7.65, 12.64)	Z = −2.75	.006
ALB, M (Q_1,Q_3)	44.20 (42.30, 45.90)	44.80 (42.73, 46.30)	43.73 (41.95, 45.30)	Z = 3.01	.003
Gonadal hormone concentrations					
LH, M (Q_1,Q_3)	2.09 (1.45, 3.02)	2.28 (1.57, 3.19)	1.90 (1.40, 2.76)	Z = 3.03	.002
E2, M (Q_1,Q_3)	2854.00 (2034.00, 3860.00)	2854.00 (2054.00, 3873.5)	2854.00 (1973.50, 3774.40)	Z = 0.16	.873
P, M (Q_1,Q_3)	0.70 (0.50, 1.02)	0.64 (0.49, 0.86)	0.73 (0.50, 1.16)	Z = −2.67	.008
IVF data					
Endometrial thickness, (Mean ± SD)	1.01 ± 0.18	1.03 ± 0.17	0.98 ± 0.18	t = 2.70	.007
Retrieved oocytes, M (Q_1,Q_3)	9 (7, 12)	9 (7, 12)	9 (7, 12)	Z = −0.38	.704
Number of high-quality embryos, M (Q_1,Q_3)	4 (2, 6)	4 (3, 6)	4 (2, 6)	Z = 1.45	.148
Number of transplanted embryos, n (%)				-	.087
1	91 (24.66)	43 (11.65)	48 (13.08)		
2	274 (74.25)	165 (44.71)	109 (29.54)		
3	4 (1.09)	2 (0.54)	2 (0.54)		
ICSI, n (%)				χ^2 = 0.375	.829
Yes	15 (4.07)	8 (3.81)	7 (4.40)		
No	300 (81.30)	173 (82.38)	127 (79.87)		
Unknown	54 (14.63)	29 (13.81)	25 (15.72)		
FSH, M (Q_1,Q_3)	6.32 (5.13, 7.52)	6.18 (4.85, 7.48)	6.48 (5.31, 7.55)	Z = 1.571	.116
AMH, M (Q_1,Q_3)	2.38 (1.46, 3.80)	2.45 (1.53, 4.02)	2.27 (1.35, 3.77)	Z = −0.945	.345
Live birth n (%)					
Yes	169 (45.80)	169 (80.48)	0 (0)		
No	200 (54.20)	41 (19.52)	159 (100%)		

ALB = albumin, ALT = alanine aminotransferase, AMH = anti-Mullerian hormone, AST = glutamic-oxalacetic transaminase, BMI = body mass index, E2 = estradiol, FSH = follicle-stimulating hormone, HCT = hematocrit, ICSI = intracytoplasmic sperm injection, LH = luteinizing hormone, P = progesterone, RBC = red blood cell, TBI = total bilirubin.

influencing the clinical pregnancy outcomes and established a random forest prediction model to predict the clinical pregnancy outcomes of those patients. The results delineated that age, BMI, 3 cycles, HCT, LH, P, FSH, and endometrial thickness were factors affecting the clinical pregnancy outcomes of patients receiving IVF-ET. The random forest prediction model had good predictive value for distinguishing patients with high risk of poor clinical pregnancy outcome and may provide a reference for the obstetricians to make timely intervention to prevent the occurrence of poor clinical pregnancy outcome in those patients.

A successful pregnancy depends on embryo quality, endometrial receptivity, and the interaction between embryo and maternal endometrium, and endometrial thickness was the most commonly investigated marker of endometrial receptivity.[14] Previously, the evidence from several meta-analyses indicated that decreased endometrial thickness led to a poor pregnancy outcome in patients receiving IVF.[15,16] The findings of these studies were allied with the results in our study, which depicted that endometrial thickness was a protective factor of good clinical pregnancy outcome and thicker endometrial thickness increased the possibility of good clinical pregnancy outcome in patients receiving IVF-ET. LH is involved in the growth and maturation of follicle.[17] High levels of LH trigger theca cells of developing follicles to produce androgens and polypeptide growth factors and increase the follicular response to follicle-stimulating hormone during follicular recruitment and selection as well as stimulate the growth of large antral follicles.[18] As observed in this study, higher LH levels on the day of HCG injection were associated with good clinical pregnancy outcome in patients receiving IVF-ET. Current studies have shown that P level on the day

