



Original article

Predictive Model for Live Birth at 12 Months After Starting In-Vitro Fertilization Treatment

Ho NA. Vu^a, Pham D. Toan^a, Ho M. Tuong^a, Vuong N. Lan^{ab*}

^aIVFMD, My Duc Hospital, 4 Nui Thanh Street, Tan Binh District, Ho Chi Minh City, Vietnam;

^bDepartment of Obstetrics and Gynaecology, University of Medicine and Pharmacy HCMC, 217 Hong Bang Street, District 5, Ho Chi Minh City, Vietnam.

Received December 31, 2017; Accepted March 31, 2018; Published online July 01, 2018

Abstract: IVF carries a considerable physical, emotional and financial burden. Therefore, it would be useful to be able to predict the likelihood of success for each couple. The aim of this retrospective cohort study was to develop a prediction model to estimate the probability of a live birth at 12 months after one completed IVF cycle (all fresh and frozen embryo transfers from the same oocyte retrieval). We analyzed data collected from 2600 women undergoing in vitro fertilization (IVF)/intracytoplasmic sperm injection (ICSI) at a single center in Vietnam between April 2014 and December 2015. All patients received gonadotropin-releasing hormone (GnRH) antagonist stimulation, followed by fresh and/or frozen embryo transfer (FET) on Day 3. Using Cox regression analysis, five predictive factors were identified: female age, total dose of recombinant follicle stimulating hormone used, type of trigger, fresh or FET during the first transfer, and number of subsequent FET after the first transfer. The area under the receiver operating characteristics curve for the final model was 0.63 (95% confidence interval [CI] 0.60–0.65) and 0.60 (95% CI 0.57–0.63) for the validation cohort. There was no significant difference between the predicted and observed probabilities of live birth (Hosmer-Lemeshow test, $p > 0.05$). The model developed had similar discrimination to existing models and could be implemented in clinical practice.

Keywords: predictive model, live birth, GnRH antagonist, embryo transfer.

1. INTRODUCTION

Significant advances in the field of assisted reproductive technology (ART) have helped numerous infertile couples fulfil their desire for parenthood. In its most recent report, the International Committee Monitoring Assisted Reproductive Technologies (ICMART) estimated that around 1.9–2.2 million ART cycles were being performed globally each year, with 480,000 babies delivered annually [1]. Between the birth of the first test-tube baby, Louise Brown, in 1978 and 2012, the total number of babies born via ART has reached 6.5 million worldwide. Despite these impressive numbers, not every couple who present for fertility treatment will be successful. On average, only half of all women starting IVF will achieve a successful live birth, with the rest remaining childless despite multiple IVF cycles [2].

In view of the considerable physical, emotional and financial burden associated with IVF, it is important to discuss the probability of success with each couple prior to their decision to commit to IVF. Common prognostic factors known to influence the live birth rate include the woman's age, cause and duration of infertility, ovarian reserve markers such as antral follicle count (AFC), anti-Mullerian hormone (AMH) levels, the number of oocytes retrieved, and embryo-related factors [3-5]. To aid physicians and patients in their decision-making process, one or more of these prognostic factors can be statistically modelled to predict IVF outcome. The power of prediction increases when several prognostic factors are used together [6, 7].

Most existing models were developed to predict pregnancy or ongoing pregnancy as the primary outcome [8,

*Address correspondence to Vuong Ngoc Lan. Department of Obstetrics and Gynaecology, University of Medicine and Pharmacy HCMC, 217 Hong Bang Street, District 5, Ho Chi Minh City, Vietnam.

Email: lanvuong@ump.edu.vn

DOI: 10.32895/UMP.MPR.2.2.5

9]. Very few have attempted to predict live birth following IVF [10-13] and only two, the Templeton [13] and Nelson [11] models, have been externally validated. However, both these models are of limited use in Asian countries because their datasets were derived from populations predominantly consisting of Caucasian women. Current evidence suggests that Asian women experience significantly lower clinical pregnancy and live birth rates with IVF compared with Caucasian women, probably due to physiologic and socioeconomic differences [14, 15]. In addition, the Templeton model pre-dated the use of ICSI whilst the model by Nelson involved fresh cycles only.

With the gradual shift from use of long gonadotropin-releasing hormone (GnRH) agonist protocols to a GnRH antagonist protocol as the standard regimen [16], plus the widespread use of ICSI and embryo vitrification in modern ART treatment [17], there is a need for newer and more accurate models to facilitate patient selection, guide clinical decisions and improve patient counselling. In this study, we aimed to develop a reliable prediction model to estimate the probability of a live birth at 12 months in patients who had completed one IVF/ICSI cycle (including all fresh and frozen embryo transfers from the same oocyte retrieval) in a controlled ovarian cycle using GnRH antagonist stimulation protocol.

2. MATERIALS AND METHODS

2.1. Development cohort

In this retrospective cohort study, we retrieved data for women who had undergone ART treatment at IVFMD, My Duc General Hospital, Ho Chi Minh City, Vietnam, between April 2014 and December 2015. Women included in the development cohort fulfilled the following criteria: age 18–45 years, embryo transfer(s) within 12 months since the start of the IVF cycle, and a maximum of 2 embryos for transfer on Day 3. Women who had IVF with donor eggs, in-vitro maturation of oocytes, cancelled cycle due to poor ovarian response or no oocyte pickup (OPU), no embryo available for transfer and uterine abnormalities (i.e. adhesion of the uterine cavity, bicornuate uterus or didelphis uterus) were excluded.

