ELSEVIER

Prospective observational cohort study: Computational models for early prediction of ongoing pregnancy in fresh IVF/ICSI-ET protocols

Huiyu Xu[a,b,c,d,1], Yuan Wei[e,1], Rui Yang[a,b,c,d,1], Guoshuang Feng[f,1], Wenhao Tang[g], Hongxia Zhang[a,b,c,d], Yilei He[a,b,c,d], Ying Feng[a,b,c,d], Rong Li[a,b,c,d,*], Jie Qiao[a,b,c,d]

[a] Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing 100191, China
[b] Key Laboratory of Assisted Reproduction (Peking University), Ministry of Education, Beijing 100191, China
[c] National Clinical Research Center for Obstetrics and Gynecology, Beijing 100191, China
[d] Beijing Key Laboratory of Reproductive Endocrinology and Assisted Reproductive Technology, Beijing 100191, China
[e] Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing 100191, China
[f] Center for Clinical Epidemiology & Evidence-based Medicine, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing 100045, China
[g] Department of Urology, Peking University Third Hospital, Beijing 100191, China

ARTICLE INFO

ABSTRACT

Purpose: This study sought to identify the significant factors related to ongoing pregnancy (OP) and to discover the most reliable model to distinguish OP from non-OP in early gestational age.

Methods: A total of 1650 cycles were enrolled in this study. Univariate Logistic Regression was used to identify the predictors included in multivariable analysis. The dataset was then randomly split into training set and test set with proportion of 70% and 30%. Forward stepwise multivariable logistic regression with 5-fold cross validation was used to build the final mathematic model. The performance of the model was determined by the arguments of test set. The area under receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and misclassification rate (MR) were then calculated for model evaluation.

Results: Seven predictors were related to OP by univariate analysis. The serum hCG level on 14th day post-embryo-transfer (hCG_{14}) and 21th day post-embryo-transfer (hCG_{21}) were linear correlated. Therefore, different multivariate regression models were built using hCG_{14} or hCG_{21}, respectively. After multivariate regression with 5-fold validation, the final indicators in model-1 were age_group, hCG_{21} and hCG_{21}/hCG_{14}, while age_group, hCG_{14}, and calculated 48-hour-rising-ratio of hCG were the significant predictors in model-2. Model-2 showed better sensitivity and NPV, lower MR, and similar specificity and PPV.

Conclusion: This study provided an effective mathematic model for early prediction of OP. The model could be of better clinical significance, especially for clinical counseling to manage patients' stress and anxiety, and for early warning of threatened miscarriage.

1. Introduction

The prediction of successful pregnancy for women who undergo in-vitro fertilization/intracytoplasmic sperm injection–embryo transfer (IVF/ICSI-ET) is desirable for both patients and clinicians. An early, accurate prediction of the pregnancy outcome would help reduce patients' level of stress and anxiety. Many serum markers, such as human chorionic gonadotrophin (hCG), estradiol, progesterone (PGN),

activins, inhibins and kisspeptin [1–3] have been investigated regarding their pregnancy outcomes. Among these biomarkers, hCG has been shown to be the best predictive indicator [4].

hCG is a glycoprotein produced by syncytiotrophoblast cells. It can be detected in embryos as early as the 6- to 8-cell stages, and increases rapidly during early pregnancy. Plasma hCG reaches an average peak level of approximately 110,000 IU/L around week 10 of gestation, after which it declines quickly to become approximately 10% of peak levels

[3]. With a similar structure to luteinizing hormone (LH), hCG stimulates corpus luteum to produce PGN before the placenta takes over. PGN is believed to relax utero-tubal musculature and affect tubal motility, which benefits for the transfer of the embryo from fallopian tube to the uterus and the endometrial receptivity [5]. The hCG, itself, was also found to stimulate endometrial angiogenesis and growth [5]. Hence, the proper level of hCG during pregnancy is important for the tubal-uterine transport and uterine implantation of the embryo—and also proper growth and development of embryo—suggesting that the hCG level, as measured by its β-subunit (β-hCG), can be a marker for early prediction of successful pregnancy.

The aim of the present study was to discover a sensitive and reliable computational model for prediction of the probability of OP among women undergoing fresh IVF/ICSI-ET protocols based on serum hCG levels on the 14th and 21st days post-ET (embryo transfer) as well as other basic and clinical characteristics, in the hope of easing the stress and anxiety that patients endure after embryo transfer, and provide early warning of threatened miscarriage induced by insufficient hCG level.