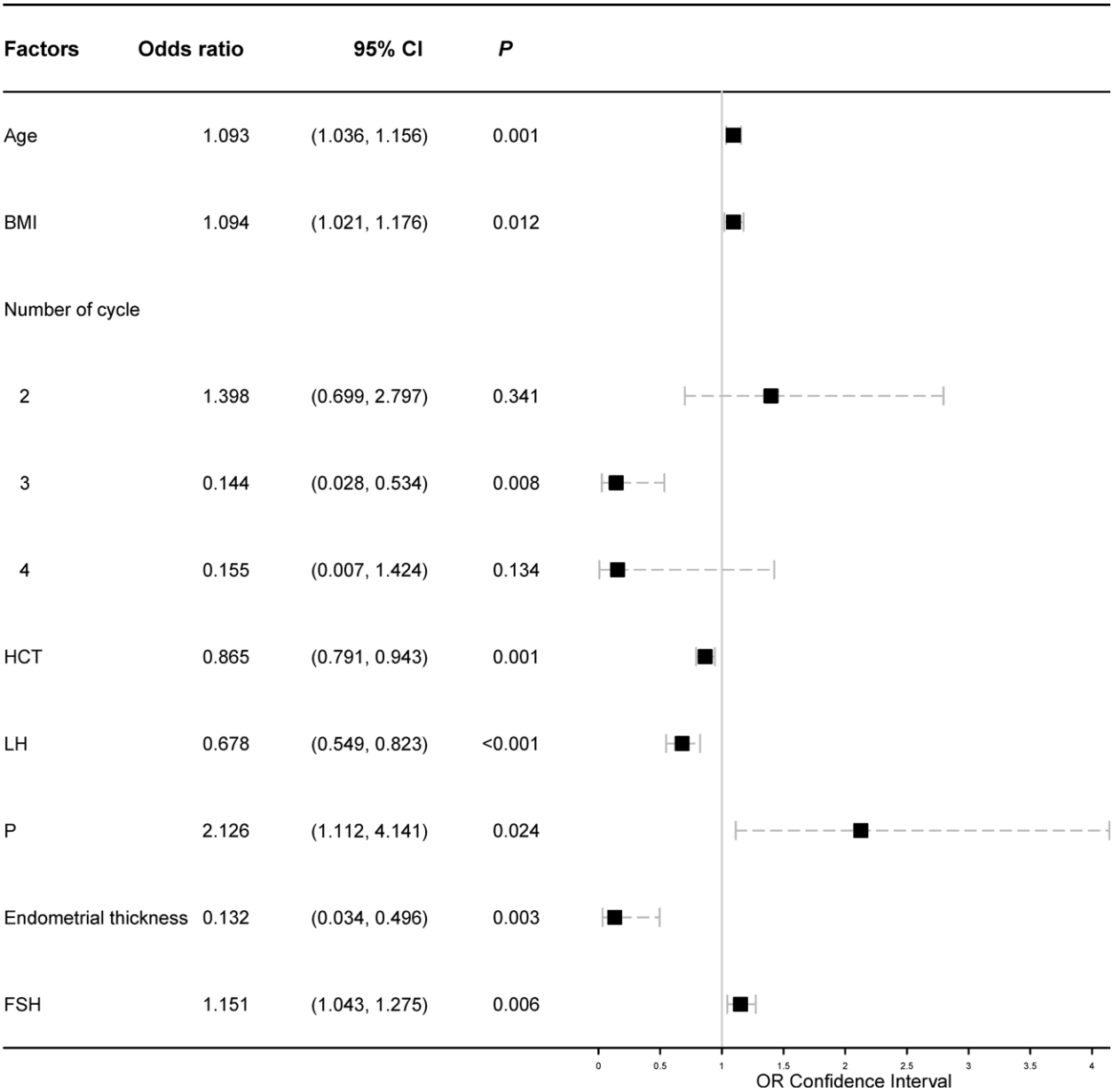

Figure 2. Forest plot of multivariate logistic regression model.