2.2. Stimulation protocol

All subjects received recombinant follicle-stimulating hormone (rFSH) stimulation in a GnRH antagonist protocol. rFSH was administered from Day 2 of the menstrual cycle at an initial dose of 150 IU, 225 IU or 300 IU per day based on the patient characteristics, including age, body mass index (BMI), AMH level and AFC. Stimulation was continued until the day of human chorionic gonadotrophin (hCG) administration; daily rFSH dose was titrated at the discretion of the attending physician. Ovarian response was monitored by serial ultrasound scan and hormone levels (serum estradiol and progesterone). hCG (250 µg/0.5 mL, Ovitrelle®; Merck Serono, Germany) was administered when at least 2 leading follicles had reached a diameter of 17 mm on ultrasound.

GnRH agonist (triptorelin 0.2 mg subcutaneous injection, Decapeptyl®; Ipsen Beaufour, France) was used for triggering oocyte maturation when there were more than 15 developing follicles on the day of trigger. Oocyte retrieval was performed via transvaginal aspiration 36 hours after triggering.

2.2.1. Embryo culture and scoring

ICSI was performed for all cycles. Fertilization was checked 16–18 hours post-insemination. Embryos were rated according to the Istanbul Consensus criteria, where “good” was defined as having Grade 1 morphology (i.e. presence of 8 cells, even cell size, less than 10% fragmentation and no multinucleation) [18]. Embryo transfer occurred on Day 3 after oocyte retrieval. The choice of single or double fresh embryo transfer was based on the number of good-quality embryos available. Additional embryos were cryopreserved for later use. Indications for a freeze-all strategy were GnRH agonist trigger, unfavorable endometrium, fluid accumulation in the endometrial cavity, hydrosalpinx not removed before IVF treatment, risk of ovarian hyperstimulation or patient preference. Patients who failed the first fresh or frozen transfer were given the option to return for subsequent FET.

2.2.2. Luteal phase support

Patients who had hCG trigger cycle were given daily vaginal pessary progesterone 400 mg (Cyclogest®; Actavis, UK) for 16 days starting from the day of OPU. For patients who underwent GnRH agonist trigger, luteal support consisted of a combination of intramuscular progesterone 50 mg per day (Rotex Medica, Germany), vaginal suppository progesterone 400 mg per day and oral estradiol valerate 6 mg per day (Progynova®; Bayer Schering, Germany) from the day of OPU. Pregnancy testing was performed 14 days after embryo transfer. A serum beta-hCG of more than 5 mIU/mL was considered positive for pregnancy. Clinical pregnancy was confirmed via transvaginal ultrasonography at 7 weeks of gestation.

2.2.3. Reproductive outcome

The primary outcome for the model was live birth, defined as the birth of a newborn after 24 weeks’ gestation that exhibited any sign of life, such as respiration, heartbeat, umbilical pulsation or movement of voluntary muscles. The birth of twins was considered as a single live birth event.

2.3. Model performance

We assessed the predictive ability of the three models based on two performance measures, i.e. discrimination and calibration. The ability to distinguish patients who will achieve live birth from those without live birth at 12 months after starting IVF was tested using receiver operating characteristics (ROC) analysis. For calibration, we compared the predictive values of the model with real observed live birth rates. To provide clinicians with a user-friendly interface for the final model, we constructed a nomogram that could be used as a graphical tool for conveying the probability of

live birth based on an individual patient profile.

2.4. Statistical analyses

Data were analysed using the R Statistical Environment version 3.3.3 on Windows platform. Firstly, descriptive analysis was applied to the development cohort, and then a bi-variable followed by multi-variable Cox proportional hazards regression model was used to determine factors that may affect the probability live birth within 12 months of starting IVF treatment. We used the following routinely recorded variables for testing:

- Pre-treatment factors: female age, BMI, baseline AMH and AFC, number of prior IVF attempts, IVF indications, duration of infertility and type of infertility.
- Treatment factors: duration of stimulation, total dose of rFSH, type of trigger (hCG or GnRH agonist), estradiol level on day of trigger, progesterone level on day of trigger, and the number of oocytes retrieved.
- Laboratory factors: total number of embryos obtained, number of good-quality embryos.
- Factors associated with embryo transfer: mode of transfer (fresh or frozen transfer during the first transfer), number of subsequent FET after first transfer.

Variables with a p value of ≤ 0.25 were selected for multi-variable analysis. To reduce the number of variables for the final model to predict live birth, we conducted Pearson's correlation analysis for paired samples. A coefficient of correlation of 0.7 ($P < 0.05$) was obtained for the following data pairs:

- Number of oocytes retrieved and total number of embryos obtained
- Total number of embryos obtained and number of good-quality embryos
- Number of embryos and number of subsequent FET after the first transfer
- Number of good-quality embryos and number of subsequent FET after the first transfer

Only one variable from each data pair was selected. Because the number of oocytes retrieved led to the embryology outcomes, and the number of subsequent FET after the first transfer directly correlates with the probability of live birth, these two variables were included for multi-variable Cox regression. Therefore, there were a total of 14 variables for regression modelling. From these 14 variables, we can expect a total of $2^{14} = 16,384$ models to be developed.

2.5. Ethical approval

The study was approved by the Institutional Review Board (approval number CS/MD/17/13). The IVFMD database contained information that was routinely collected from all patients, including baseline demographics, cycle data and outcome data. Patient information was kept confidential and patient consent was obtained at the start of IVF treatment

allowing use of their data for research purposes. No medical ethical approval was required for this research.

3. RESULTS

3.1. Development cohort

From April 2014 to December 2015, a total of 4551 women underwent IVF/ICSI at My Duc Hospital. Approximately 2810 women (61.7%) fulfilled the inclusion/exclusion criteria and formed the development cohort. Two hundred and ten women returned to their home town to give birth and could not be contacted and were therefore lost to follow-up, leaving a dataset of 2600 women for model development (Figure 1). Table 1 shows the baseline demographics and IVF cycle characteristics.