2. Materials and methods

2.1. Data collection

This is a prospective observational cohort study performed from December 2015 to June 2016. Ethical approval was acquired from the institutional review board of Peking University Third Hospital. Datasets of all fresh embryo-transfer cycles were prospectively recorded. Data was entered into a computerized database by clinical support staff. The database collects basic characteristics of patients: basic sex hormone levels, number of antral follicles, the primary, secondary, third, fourth, and the fifth causes of infertility, endometrial thickness on hCG trigger day, details of controlled ovarian stimulation protocol, number of previous IVF/ICSI-ET attempts, number of oocyte retrieved, insemination method, date of insemination, date of embryo transfer (ET), number of ET, the serum concentration of 14th and 21th day post-ET (hCG_{14} and hCG_{21}), specific fertilization results, pregnancy types, including ectopic pregnancy, biochemistry pregnancy, the first-trimester abortion, ongoing pregnancy, the second-trimester abortion, the last-trimester abortion, live birth or not, number of gestational sacs, number of live births, and cause of infant mortality. A total of 6612 fresh ET cycles were carried out. Of these, the transfers which met the inclusion and exclusion criteria were selected. The inclusion criteria were: 1) serum hCG level was > 5 IU/L on hCG_{14} and hCG_{21}, 2) serum hCG levels were tested exactly on day 14th and 21st post-ET, reassured by date of hCG examination and date of ET. The concentration of serum hCG recorded in our IVF-registration system was further verified by double-checking the concentration in our laboratory information system. The specific day post-ET was reassured by checking the date of ET and the date of serum hCG examination. Cycles with hCG levels examined in other labs, with hCG concentration not available in our own laboratory information system, or incomplete information in our IVF/ICSI-ET registration system, or embryos not transferred exactly on day 3th or 5th post-insemination, or hCG were not tested on the exact 14th or 21th day post-ET were further excluded. Of these, 1650 cycles were selected. The cycles were further divided into two groups according to OP. The flowchart of this process is shown in Fig. 1.

2.2. Definition of cause of infertility in our study

The causes of infertility in our IVF/ICSI-ET registration system are different. We selected patients with the most common cause of infertility from our registration system in this study. Male infertility included primary and secondary male infertility, and was defined according to the WHO manual for the standardized investigation and diagnosis of the infertile couples [6]. Endometriosis was defined as the presence of endometrial glands and stroma in elsewhere instead of the uterine cavity, with a combination of dysmenorrhea and dyspareunia. Oligo-ovulation or annovulatory infertility corresponded to patients with no increase of basal body temperature (BBT) by tracking BBT each day, together with no dominant follicle-development showed by transvaginal sonography (TVS) monitoring during menstrual cycles, and no increase of serum progesterone by monitoring serum level of progesterone in late menstrual cycle. The above manifestations lasted more than three menstrual cycles within one year. Tubal-factor infertility was diagnosed by laparoscopic evaluation when the fallopian-tube infertility was suggested by salpingography. Unexplained infertility referred to the couples in whom all the standard investigations-such as tests of ovulation, tests of tubal patency and semen analysis-were normal, with a repeatedly failed history of guided sexual intercourse or intra-uterine insemination.

2.3. IVF/ICSI-ET protocols

The controlled ovarian stimulation protocols used in our center included GnRH agonist long protocol, GnRH agonist short protocol, GnRH agonist ultra-long protocol and GnRH antagonist protocols which were described previously [7]. When two or more leading follicles reached a diameter of 18 mm, 5000–10,000 IU human hCG was administered. Transvaginal ovum collection was performed 36–38 h later. The collected oocytes were fertilized by in vitro fertilization or intracytoplasmic sperm injection (ICSI). After 3–5 days of culture, one or two embryos were freshly transferred. PGN supplementation was provided for luteal support until 8–10 weeks of pregnancy in natural cycles and COS cycles or 10–12 weeks of artificial cycles respectively.

2.4. Pregnancy outcome

According to our routine clinical practice, every patient who went through IVF/ICSI-ET was asked to take a serial serum hCG assay at 14th and 21th day post-ET to preliminarily assess pregnant status, and to take transvaginal sonography 30 or 37 days post-ET to conform the position of the gestational sac. The follow-up was ended three months after childbirth. Different types of pregnancy outcomes were recorded on our Assisted Reproductive Technology (ART) registration system. Ectopic gestation was defined by the gestation sac outside the uterus, which was detected by TVS or laparoscopy. Biochemical pregnancy meant a temporary rise of serum hCG without the gestational sac inside or outside the uterus under ultrasound examination on 30 or 37 days post-ET. OP was defined as intrauterine pregnancies lasting > 12 weeks of gestation. In conclusion, the Non-OP group we established here includes biochemical pregnancy, ectopic pregnancy and early miscarriage of intrauterine pregnancy within 12 weeks of gestation. While the OP group contains all the pregnancies lasting > 12 gestational weeks, including live-births, dead-births and mid-trimester or last trimester miscarriage.

2.5. hCG assays

On 14th and 21th day post-ET, the serum hCG level of each patient was assessed by an Access UniCel DxI 800 chemiluminescence system (Beckman Coulter, Inc., Brea, U.S.A.) using the Access Total β-hCG (5th IS) assay, which is standardized to the highly purified World Health Organization (WHO) 5th International Standard (IS) for chorionic gonadotropin. The inter-assay variation was 7.9% in low-level of Bio-Rad immunoassay plus control, 7.37% in mid-level control and 4.09% in high-level control. Testing was performed from December 2015 to June 2016.