Table 3

The equilibrium test of training set and testing set.

Variable	Total (n = 369)	Testing set (n = 74)	Training set (n = 295)	Statistical magnitude	*P*
Age, Mean ± SD	31.53 ± 4.39	31.27 ± 4.69	31.59 ± 4.31	t = –0.56	.576
BMI, Mean ± SD	22.61 ± 3.41	22.43 ± 3.38	22.65 ± 3.42	t = –0.68	.622
Number of cycle, M (Q_1, Q_3)	1.00 (1.00, 1.00)	1.00 (1.00,1.00)	1.00 (1.00,1.00)	Z = –0.454	.649
HCT, Mean ± SD	39.69 ± 2.76	39.46 ± 2.68	39.74 ± 2.79	t = 0.77	.440
LH, M (Q_1, Q_3)	2.09 (1.45,3.02)	1.90 (1.10, 2.94)	2.10 (1.51, 3.05)	Z = –1.922	.055
P, M (Q_1, Q_3)	0.70 (0.50,1.02)	0.69 (0.47, 0.97)	0.70 (0.50, 1.04)	Z = –0.457	.648
Endometrial thickness, Mean ± SD	1.01 ± 0.18	0.99 ± 0.17	1.01 ± 0.18	t = 1.02	.309
FSH, M (Q_1, Q_3)	6.32 (5.13, 7.52)	6.32 (5.34, 7.43)	6.34 (5.04, 7.59)	Z = 0.023	.982
Clinical pregnancy, n (%)				χ^2 = 1.041	.308
No	210 (56.91)	46 (62.16)	164 (55.59)		
Yes	159 (43.09)	28 (37.84)	131 (44.41)		

BMI = body mass index, E2 = estradiol, LH = luteinizing hormone, P = progesterone.