The overall rate of at least one live birth from the whole dataset was 39.2%. Live birth rate per cycle (inclusive of all fresh and frozen embryo transfers) at 12 months was 28.6%. Live birth rate was 50.0% for women aged ≤ 25 years, 45.5% for women aged 26–30 years, 40.2% for women aged 31–35 years, 30.7% for women aged 36–40 years and 8.2% for women aged >40 years. The cumulative live birth rate in women who had undergone fresh embryo transfer followed by subsequent FET cycle(s) was 40.8%. This was higher than the 35.5% live birth rate for patients managed using an initial freeze-only strategy.

3.2. Model development

Bi-variable associations between potential predictors for live birth at 12 months following IVF are shown in Table 2. There were sixteen important predictive variables: female age, AMH, AFC, duration of infertility, type of infertility, IVF indications, number of IVF attempts, duration of stimulation, total dose of rFSH used, type of trigger, progesterone level on the day of trigger, number of oocytes retrieved, number of embryos, number of good embryos, fresh or frozen transfer for the first transfer, and the number of subsequent FET after the first transfer. We applied Bayesian model averaging to all possible models (Figure 2), using the Bayesian informative criteria (BIC) approximation and posterior probability for 'best model' selection. A model that has the lowest BIC and highest posterior probability is considered the best. Therefore, we selected the 3 best models, which had comparable BIC and posterior probability (Table 3); the remaining two models had higher BIC and lower posterior probability:

- Model I consisted of five predictive factors: female age, total dose of rFSH used, type of trigger, fresh or frozen transfer during the first transfer, and number of subsequent FET after the first transfer.

- Model II consisted of seven predictive factors: female age, AMH, total dose of rFSH used, type of trigger, fresh or frozen transfer during the first transfer, number of subsequent FET after the first transfer, and progesterone level on the day of trigger.

- Model III consisted of six predictive factors: female age, total dose of rFSH used, type of trigger, fresh or frozen transfer during the first transfer, number of FET after the first

transfer, and progesterone level on the trigger day.

3.3. Missing data

Missing data occurred for BMI (1.03%), AMH (3.42%), AFC (9.42%), duration of infertility (6.73%), and estradiol and progesterone levels on the day of trigger (6.31% and 6.54%, respectively) (Figure 3). To avoid potential bias and statistical inaccuracy due to loss of data, we used multi-variable

imputation by chained equation (MICE) method to complete multi-variable predictors with missing data. The MICE algorithm generates plausible synthetic values for predictors with missing information within a column, conditional on all other columns in the data. For this purpose, we assumed that the missing data occurred randomly. Imputation was performed using R package *mice* (version 2.30)

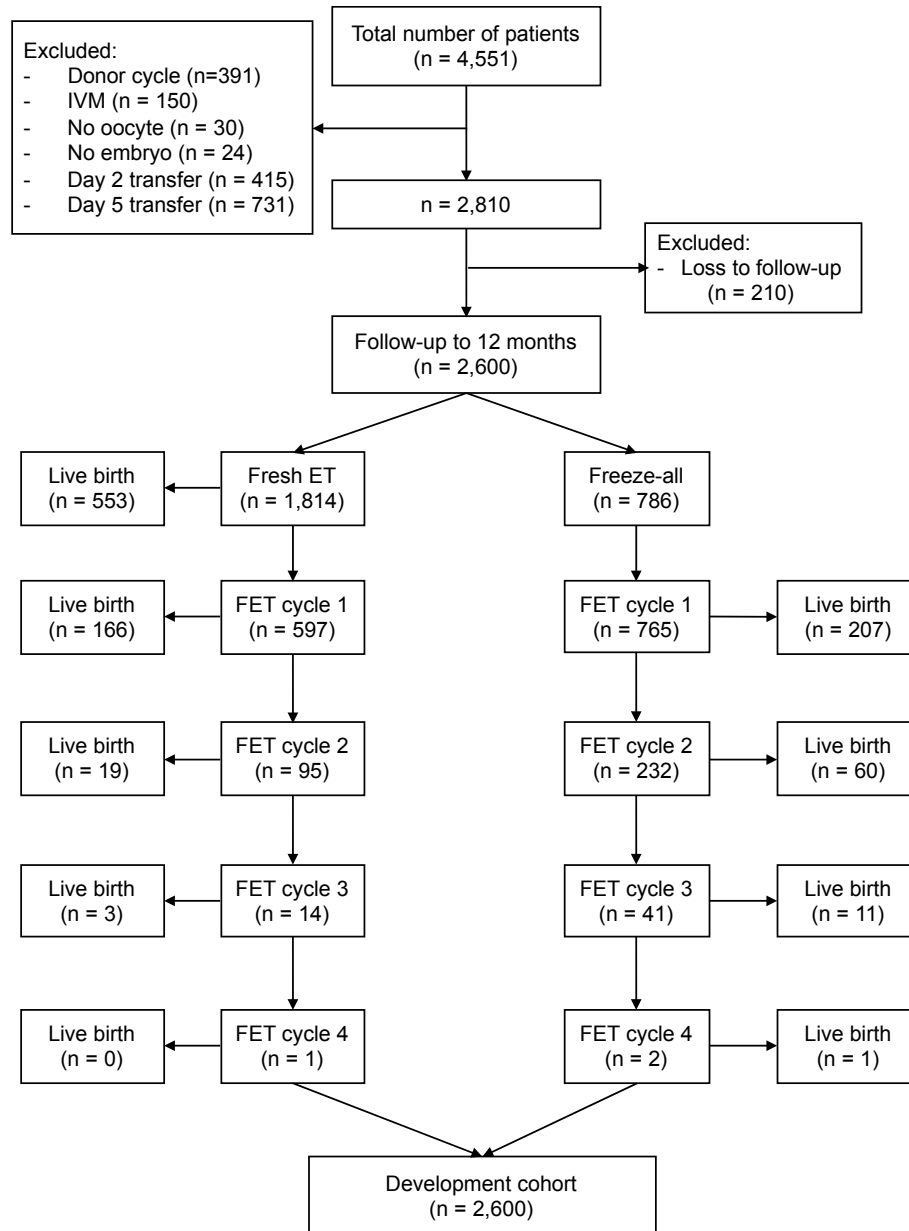


Figure 1. Definition of eligible development cohort.