2.6. Statistical analysis

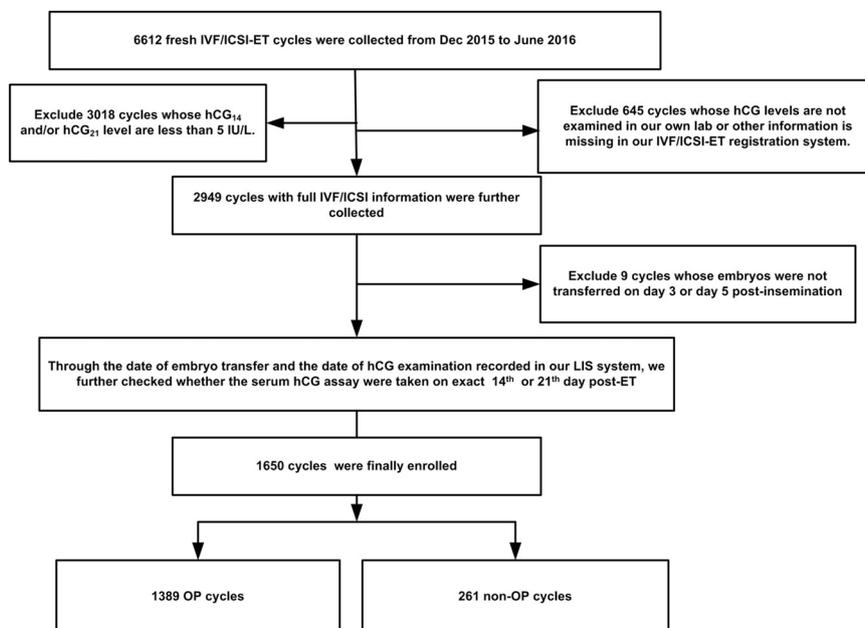The 48-hour-rising-ratio of hCG was calculated as: (β-hCG on day 21

Fig. 1. Flowchart of the data selection strategy. hCG_{14} and hCG_{21} denote the serum hGC level on 14th and 21th day post-ET. LIS is short for laboratory information system.

/ β-hCG on day 14) $^{1/3.5}$. For the statistical description, we use percentages for categorical variables and mean ± standard deviation or median (interquartile range) for normally or non-normally distributed continuous variables. Categorical data were compared using the chi-squared test, and continuous data using the independent samples *t*-test or Wilcoxon rank sum test where appropriate.

The establishment step of prediction model was as follows: first, the dataset was randomly partitioned into a training set and a test set by the proportion of 0.7:0.3; second, the forward selection logistic regression with 5-fold cross-validation was used to construct a prediction model for OP in the training set. The criteria for variables selection was the scaled negative log-likelihood, which considered the one with smallest scaled negative log-likelihood as the best. Then the model was applied to the test data, and the sensitivity, specificity, misclassification rate (MR), PPV, NPV, and AUC were calculated for model evaluation. The data was analyzed with JMP PRO 14.0 software from SAS Institute Inc., and a two-sided *p* value of < 0.05 was considered statistically significant.

3. Results

3.1. hCG levels and ratios in different pregnancy outcomes

During the study period, a total of 6612 fresh IVF/ICSI-ET cycles were enrolled in our study. Of these, 1650 cycles which met the inclusion and exclusion criteria were finally selected. The basic and clinical characteristics in relation to OP were shown in Table 1. The distribution of hCG levels and ratios among different pregnancy outcome groups were shown in Table 2. The median and 95% CI (confidence interval) of hCG_{14}, the serum hCG level of 14th day post-ET, were 884 (222–3014) mIU/mL in OP, and 331 (28–1886) mIU/mL in non-OP group. The median and 95% CI of hCG_{21} levels were 16,311 (4522–45,203) mIU/mL in OP group, and 3667 (11–20,608) in non-OP group (*p* < 0.0001). The median and 95% CI level of calculated 48 hour ratios of hCG rising were 2.28 (1.88–2.73) and 1.97 (0.5–2.95) in OP and non-OP groups, respectively.

Table 1
Clinical and biological data in relation to OP in fresh IVF/ICSI-ET cycles.

	OP	
	Yes (*n* = 1389)	No (*n* = 261)
Age (years)	31.9 ± 4.0	33.9 ± 4.7
BMI	22.1(20.3–24.5)	22.2 (20.2–24.7)
Main cause of infertility, n (%)		
Male infertility	442 (85.8%)	73 (14.2%)
Endometriosis	54 (75%)	18 (25%)
Annovulatory infertility	171 (80.7%)	41 (19.3%)
Tubal factor	458 (83.2%)	92(16.8%)
Unexplained and others	264 (87.7%)	37(12.3%)
Endometrial thickness on hCG day (mm)	11.0 ± 1.6	10.7 ± 1.6
hCG_{14}*	884 (576–1342)	331 (155–563)
hCG_{21}*	16,311 (11206–23,907)	3667 (587–8278)
48 h rising ratio of hCG	2.3 ± 0.2	1.8 ± 0.6
ET on 3th or 5th day post insemination, n (%)		
Day 3	1366 (84.6%)	249 (15.4%)
Day 5	23 (65.7%)	12 (34.3%)
Number of embryo transferred, n (%)		
1	79 (71.2%)	32 (28.8%)
2	1310 (85.1%)	229 (14.9%)

ET, embryo transfer; * hCG_{14}, serum level of hCG on 14th day post-ET; * hCG_{21}, serum level of hCG on 21th day post-ET.