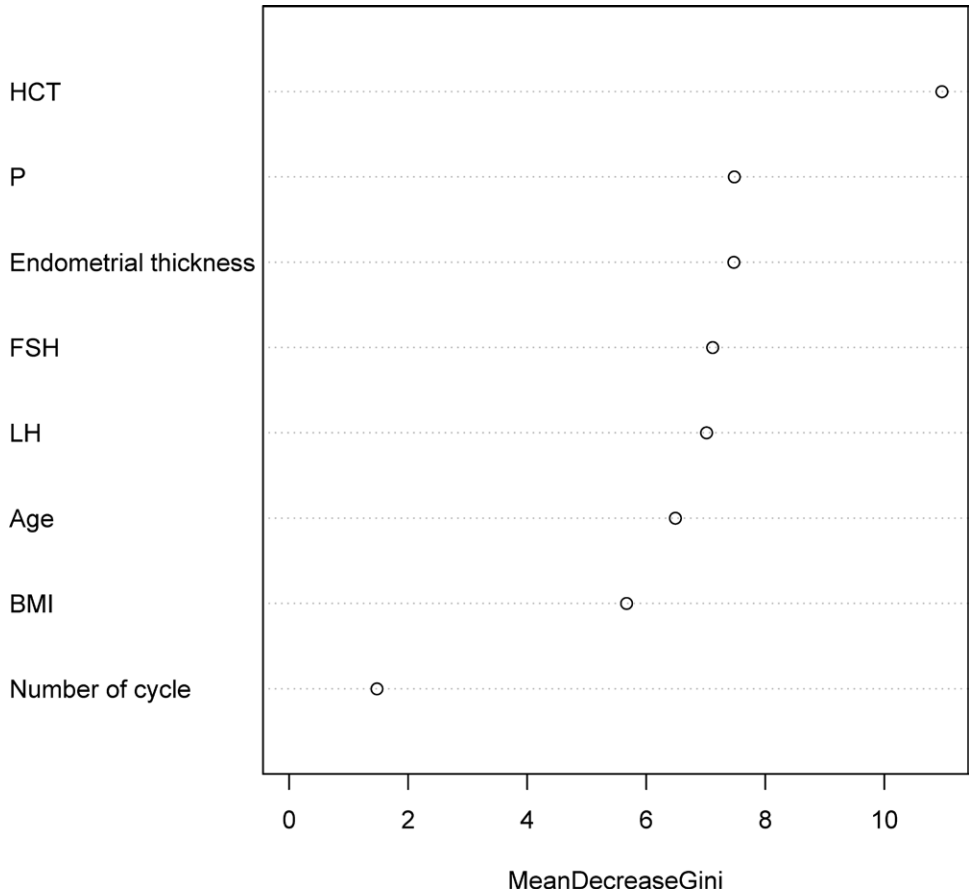

Figure 3. The variable importance assessed by Mean Decrease Gini.

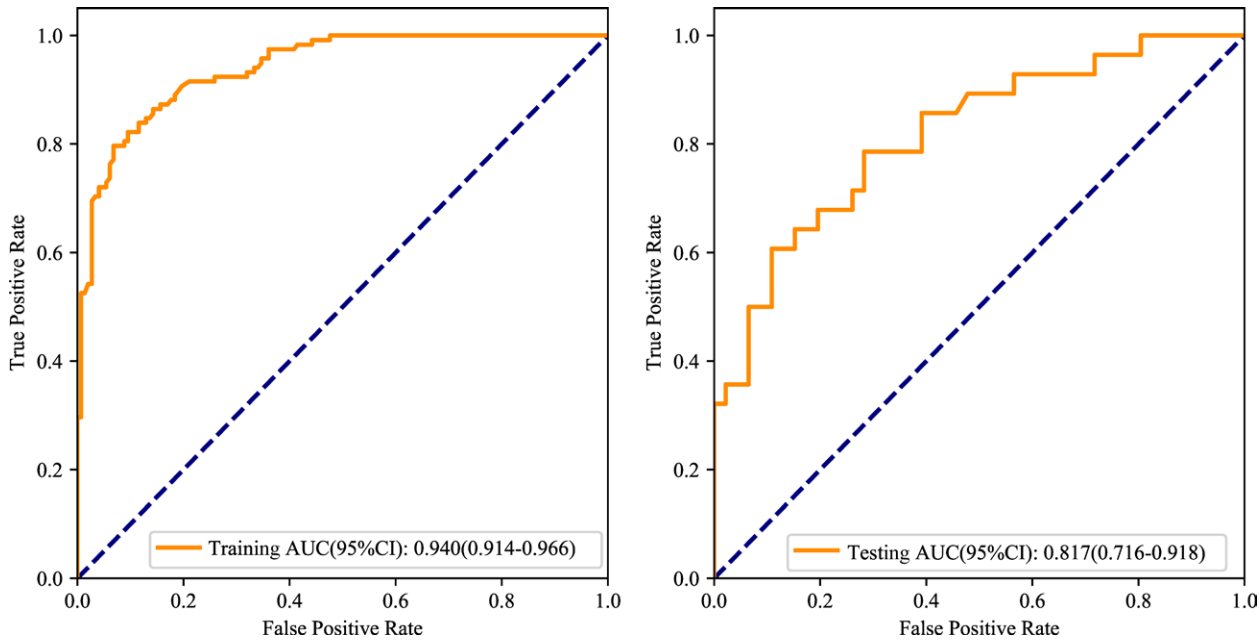

Figure 4. ROC curve on the predicative value of the random forest model. ROC = receiver operating characteristic curve.

of HCG injection was a predictor of P elevation.[19] A previous study including 8649 IVF cycles revealed that the elevated P levels were associated with a lower pregnancy rate.[20] A cohort study by Venetis et al included 3296 patients receiving IVF and showed that elevated P level increased the risk of obstetric complications and decreased the pregnancy rates.[21] Here in the current study, P level was an independent risk factor for poor pregnancy outcome in patients with IVF-ET. During a cycle of IVF, women receive daily doses of FSH to induce multi-follicular development in the ovaries, which is associated with the

Table 4

The predictive value of random forest model for the clinical pregnancy outcomes.