A total of 1,814 fresh embryo transfers were performed. Of these, 553 resulted in live birth, 597 had no live birth but had frozen embryos available and underwent frozen embryo transfer, and 664 had no live birth but no subsequent frozen transfer (so had fresh embryo transfer only). ET = embryo transfer; FET = frozen embryo transfer; IVM = in vitro maturation.

Table 1. Patient demographics and cycle characteristics in the development and validation cohort

Characteristics	Development cohort (n = 2600)		Validation cohort (n = 1416)	
	No live birth (n = 1,580)	Live birth (n = 1,020)	No live birth (n = 873)	Live birth (n = 543)
Age – years (mean ± SD)	33.0 ± 4.7**	31.4 ± 4.0**	33.5 ± 4.9**	31.8 ± 4.5**
Age group – years (%)				
≤25	71 (4.5)	71 (7.0)	33 (3.8)	40 (7.4)
26–30	433 (27.4)	362 (35.5)	214 (24.5)	166 (30.6)
31–35	619 (39.2)	416 (40.8)	326 (37.3)	224 (41.3)
36–40	368 (23.3)	163 (16.0)	230 (26.3)	97 (17.9)
>40	89 (5.6)	8 (0.8)	70 (8.0)	16 (2.9)
BMI – kg/m ² (mean ± SD)	20.9 ± 2.3	21.0 ± 2.3	21.0 ± 2.4	21.0 ± 2.3
Anti-Müllerian hormone – ng/mL (mean ± SD)	5.4 ± 3.6**	6.3 ± 3.9**	4.1 ± 2.9**	5.4 ± 3.6**
Antral follicle count – n (mean ± SD)	13.0 ± 6.8**	14.9 ± 6.7**	11.2 ± 6.4**	12.7 ± 6.4**
Duration of infertility – year (mean ± SD)	5.0 ± 3.5*	4.6 ± 3.3*	5.0 ± 3.5	4.8 ± 3.2
Type of infertility – n (%)	*	*		
Primary	620 (39.2)	453 (44.4)	512 (58.6)	336 (61.9)
Secondary	960 (60.8)	567 (55.6)	361 (41.4)	207 (38.1)
IVF indications – n (%)	**	**	**	**
Male factor infertility	546 (34.6)	434 (42.5)	266 (30.5)	214 (39.4)
Tubal	321 (20.3)	198 (19.4)	168 (19.2)	103 (19.0)
Diminished ovarian reserve	218 (13.8)	64 (6.3)	184 (21.1)	69 (12.7)
Unexplained	219 (13.9)	133 (13.0)	129 (14.8)	64 (11.8)
Ovulation disorder	129 (8.2)	116 (11.4)	54 (6.2)	57 (10.5)
Endometriosis	47 (3.0)	19 (1.9)	22 (2.5)	9 (1.7)
Others	100 (6.3)	56 (5.5)	50 (5.7)	27 (5.0)
Number of IVF attempts – n (%)	*	*	*	*
1	1181 (74.7)	823 (80.7)	680 (77.9)	422 (77.7)
2	262 (16.6)	146 (14.3)	123 (14.1)	86 (15.8)
3	89 (5.6)	38 (3.7)	46 (5.3)	19 (3.5)
4	25 (1.6)	7 (0.7)	17 (1.9)	6 (1.1)

	Development cohort (n = 2600)		Validation cohort (n = 1416)	
5	15 (0.9)	3 (0.3)	6 (0.7)	5 (0.9)
≥6	8 (0.5)	3 (0.3)	1 (0.1)	5 (0.9)
Duration of stimulation – days (mean ± SD)	9.2 ± 1.5	9.1 ± 1.4	9.1 ± 1.4	9.0 ± 1.3
Total dose of rFSH used – IU (mean ± SD)	2648.3 ± 909.9**	2424.4 ± 843.5**	2997.5 ± 837.0**	2725.6 ± 781.8**
Type of trigger – n (%)		*	*	*
Agonist trigger	174 (11.0)	98 (9.6)	93 (10.7)	87 (16.0)
hCG trigger	1406 (89.0)	922 (90.4)	780 (89.3)	456 (84.0)
Estradiol level on day of trigger – pg/mL (mean ± SD)	7170.6 ± 6135.3	7440.3 ± 6351.8	5235.9 ± 4773.2**	6453.2 ± 5716.6**
Progesterone level on day of trigger – ng/mL (mean ± SD)	1.3 ± 1.0	1.3 ± 1.1	1.2 ± 1.1	1.2 ± 0.8
<1.5	1046 (70.8)	719 (75.4)	665 (80.2)	411 (80.9)
≥1.5	431 (29.2)	234 (24.6)	164 (19.8)	97 (19.1)
Number of oocyte retrieved – n (mean ± SD)	12.4 ± 6.1**	14.2 ± 6.1**	11.2 ± 5.7**	13.5 ± 6.4**
Proportion of women with the number of oocytes retrieved – n (%)				
1 – 3	28 (1.8)	5 (0.5)	18 (2.1)	6 (1.1)
4 – 9	485 (30.7)	217 (21.3)	349 (40.0)	151 (27.8)
10 – 15	651 (41.2)	460 (45.1)	305 (34.9)	206 (37.9)
>15	416 (26.3)	338 (33.1)	201 (23.0)	180 (33.1)
Number of embryos – n (mean ± SD)	5.6 ± 3.3**	7.0 ± 3.3**	5.5 ± 3.4**	7.0 ± 3.6**
Number of good embryos – n (mean ± SD)	4.0 ± 2.9**	5.3 ± 3.0**	3.8 ± 2.9**	5.0 ± 3.2**
Fresh or frozen transfer during first ET – n (%)	*	*	*	*
Fresh ET	1073 (67.9)	741 (72.6)	585 (67.0)	347 (63.9)
Frozen ET	507 (32.1)	279 (27.4)	288 (33.0)	196 (36.1)
Number of subsequent FET after first transfer – n (%)	**	**	**	**
0	923 (58.4)	477 (46.8)	364 (41.7)	159 (29.3)
1 – 2	489 (30.9)	357 (35.0)	388 (44.4)	288 (53.0)
3 – 4	137 (8.7)	146 (14.3)	117 (13.4)	83 (15.3)
≥5	31 (2.0)	40 (3.9)	4 (0.5)	13 (2.4)