3.2. Univariate analysis to identify predictors related to OP

Univariate analysis was performed to evaluate each variable's effect on OP. The continuous variable of age was first converted into categorical variable of age_group at age cut-off value of 35 years old. Other predictors included in our analysis were BMI, main cause of infertility, endometrial thickness on hCG trigger day, hCG_{14}, hCG_{21}, 48-hour-rising-ratio of hCG, number of ET, ET on 3 or 5 days post-insemination, serum basal follicular stimulating hormones (FSH) level, serum basal LH level, serum basal estradiol (E2) level and antral follicle counts (AFC). Continuous data were subjected to univariate logistic regression, while the categorical data were subjected to chi-square test. The results were shown in Table 3. Seven predictors were significantly related to

Table 2
The distribution of hCG levels and ratios among different pregnancy outcome groups.

Variable	Outcome	n	Percentile						
			0.5th	2.5th	25th	50th	75th	97.5th	99.5th
hCG on day 14th post-ET (mIU/mL)	OP	1389	119	222	576	884	1342	3014	3757
	non-OP	261	8	28	155	331	563	1886	2164
hCG on day 21th post-ET (mIU/mL)	OP	1389	2421	4522	11,206	16,311	23,907	45,203	55,017
	non-OP	261	6	11	587	3667	8278	20,608	30,182
48-Hour-rising-ratio of hCG	OP	1389	1.72	1.88	2.14	2.28	2.43	2.73	2.88
	non-OP	261	0.36	0.50	1.49	1.97	2.21	2.95	3.41

OP, including age_group, endometrial thickness on hCG trigger day, hCG_{14}, hCG_{21}, hCG_{21}/hCG_{14} or calculated 48-hour-rising-ratio of hCG rising, number of ET, and ET on 3 or 5 days post-insemination, with all p-values < 0.05. To be specific, if the value of parameter estimation is positive, the probability of OP increases with the increase of the predictor, and if the value of parameter estimation is negative, the probability of OP increases with the decrease of the predictor. According to the results from Table 3, the probability of OP in blastocyst transfer (ET on 5 days post insemination) cycles was lower than in cleavage transfer (ET on 3 days post insemination) cycles. The chance of OP was lower in the women aged ≥ 35 years old group than < 35 years old group. The probability of OP increases with the increasing of endometrial thickness, hCG_{14}, hCG_{21} and 48-hour-rising-ratio of hCG. Furthermore, patients with two embryos transferred had a better chance to have OP compared to those who had one embryo transferred.

3.3. Multivariable regression model for predicting OP

The hCG_{14} and hCG_{21} were linear correlated, with correlation coefficient of 0.858, which indicated that these two indicators couldn't be placed in one model simultaneously. Therefore, we built two separate multivariate regression models using hCG_{14} and hCG_{21}, respectively.

Firstly, the dataset was randomly split into a training set and a validation set with proportion of 70% and 30%. Secondly, in the training set, to find the ideal model, the forward stepwise logistic regression added variables to the regression model one by one according to their significance, followed by a 5-fold cross-validation to find the appropriate stopping criteria, which controlled model complexity. The scaled negative log-likelihood was used to evaluate the appropriate number of predicting indicators included for the model. The lower the scaled negative log-likelihood is, the better the model will be. The green

zone in Fig. 2 identified the model which is comparable to the best model. The included indicators in model-1 were age_group, hCG_{21} and hCG_{21}/hCG_{14}. While age_group, hCG_{14}, and calculated 48-hour-rising-ratio of hCG were the statistical significant predictors in model-2. The variable selection process of the two models was shown in Fig. 2. Finally, the two models were verified in their corresponding validation set.

Both model-1 and model-2 showed good discriminatory performances. The AUC, sensitivity, specificity, PPV, NPV, and MR of model-1 and model-2 were indicated in Table 4. As we know, we built the mathematic model in the training set, and verify its efficacy in the validation set. Therefore, the performance of each model was determined by the arguments of validation set. The AUC and MR in model-1 were higher than model-2. While the sensitivity and the NPV in model-2 were higher than in model-1, with the specificity and PPV in the model were almost the same with model-1. In addition, model 2 can predict the results earlier using hCG_{14} rather than hCG_{21}. Therefore, model-2 may be more ideal to help manage patient anxiety.