Predictive value	Training set	Testing set
AUC (95%CI)	0.940 (0.914–0.966)	0.817 (0.716–0.918)
Sensitivity (95%CI)	0.797 (0.724–0.869)	0.786 (0.634–0.938)
Specificity (95%CI)	0.932 (0.891–0.973)	0.717 (0.587–0.848)
PPV (95%CI)	0.904 (0.847–0.961)	0.629 (0.468–0.789)
NPV (95%CI)	0.851 (0.796–0.906)	0.846 (0.733–0.959)

AUC = area under the curve, NPV = negative predictive value, PPV = positive predictive value.

quantity and quality of oocytes retrieved.[22] In the present study, the level of FSH was also a vital predictor for pregnancy outcome in patients receiving IVF-ET.

Normal values of HCT in women of childbearing age ranges from 36% to 48% and patients with a decreased HCT may be due to anemia.[23] As shown in our results, higher HCT may be associated with good pregnancy outcome in patients receiving IVF-ET. This suggested that obstetricians should be careful with patients with low HCT or anemia and make timely interventions to increase HCT. Age of women affecting the clinical pregnancy outcomes were frequently reported by a variety of studies,[24,25] which supported the results in the current study. Patients receiving IVF-ET with older age should be with caution. Overweight and obesity was associated with decreased pregnancy in women receiving IVF.[26,27] Li et al revealed that polycystic ovary syndrome patients with dyslipidemia having a higher BMI increased the dosage of gonadotropin and decreased the clinical pregnancy rate.[28] Qian et al found that increased BMI might affect ovulation induction response in early follicular phase prolonged protocol IVF/intracytoplasmic sperm injection patients, leading to the increase of gonadotropin dosage and the extension of gonadotropin induction days, but no significant difference in pregnancy outcome was found among different BMI groups.[29] In this study, higher BMI was associated with lower pregnancy rate, women with high BMI should be reminded to lose weight and control their BMI. A previous study also identified that women undergoing a controlled ovarian hyperstimulation did not show any alteration to renal and hepatic functions, and in the current study, liver function was not found to be important variables associated with the pregnancy outcome in women receiving IVF. The current coronavirus disease 2019 (COVID-19) pneumonia pandemic was reported to have effects throughout pregnancy including increased risk of hypoxemic respiratory failure, and thromboembolic events with associated mortality.[30,31] Global efforts are required to decrease the effects of COVID-19 pneumonia on the women undergoing IVF-ET, labor and neonatal health and fetal growth and development.[32]

The present study assessed factors associated with the clinical pregnancy outcomes of patients receiving IVF-ET and established a random forest model for predicting the pregnancy outcomes of patients receiving IVF-ET. We included not only hormone levels and maternal age, but also some other demographic and laboratory indicators as predictors in the random forest model, and the results might be more reliable than previous studies. The AUC value of random forest model was 0.940, which had a high accuracy for predicting the pregnancy outcomes of patients. The internal validation of the prediction model was also performed in the testing set and the AUC value were 0.817, suggesting the good predictive ability of the model. All the findings demonstrated the random forest model was useful for predicting the pregnancy outcomes of patients receiving IVF-ET, which might help identify patients receiving IVF-ET with a high risk of poor clinical pregnancy outcomes and provide intervention in those patients to prevent the occurrence of poor pregnancy outcomes. Therefore, the clinical pregnancy outcomes might be improved and the expense of the patients might be reduced.

This study was a nested case-control study, the data on exposure and confounding collected before the occurrence of cases, which decreased the potential for recall bias and uncertainty regarding the temporal sequence between exposure and case onset. The limitations of these study were that the sample size was small and external validation of our results was not performed. Studies with large scale of sample size and external validation were required to verify the findings of our study.

This nested case-control study collected the clinical data of 369 patients receiving IVF-ET. The predictors of patients with poor pregnancy outcomes were analyzed and we observed that age, BMI, 3 cycles, HCT, LH, P, endometrial thickness and FSH. A random forest model was established based on the predictors to predict the clinical pregnancy outcomes of patients and exhibited good predictive ability. The results of our study might provide a novel method to identify patients receiving IVF-ET with a high risk of poor clinical pregnancy outcomes and provide intervention in those patients to prevent the occurrence of poor pregnancy outcomes.