*P < 0.05; **P < 0.001.

BMI = body mass index; ET = embryo transfer; FET = frozen embryo transfer; hCG = human chorionic gonadotropin; IVF = in vitro fertilization; rFSH = recombinant follicle stimulating hormone; SD = standard deviation.

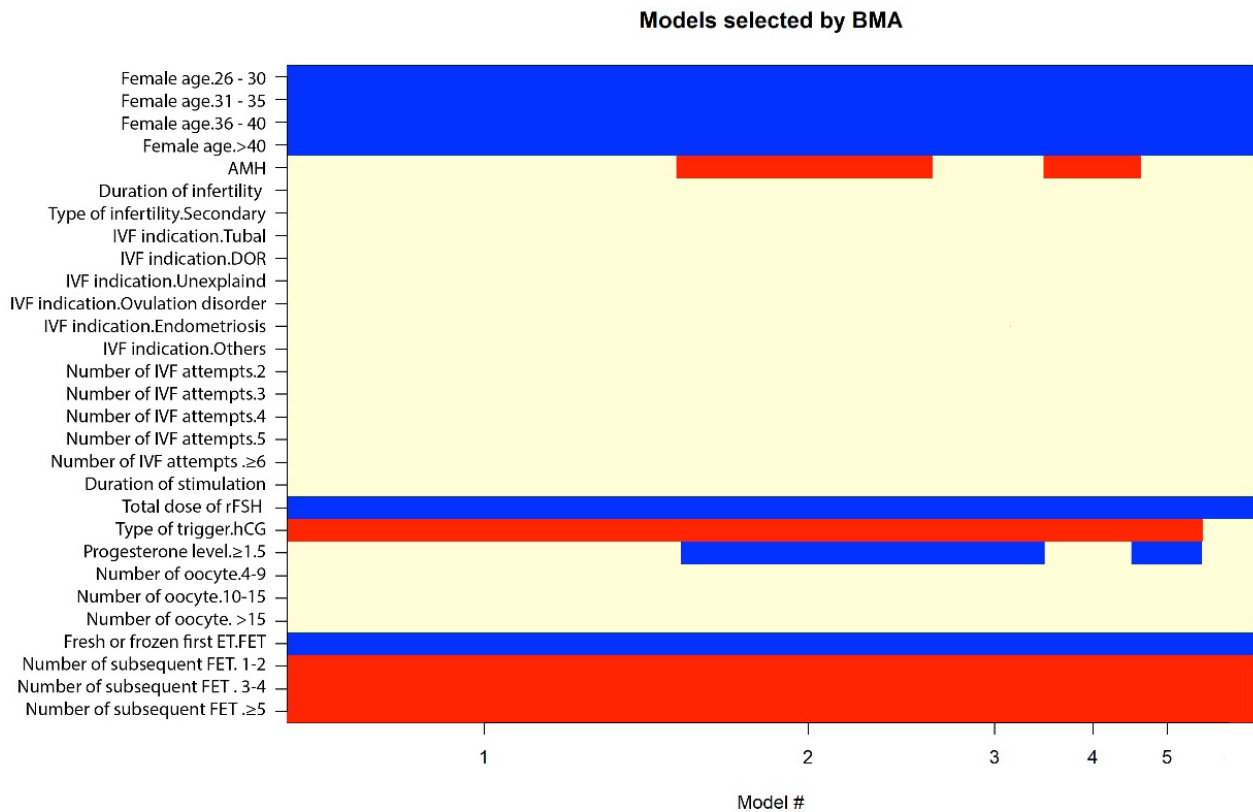


Figure 2: Models selected by Bayesian Model Averaging (BMA).

Variable estimates are negative (red) or positive (blue); uncoloured variables were not included in the model).

On the x-axis, models are listed in order of decreasing posterior model probability.

3.4. Model performance

There were no significant differences in the area under the ROC curve (AUROC) between the three models (DeLong test, $p > 0.05$) (Figure 3). Discrimination, assessed by determining the AUC, was similar for all three models. The AUC was 0.63 (95% confidence interval [CI] 0.60–0.65) for Model I, and 0.63 (95% CI 0.61–0.65) for Models II and III. The calibration plots for all three models showed good agreement between the predicted and observed probabilities of live birth (Hosmer-Lemeshow test, $p > 0.05$) (Figure 4).