3.4. Grouping the population according to the distribution between predicted probability of OP and incidence of OP

We attempted to divide the cohort into more groups based on the models' predicted probability, thus avert from misclassification of OP or non-OP predicted by probability of 50%, which is a default setting of logistic regression from any statistic software. To be specific, the predicted probability of $> 50\%$ would be grouped into OP by default settings in the software, otherwise it is non-OP. To explore the relationship between the predicted probability of OP and the incidence of OP, we made a histogram between them. The predicted probability was further classified into 20 groups by every 5 percentage (Fig. 3). From this figure, we can see that the incidence of OP was significantly different in
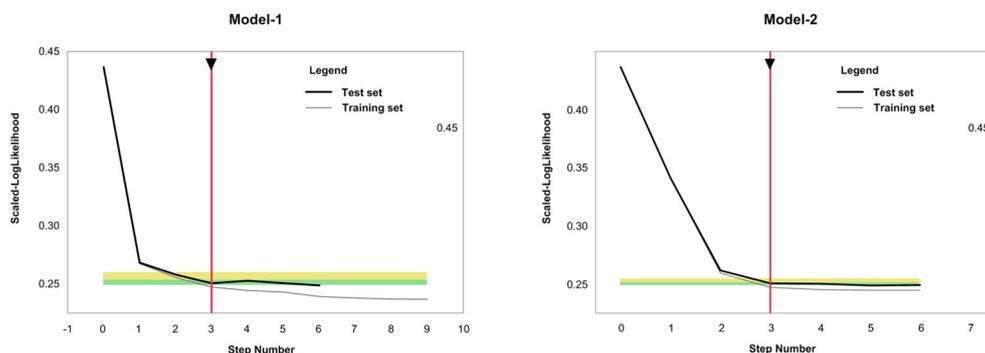
Table 3
Univariate analysis of the effect of each predicting variable on OP.

		Parameter estimation	Standard error	Wald χ^2	P value
Age	≥ 35 vs. < 35	-0.9194	0.14	44.07	< 0.0001
BMI		-0.005	0.02	0.08	0.779
Endometrial thickness on hCG day		0.1266	0.04	8.65	0.003
ET on day 3 or day 5 post insemination, n (%)	5 vs. 3	-1.0516	0.36	8.41	0.004
Number of embryo transferred, n (%)	2 vs. 1	0.8404	0.22	14.40	< 0.001
Main cause of infertility, n (%)	1 vs. 4[a]	0.1958	0.17	1.32	0.250
	2 vs. 4[a]	-0.5065	0.30	2.94	0.086
	3 vs. 4[a]	-0.1770	0.21	0.72	0.395
	5 vs. 4[a]	0.360	0.21	2.95	0.086
Serum bFSH level (IU/L)		-0.0298	0.02	1.55	0.214
Serum bLH level (IU/L)		0.0943	0.06	2.86	0.091
Serum E_2 level (IU/L)		0.0021	0.00	1.49	0.223
Antral follicle count (AFC)		-0.0241	0.02	1.20	0.274
hCG_{14}		0.0030	0.00	151.13	< 0.001
hCG_{21}		0.0003	0.00	262.96	< 0.001
48 h rising ratio		3.7807	0.28	178.04	< 0.001

[a] 1 denotes male infertility, 2 denotes endometriosis, 3 denotes annovulatory infertility, 4 is short for tubal factor infertility, 5 denotes is short for unexplained infertility and others.

Fig. 2. The relationship between scaled negative log-likelihood and the step number (number of variables) in training sets and test sets of model-1 and model-2. The test set reflects the effect of the model built in training set. The green zone identifies the model which is comparable to the best model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
The performance of model-1 and model-2.

	Model-1		Model-2	
	hCG$_{21}$ + 48 h rising ratio		hCG$_{14}$ + 48 h rising ratio	
	Training set	Test set	Training set	Test set
AUC	0.899	0.923	0.885	0.903
Sensitivity	0.972	0.982	0.980	0.993
Specificity	0.593	0.539	0.531	0.539
PPV	0.927	0.919	0.917	0.920
NPV	0.800	0.849	0.835	0.933
MR	0.088	0.088	0.091	0.079

AUC, area under ROC curve; PPV, positive predictive value; NPV, negative predictive value; MR, Misclassification rate;

most of the 20 groups. Dividing patients into more groups can help minimize the bias generated by the default setting of using probability of 50%.

As shown in Fig. 3, the incidence of OP increased with the predicted probability of OP. When predicted probability was < 20%, no OP was discovered. The incidence of OP continuously increased with the increase in predicted probability from the range between 20% and 65% to the range between 65% and 85%, and then the range of between 85% and 95%. The incidence of OP was the highest when predicted probability was > 95%.