Acknowledgments

None.

Author contributions

HY, MD designed the study. HY wrote the manuscript. FL and YM collected, analyzed and interpreted the data. MD critically reviewed, edited and approved the manuscript. All authors read and approved the final manuscript.
Conceptualization: Hongya Yang, Man Di.
Data curation: Fang Liu, Yuan Ma.
Formal analysis: Fang Liu, Yuan Ma.
Investigation: Fang Liu, Yuan Ma.
Methodology: Fang Liu, Yuan Ma.
Supervision: Hongya Yang.
Writing – original draft: Hongya Yang.
Writing – review & editing: Hongya Yang, Man Di.

References

[1] Broughton DE, Moley KH. Obesity and female infertility: potential mediators of obesity's impact. Fertil Steril. 2017;107:840–7.
[2] Thoma ME, McLain AC, Louis JF, et al. Prevalence of infertility in the United States as estimated by the current duration approach and a traditional constructed approach. Fertil Steril. 2013;99:1324–1331.e1.
[3] Houtchens MK, Edwards NC, Hayward B, et al. Live birth rates, infertility diagnosis, and infertility treatment in women with and without multiple sclerosis: data from an administrative claims database. Mult Scler Relat Disord. 2020;46:102541.
[4] Sacha CR, Harris AL, James K, et al. Placental pathology in live births conceived with in vitro fertilization after fresh and frozen embryo transfer. Am J Obstet Gynecol. 2020;222:360.e1–360.e16.
[5] Sunderam S, Kissin DM, Crawford SB, et al. Assisted reproductive technology surveillance - United States, 2013. MMWR Surveill Summ. 2015;64:1–25.
[6] Yang X, Li Y, Li C, et al. Current overview of pregnancy complications and live-birth outcome of assisted reproductive technology in mainland China. Fertil Steril. 2014;101:385–91.
[7] Bidzan M, Podolska M, Bidzan L, et al. [Personality traits and the feeling of loneliness of women treated for infertility]. Ginekol Pol. 2011;82:508–13.
[8] Koedooder R, Singer M, Schoenmakers S, et al. The ReceptIVFity cohort study protocol to validate the urogenital microbiome as predictor for IVF or IVF/ICSI outcome. Reprod Health. 2018;15:202.
[9] Vaegter KK, Lakic TG, Olovsson M, et al. Which factors are most predictive for live birth after in vitro fertilization and intracytoplasmic sperm injection (IVF/ICSI) treatments? Analysis of 100 prospectively recorded variables in 8,400 IVF/ICSI single-embryo transfers. Fertil Steril. 2017;107:641–648.e2.