3.5. Model validation

To validate Models I, II and III, we obtained data from a new cohort of patients undergoing IVF treatment at the same centre over different period (January to June 2016). A total of 1416 patients met the inclusion and exclusion criteria (Figure 5). The baseline demographics and IVF cycle characteristics of the validation cohort were comparable to those of the development cohort (Table 1). The overall birth rate of the validation cohort was 38.3% (543/1416), which was comparable to the live birth rate of the development cohort (39.2%). When Models I, II and III were applied to the validation cohort, the AUC was 0.60 (95% CI 0.57–0.63) for all three models (Figure 6). This was slightly lower than the value obtained from the development cohort.

3.6. Final model

All three models were comparable with regards to BIC, posterior probability, performance and validation. We selected Model I as the most parsimonious model because it required the least number of predictive factors for input. The nomogram showing probability of live birth in an individual patient is shown in Figure 7. The C-index for the nomogram is the AUCROC of Model I (0.63, 0.60–0.65).

4. DISCUSSION

In our study, we identified five factors that significantly influenced the probability of live birth. As expected, age was one of the most important factors influencing the outcome of IVF/ICSI. The hazard ratio for live birth in women over the age of 40 years was considerably lower at 0.15 (95% CI 0.07–0.31, $p < 0.001$) versus 0.59 (95% CI 0.44–0.78, $p < 0.001$) in the 36–40 years age group. In the bi-variable analysis, we demonstrated that baseline AMH and AFC were significantly associated with live birth. Studies have shown both AMH and AFC to be good predictors of ovarian response independent of age, suggesting that there may be a positive role for the inclusion of biomarkers for ovarian reserve in the prediction for live birth [19–21]. However, AMH and AFC were excluded from our final multi-variable model because we did not find any significant difference in discrimination when these variables were added either alone or together.

This was in agreement with the findings of another study [22], in which the combination of AMH (with or without AFC) and age only correctly classified an additional < 2% of subjects compared with the model including age alone.

Another important predictive factor identified in our study was the duration of infertility. The mean duration of infertility for women with no live birth was 5.0 ± 3.5 years compared with 4.6 ± 3.3 years for women with live birth. Women with a longer duration of infertility were less likely to achieve live birth, with a hazard ratio (HR) of 1.20 (95% CI 1.05–1.36, $p = 0.007$) for those who were infertile for ≥ 5 years versus those who were infertile for 4 years or less. We excluded this variable from the final model for the same reason as AMH and AFC. Another significant predictive factor was rFSH consumption. The mean total rFSH dose used was 2648.3 ± 909.9 IU for women with no live birth and 2424.4 ± 843.5 IU for those with live birth. For every increase of 300 IU rFSH, the probability of live birth was reduced by 37% (HR 0.93 [CI 95% 0.91–0.95], $p < 0.001$).

Not all women in our study used all their cryopreserved embryos within the 1-year follow-up period. Three hundred and eighteen women had not achieved live birth but still had remaining frozen embryos. In two different studies [23, 24], the minimum follow-up time from the start of IVF cycle was 2 years. Nevertheless, our follow-up period for assessing live birth at 12 months can be justified because preliminary analysis showed that the median time from the start of treatment to live birth at our center was 9.5 months. In addition, the proportion of patients returning after 12 months was low (<8%). Furthermore, IVF is not subsidized or covered by health insurance in Vietnam meaning that patients pay for their own IVF treatment and usually do not have the financial means to pursue treatment for up to 2 years.

To our knowledge, this is the first predictive model for live birth to be derived from an Asian population. We assessed the probability of live birth in women receiving ICSI and GnRH antagonist treatment, both of which are favoured and practiced in many countries of this region [25–29]. Existing multi-variable models for predicting live birth had been derived based on GnRH agonist cycles [12, 30], excluded ICSI or FET [11, 13], or were limited to fresh single embryo transfer when GnRH antagonist treatment was used [31, 32]. Hence, a major strength of our model is its relevance to current ART practices, particularly in the region from which data were used to generate the model. The nomogram developed is the first of its kind to allow prediction of live birth based on pre-treatment, post-stimulation, laboratory and embryo transfer parameters.

Of existing models that used live birth as the primary outcome, only the Nelson model has been validated in an Asian country (Singapore) [33]. However, this showed that the Nelson model fitted poorly with the local study population. This is likely due to differences in ethnicity and the legally-mandated younger age limit for IVF treatment (≤ 45 years). Unlike the Nelson model, ours was derived from an Asian population and included a younger age group (the maximum age for IVF treatment in our study was 45 years versus 50 years in the Nelson model). Therefore, our prediction model may have greater reproducibility and generalizability in countries whose population share similar age-related biological or physiological attributes with our patients.

We acknowledge several shortcomings in our study.

Firstly, the number of women returning to use frozen embryos within 12 months after the first embryo transfer was low. In addition, the AUC of our final model and validation cohort demonstrated poor discriminative ability to separate women with and without live birth at 1 year with an AUC of 0.63 and 0.60, respectively. An AUC of 1 indicates perfect discrimination whereas an AUC of 0.5 indicates no discrimination. A model is considered to have poor performance if the AUC lies between 0.5 and 0.7, fair performance if the AUC lies between 0.7 and 0.8, and good performance if the AUC lies between 0.8 and 0.9 [9]. Existing models for live birth following IVF treatment that have been externally validated had shown similarly poor discrimination. The Templeton and Nelson models each had an AUC of 0.63 [11, 34], while the model by Dhillon *et al.* had an AUC of 0.62 [10]. For IVF models, achieving fair or good discrimination is unfeasible because the AUC typically does not exceed 0.67. As such, calibration is considered a more meaningful measure of model performance than discrimination [35]. Our model had good calibration as measured using the Hosmer-Lemeshow goodness-of-fit statistic, indicating close agreement between predicted and observed live births.