4. Discussion

In this analysis of 1650 fresh IVF/ICSI-ET cycles, we constructed mathematic models using hCG$_{14}$ or hCG$_{21}$ combined with hCG ratio and age_group for prediction of OP. hCG$_{14}$ model (Model-2) showed better sensitivity and NPV, lower MR, and similar specificity and PPV, with AUC, sensitivity, specificity, PPV, NPV, and MR of 0.903, 0.993, 0.539, 0.920, 0.933, and 7.9%, respectively. We further grouped the cohort
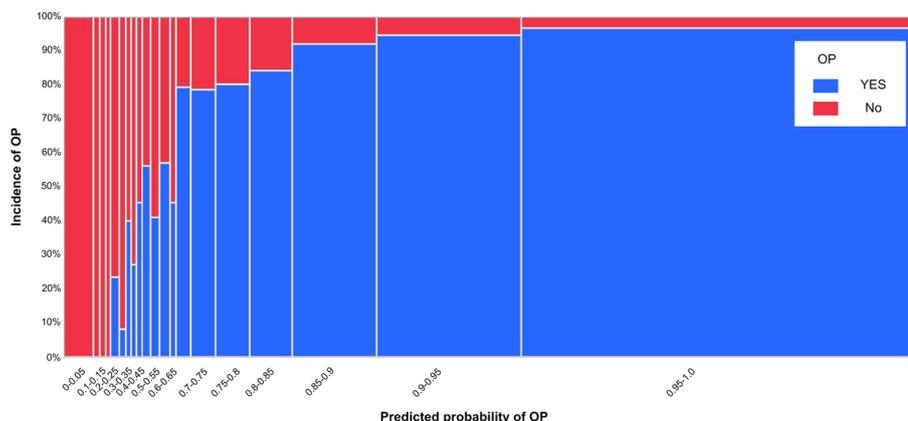
into sub-groups according to the distribution of predicted probability of OP and incidence of OP, thus averted from misclassification of OP or non-OP predicted by probability of 50%. The mathematical model in this study could be useful for clinical counseling to manage patients' stress and anxiety, and may be beneficial for early warning of threatened miscarriage.

The level of hCG is a key factor for successful embryo implantation in the first trimester of pregnancy. To date, there are a few studies using serum hCG levels to predict OP or viable pregnancy. For instance, Bjercke et al. [8] used hCG level on 12 days post-ET (hCG$_{12}$) to predict non-OP (early pregnancy loss, including biochemical pregnancy, ectopic pregnancy and first-trimester abortion) using 417 IVF cycles. The NPV at their cut-off value was 90%, which corresponds to our PPV of 92%. Furthermore, their PPV at their cut-off value was 60%, which corresponds to our NPV of 93.3%. However, Bjercke's study did not show the sensitivity, specificity or MR of their model. Poikkeus's [3] also used hCG$_{12}$ levels to predict viable pregnancies of > 22 weeks. The sensitivity at their cut-off values was 80% with the specificity at 82%. The PPV and NPV were 87% and 74%, respectively. Their study didn't show the AUC and MR. Our AUC, PPV, NPV, sensitivity, specificity and MR were shown in Table 4. We also reviewed other literatures, including the latest literatures, and discovered that the performance of our model is aligned with other studies.

We have actually predicted different pregnancy outcomes: 1) predicting live births, 2) predicting ongoing pregnancies and 3) predicting early intrauterine pregnancies/ectopic pregnancies/biochemical pregnancies. The predictive effect is getting better and better. I think this may be related to the function of hCG itself. hCG plays an important role in the maintenance of early embryo growth. It continues to rise until 12 weeks of gestation and then decreases. This may be the reason why the effect of using early hCG level to predict pregnancy outcomes of > 12 gestational weeks was not satisfactory. In this study, we focused on ongoing pregnancies of lasting > 12 weeks of gestation, which include live-births, dead-births and mid-trimester or last-trimester miscarriage.

Fig. 3. Distribution between incidence of OP and predicted OP probability.
The ordinate signifies the incidence of OP, and the abscissa indicates the predicted probability of OP, the wider the abscissa, the larger the sample size in each sub-group. Blue means the OP group, red signifies the non-OP group. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In addition, in our prediction model, we use the hCG level of the 14th and 21th day after embryo transfer. This is also our clinical practice, which established based on many years of exploring and summary. We find that the prediction of some pregnancies (such as ectopic pregnancy) may not be well predicted if the hCG examination is too early or the detection interval is too short. Of course, in clinical practice, due to poor patient compliance, the date of hCG examination is often not accurate the 14th and 21th day after embryo transfer. We will then optimize this model, so that other day of hCG examination, for example 12th and 20th day after embryo transfer, can also be predicted by our model.

The limitation of our model is its low specificity (53.9%), that is, a great proportion of the non-OP is not predicted. The primary reason of the low specificity may be that we only include people with hCG_{14} and $hCG_{21} > 5\,IU/mL$, among whom a considerable part of non-OP was first-trimester abortion or ectopic pregnancy, accounting for 77.8% of all non-OP. We further discovered that, in the group of predicted OP probability of > 60%, the non-OP in this cohort were mainly first-trimester abortion, which happened at least 30 days after embryo transfer, indicating that a great proportion of their embryo implantations were relatively normal at 21th day post-ET. The failure of pregnancy may result from the natural selection of embryos for not passing the checkpoint of embryonic development in the first trimester.