[10] Goldman RH, Farland LV, Thomas AM, et al. The combined impact of maternal age and body mass index on cumulative live birth following in vitro fertilization. Am J Obstet Gynecol. 2019;221:617.e1617.e611–617.e13.
[11] Paulson RJ, Sauer MV, Lobo RA. Factors affecting embryo implantation after human in vitro fertilization: a hypothesis. Am J Obstet Gynecol. 1990;163:2020–3.
[12] Kim SK, Kim H, Oh S, et al. Development of a novel nomogram for predicting ongoing pregnancy after in vitro fertilization and embryo transfer. Obstet Gynecol Sci. 2018;61:669–74.
[13] Zhao F, Lan Y, Chen T, et al. Live birth rate comparison of three controlled ovarian stimulation protocols for in vitro fertilization-embryo transfer in patients with diminished ovarian reserve after endometrioma cystectomy: a retrospective study. J Ovarian Res. 2020;13:23.
[14] Gadalla MA, Huang S, Wang R, et al. Effect of clomiphene citrate on endometrial thickness, ovulation, pregnancy and live birth in anovulatory women: systematic review and meta-analysis. Ultrasound Obstet Gynecol. 2018;51:64–76.
[15] Craciunas L, Gallos I, Chu J, et al. Conventional and modern markers of endometrial receptivity: a systematic review and meta-analysis. Hum Reprod Update. 2019;25:202–23.
[16] Gao G, Cui X, Li S, et al. Endometrial thickness and IVF cycle outcomes: a meta-analysis. Reprod Biomed Online. 2020;40:124–33.
[17] Schwarze JE, Crosby JA, Zegers-Hochschild F. Addition of neither recombinant nor urinary luteinizing hormone was associated with an improvement in the outcome of autologous in vitro fertilization/intracytoplasmatic sperm injection cycles under regular clinical settings: a multicenter observational analysis. Fertil Steril. 2016;106:1714–1717.e1.
[18] Revelli A, Chiado A, Guidetti D, et al. Outcome of in vitro fertilization in patients with proven poor ovarian responsiveness after early vs. mid-follicular LH exposure: a prospective, randomized, controlled study. J Assist Reprod Genet. 2012;29:869–75.
[19] Venetis CA, Kolibianakis EM, Bosdou JK, et al. Basal serum progesterone and history of elevated progesterone on the day of hCG administration are significant predictors of late follicular progesterone elevation in GnRH antagonist IVF cycles. Hum Reprod. 2016;31:1859–65.
[20] Shufaro Y, Sapir O, Oron G, et al. Progesterone-to-follicle index is better correlated with in vitro fertilization cycle outcome than blood progesterone level. Fertil Steril. 2015;103:669–74.e3.
[21] Venetis CA, Kolibianakis EM, Bosdou JK, et al. Estimating the net effect of progesterone elevation on the day of hCG on live birth rates after IVF: a cohort analysis of 3296 IVF cycles. Hum Reprod. 2015;30:684–91.
[22] Luo X, Pei L, He Y, et al. High initial FSH dosage reduces the number of available cleavage-stage embryos in a GnRH-antagonist protocol: real-world data of 8,772 IVF cycles from China. Front Endocrinol. 2022;13:986438.
[23] Khoigani MG, Goli S, Hasanzadeh A. The relationship of hemoglobin and hematocrit in the first and second half of pregnancy with pregnancy outcome. Iran J Nurs Midwifery Res. 2012;17:S165–70.
[24] Robillard PY. Epidemiological evidence that severe obese women (pre-pregnancy BMI ≥40 kg/m(2)) should lose weight during their pregnancy. J Matern Fetal Neonatal Med. 2022;35:6618–23.
[25] Liu J, Zheng J, Lei YL, et al. Effects of endometrial preparations and transferred embryo types on pregnancy outcome from patients with advanced maternal age. Syst Biol Reprod Med. 2019;65:181–6.
[26] García-Ferreyra J, Carpio J, Zambrano M, et al. Overweight and obesity significantly reduce pregnancy, implantation, and live birth rates in women undergoing In Vitro Fertilization procedures. JBRA Assist Reprod. 2021;25:394–402.
[27] Di Gregorio LM, Zambrotta E, Di Guardo F, et al. Effects of body mass index and biochemical lipid levels on reproductive outcomes during an intracytoplasmic sperm injection: a retrospective study. Int J Fertil Steril. 2019;13:190–5.
[28] Li X, Ding W, Liu JY, et al. Effects of dyslipidemia on IVF/ICSI pregnancy outcome in patients with polycystic ovary syndrome. Zhonghua Fu Chan Ke Za Zhi. 2018;53:402–8.
[29] Qian Y, Zhang J, Jiang CY, et al. Impact of BMI on fresh cycle embryo transfer outcome of IVF/ICSI in patients with early follicular phase prolonged protocol. Zhonghua Fu Chan Ke Za Zhi. 2021;56:868–75.
[30] Wastnedge EAN, Reynolds RM, van Boeckel SR, et al. Pregnancy and COVID-19. Physiol Rev. 2021;101:303–18.
[31] Dashraath P, Wong JLJ, Lim MXK, et al. Coronavirus disease 2019 (COVID-19) pandemic and pregnancy. Am J Obstet Gynecol. 2020;222:521–31.
[32] Ntambara J, Chu M. The risk to child nutrition during and after COVID-19 pandemic: what to expect and how to respond. Public Health Nutr. 2021;24:3530–6.