Another potential limitation is that the live birth rate may be affected by confounding factors that potentially arise as the pregnancy progresses. For example, prematurity and pregnancy complications, such as hypertension, gestational diabetes, and intrauterine growth retardation could reduce the probability of live birth. Lastly, as with any retrospective study designed to analyse pre-existing data, the presence of biases cannot be excluded [36]. To mitigate random error and statistical bias, we used a relatively large sample size ($n = 2600$) to construct this model.

It is important for clinicians to manage patient expectations about their chances for successful outcomes at different stages of the IVF cycle. The main function of our model is to create a patient counselling tool that accounts for all relevant predictors of live birth starting from the time of presentation to the day of embryo transfer. By collectively assessing these factors, clinicians will be able to provide patients with a more accurate prognosis than with pre-treatment information alone.

In places where access to IVF treatment is regulated by insurance reimbursement or legislative policies, patients with poor pre-treatment factors (e.g. advanced age or low ovarian reserve) are often denied IVF treatment due to their low chance of achieving live birth and advised to consider donor IVF or adoption instead. This is often contrary to the patients' desire to have their own genetic offspring. From the clinician's perspective, predicting reproductive outcomes is a dynamic process. For example, 35/2600 patients (1.3%) in our development cohort were aged >38 years and had AMH ≤ 1.25 ng/mL, but 22/35 (62.9%) had at least 1 good embryo for transfer. Therefore, the probability of live birth should be adjusted according to the patient's response to IVF treatment up to the day of embryo transfer.

5. CONCLUSIONS

The discriminative ability of our model is comparable with previous models and shows good calibration. It has potential for implementation in clinical practice, particularly in settings where similar IVF/ICSI practices described in this study are employed. Our model should provide a more individualized and objective means for counselling patients

about treatment outcomes following ovarian stimulation and oocyte retrieval, and to advise on the probability of live birth with fresh versus freeze-all cycles, as well as manage the patient's expectations about subsequent FET

cycles if the initial fresh transfer was unsuccessful. Future studies should include geographical external validation and impact analyses to evaluate the true benefits of integrating this model as part of the IVF patient care process.

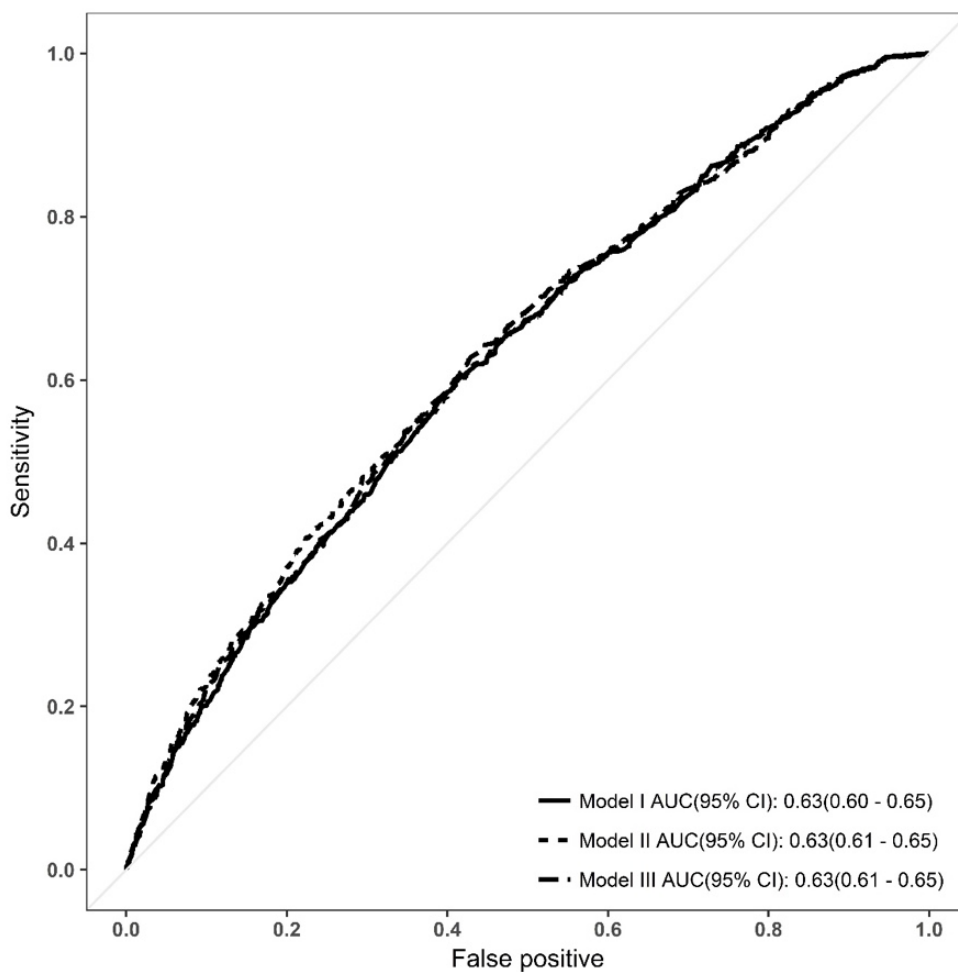


Figure 3. Area under the receiver operating characteristics (AUROC) curve for Models I, II and III.

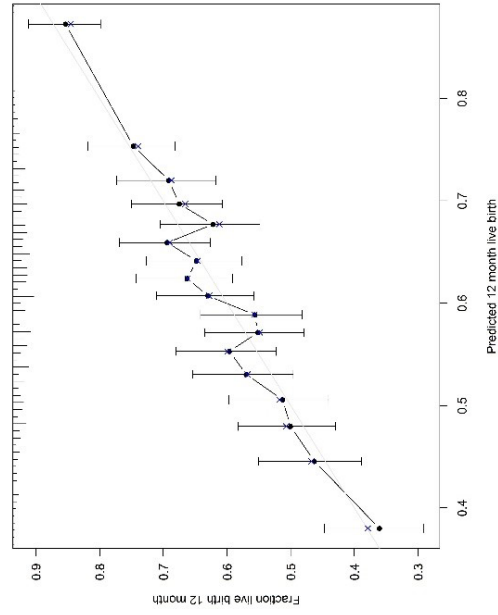
AUC = area under the curve; CI, confidence interval.