It is accepted that the initial evaluation of a woman with early pregnancies of unknown location includes measurement of the serum hCG concentration. It is commonly known that serum hCG doubles every 2 days. Deviations from this doubling pattern were highly susceptive of a first-trimester abortion, or an ectopic pregnancy, or biochemical pregnancy, all of which constitute the non-OP population. Recently, it is recognized that the hCG ratio is much more complicated than it first appeared. We still didn't draw the conclusion of how to discriminate OP from non-OP cohort. The underlying mechanism may lie in the heterogeneity in the following aspects: 1) the enrolled population, 2) the included predicting indicators; 3) the definition of pregnancy outcome; and 4) the statistical analysis. The logistic regression models have better diagnostic performance in predicting pregnancy outcomes [9]. Our regression model, using serum hCG_{14}, 48-hour-rising-ratio of hCG and age performed well in predicting OP. Furthermore, our study prospectively included all the fresh IVF/ICSI-ET cycle with complete and accurate basic and clinical characteristics, which indicates a good representativeness.

Although there are reports using anti-Mullerian hormone (AMH), FSH, E_2, and AFC to predict pregnancy outcomes, [10–13] these factors are strongly related to age [14,15], which were also confirmed in our data. These associations, on one hand, can hinder the accuracy of regression model; on the other hand, it can also increase the complexity of the model. Furthermore, we previously discovered that AMH is not related to clinical pregnancy [16]. Furthermore, AMH is more generally accepted to be a predictor of the number rather than the quality of the follicles. Thus AMH was not included in the initial design of this prospective study. Moreover, the inclusion of AFC and other basic level of sex hormones, including FSH, LH, and E2 did not contribute to the prediction of OP, which meant that AMH might not be a good predictor of OP either.

It should be noted that the ET on three days (cleavage) or five days (blastocyst) post-insemination is of significant clinical value for predicting OP. However, these two variables were not included in the final multivariate regression model. The underlying mechanism may be as follow: 1) It is acknowledged that the serum hCG level of cleavage transfer should be lower than blastocyst transfer. The earlier the gestational day, the more significant the difference will be. The differences of hCG level induced between the two days may be largely weakened by 14th or 21th post-ET; 2) In fresh IVF/ICSI-ET cycles, blastocyst transfer is not a routine approach. The proportion of ET on day 5 post insemination is too small (35 out of 1650 cycles) in our fresh IVF/ICSI-ET data, it may not reflect the actual role of cleavage or blastocyst transfer on OP.

Clinical pregnancy rate, live-birth rate and OP rate increase significantly with increased endometrial thickness [12,17–19]. Dickey et al. reported that endometrial thickness was positively correlated to the number of follicles > 15 mm in diameter and pregnancy outcome [17]. Others also found that the pregnancy rate increased with increased thickness of the endometrium [12,18,19]. We discovered that the endometrial thickness had statistical significance in univariate regression analysis but not in the variable selection process, which means that the contribution of endometrial thickness on the prediction of OP is limited.

Physicians should be aware that the studies previously reported were performed using different hCG assays, and the reference intervals and cut-offs generated from these studies varied based on the hCG assay used. The hCG assays are standardized to different international standards (I·S), which contain different combinations of β-hCG variants. While many studies used assays standardized to the Second I·S of β-hCG, some of more recent studies have used the Third IS, with only a few studies using the Fourth I.S. The Fifth I.S, which removed free and nicked variants of β-HCG, can achieve better harmonization among different β-hCG assays, thus more beneficial for the standardization of β-hCG. As far as we are aware, the Access Total β-hCG (5th I·S) assay we used in this study is the first and only chemiluminescent immunoassay that moved to the new 5th I.S. Given the fact that there is no longer any storage of the 3rd or 4th IS materials available, we would expect other manufacturers will eventually move to the 5th IS. Our study is the first one using the 5th IS hCG assay combined with other prognostic factors to predict OP, which provides reliable and important medical decision levels for physicians.

The drawback of this study is its design of non-randomized control trial study. However, as a single-center cohort study, there is no strict limit on the selection of subjects, so it averts from selection bias. Therefore, our result is more aligned with clinical practice. We further grouped the cohort into more sub-groups according to the distribution between predicted probability of OP and incidence of OP, therefore avert from misclassification of OP or non-OP predicted by probability of 50%, and thus, our model can be of better clinical significance, especially for clinical counseling to manage patients' stress and anxiety, and for early warning of threatened miscarriage induced by insufficient hCG level.

Author's roles

HYX, YW, RY participated in the design, data collection, and manuscript writing. GSF participated in statistical analysis and a part of manuscript writing. WHT, HXZ, YLH, and YF participated in data collection and verification. JQ conceived the study and participated in its design. RL conceived the study, edited this manuscript, provided financial support and finally approved the final submission. All authors read and approved the final manuscript.

Funding

Conflict of interests

None declared.

References

[1] Z. Jamil, S.S. Fatima, S. Arif, F. Alam, R. Rehman, Kisspeptin and embryo implantation after ICSI, Reprod. BioMed. Online 34 (2016) 147–153.

[2] J. Johns, S. Muttukrishna, M. Lygnos, N. Groome, E. Jauniaux, Maternal serum hormone concentrations for prediction of adverse outcome in threatened miscarriage, Reprod. BioMed. Online 15 (2007) 413–421.

[3] P. Poikkeus, V. Hiilesmaa, A. Tiitinen, Serum HCG 12 days after embryo transfer in predicting pregnancy outcome, Hum. Reprod. 17 (2002) 1901–1905.

[4] C.C. Lawler, N.M. Budrys, A.K. Rodgers, A. Holden, R.G. Brzyski, R.S. Schenken, Serum beta human chorionic gonadotropin levels can inform outcome counseling after in vitro fertilization, Fertil. Steril. 96 (2011) 505–507.

[5] Filicori M, Fazleabas AT, Huhtaniemi I, Licht P, Rao Ch V, Tesarik J, et al. Novel concepts of human chorionic gonadotropin: reproductive system interactions and potential in the management of infertility. Fertil. Steril. 2005;84:275–84.

[6] F.C. PJ Rowe, WHO Manual for the Standardized Investigation and Diagnosis of the Infertile Male. Book, (2000), p. 0521774748.

[7] H. Chi, J. Qiao, H. Li, P. Liu, C. Ma, Double measurements of serum HCG concentration and its ratio may predict IVF outcome, Reprod. BioMed. Online 20 (2010) 504–509.

[8] S. Bjercke, T. Tanbo, P.O. Dale, L. Morkrid, T. Abyholm, Human chorionic gonadotrophin concentrations in early pregnancy after in-vitro fertilization, Hum. Reprod. 14 (1999) 1642–1646.

[9] van Mello NM, Mol F, Opmeer BC, Ankum WM, Barnhart K, Coomarasamy A, et al. Diagnostic value of serum hCG on the outcome of pregnancy of unknown location: a systematic review and meta-analysis. Hum. Reprod. Update. 2012;18:603–17.

[10] R. Tal, O. Tal, B.J. Seifer, D.B. Seifer, Antimullerian hormone as predictor of implantation and clinical pregnancy after assisted conception: a systematic review and meta-analysis, Fertil. Steril. 103 (2015) 119–130 (e3).

[11] P.B. Maseelall, A.E. Hernandez-Rey, C. Oh, T. Maagdenberg, D.H. McCulloh, P.G. McGovern, Antral follicle count is a significant predictor of livebirth in in vitro fertilization cycles, Fertil. Steril. 91 (2009) 1595–1597.

[12] N. Noyes, B.S. Hampton, A. Berkeley, F. Licciardi, J. Grifo, L. Krey, Factors useful in predicting the success of oocyte donation: a 3-year retrospective analysis, Fertil. Steril. 76 (2001) 92–97.

[13] M.I. Shamonki, J.L. Frattarelli, P.A. Bergh, R.T. Scott, Logarithmic curves depicting initial level and rise of serum beta human chorionic gonadotropin and live delivery outcomes with in vitro fertilization: an analysis of 6021 pregnancies, Fertil. Steril. 91 (2009) 1760–1764.

[14] Gynecologists TACoOa, Committee opinion no. 618: ovarian reserve testing, Obstet. Gynecol. 125 (2015) 268–273.

[15] van Rooij IA, Broekmans FJ, Scheffer GJ, Looman CW, Habbema JD, de Jong FH, et al. Serum antimullerian hormone levels best reflect the reproductive decline with age in normal women with proven fertility: a longitudinal study. Fertil. Steril.. 2005;83:979–87.

[16] H. Xu, L. Zeng, R. Yang, Y. Feng, R. Li, J. Qiao, Retrospective cohort study: AMH is the best ovarian reserve markers in predicting ovarian response but has unfavorable value in predicting clinical pregnancy in GnRH antagonist protocol, Arch. Gynecol. Obstet. 295 (2017) 763–770.

[17] R.P. Dickey, T.T. Olar, D.N. Curole, S.N. Taylor, P.H. Rye, Endometrial pattern and thickness associated with pregnancy outcome after assisted reproduction technologies, Hum. Reprod. 7 (1992) 418–421.

[18] P. Kovacs, S. Matyas, K. Boda, S.G. Kaali, The effect of endometrial thickness on IVF/ICSI outcome, Hum. Reprod. 18 (2003) 2337–2341.

[19] K.S. Richter, K.R. Bugge, J.G. Bromer, M.J. Levy, Relationship between endometrial thickness and embryo implantation, based on 1,294 cycles of in vitro fertilization with transfer of two blastocyst-stage embryos, Fertil. Steril. 87 (2007) 53–59.