Table 2. Bi-variable associations of predictive factors for live birth.

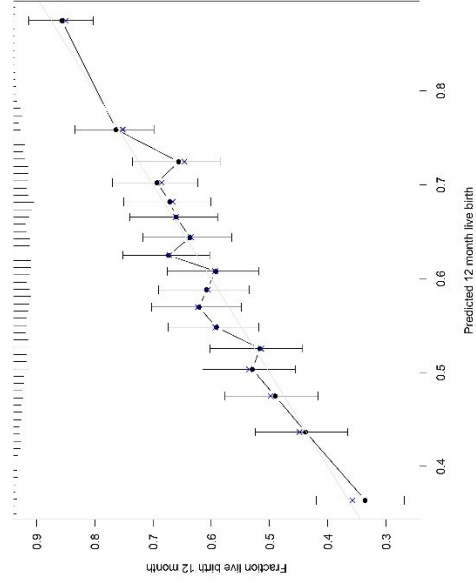
Characteristics	N = 2600	Unit	Hazard ratio (95% CI)	P-value
Age – year	2600	+1	0.95 [0.94, 0.96]	<0.001
≤25			Reference	
26 – 30			0.91 [0.70, 1.17]	0.444
31 – 35			0.77 [0.60, 0.99]	0.041
36 – 40			0.56 [0.42, 0.73]	<0.001
>40			0.13 [0.06, 0.27]	<0.001
BMI – kg/m ²	2573	+1	1.01 [0.98, 1.03]	0.668
Anti-Müllerian hormone – ng/mL	2511	+1	1.04 [1.03, 1.06]	<0.001
Antral follicle count – n	2355	+1	1.03 [1.02, 1.04]	<0.001
Duration of infertility – year	2425	+1	0.97 [0.96, 0.99]	0.011
Type of infertility – n (%)	2600			
Primary			Reference	
Secondary			0.86 [0.76, 0.97]	0.016
IVF indication – n (%)	2600			
Male			Reference	
Tubal			0.82 [0.70, 0.98]	0.024
Diminished ovarian reserve			0.45 [0.34, 0.58]	<0.001
Unexplained			0.84 [0.69, 1.02]	0.085
Ovulation disorder			1.09 [0.89, 1.34]	0.407
Endometriosis			0.56 [0.35, 0.88]	0.013
Others			0.74 [0.56, 0.98]	0.038
Number of IVF attempts – n (%)	2600			
1			Reference	
2			0.82 [0.69, 0.98]	0.028
3			0.68 [0.49, 0.95]	0.023
4			0.46 [0.22, 0.97]	0.043
5			0.36 [0.12, 1.12]	0.078
≥6			0.63 [0.20, 1.97]	0.432

Characteristics	N = 2600	Unit	Hazard ratio (95% CI)	P-value
Duration of stimulation – days	2600		0.94 [0.90, 0.98]	0.005
Total dose of rFSH – IU	2600	+300	0.93 [0.91, 0.95]	<0.001
Type of trigger – n (%)	2600			
Agonist trigger			Reference	
hCG trigger			1.20 [0.97, 1.47]	0.092
Estradiol level on day of trigger – pg/mL	2436	+1500	1.00 [0.98, 1.01]	0.855
Progesterone level on day of trigger – ng/mL	2430	+1	0.94 [0.87, 1.01]	0.096
<1.5			Reference	
≥1.5			0.78 [0.67, 0.91]	0.001
Number of oocyte retrieved – n	2600	+1	1.03 [1.02, 1.04]	<0.001
1 – 3			Reference	
4 – 9			2.20 [0.91, 5.35]	0.081
10 – 15			3.11 [1.29, 7.50]	0.012
>15			3.37 [1.39, 8.14]	0.007
Number of embryos – n	2600	+1	1.07 [1.05, 1.09]	<0.001
Number of good embryos – n	2600	+1	1.08 [1.06, 1.10]	<0.001
Fresh or frozen transfer during first ET – (n, %)	2600			
FET			Reference	
Fresh ET			1.39 [1.21, 1.59]	<0.001
Number of subsequent FET after first transfer – n (%)	2600			
0			Reference	
1 – 2			1.30 [1.13, 1.49]	<0.001
3 – 4			1.61 [1.34, 1.94]	<0.001
≥5			1.80 [1.30, 2.48]	<0.001

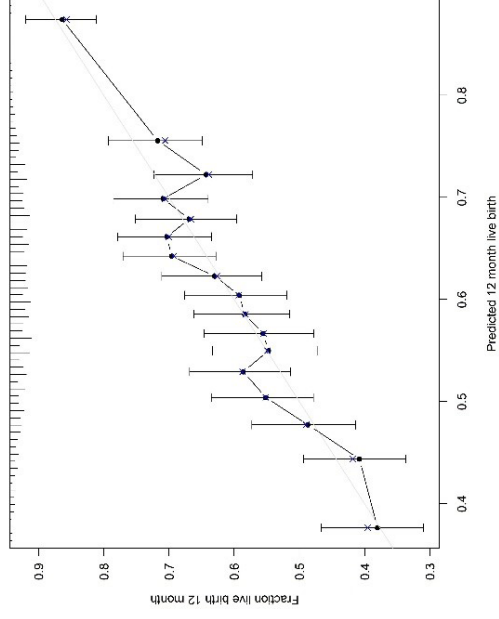
BMI = body mass index; ET = embryo transfer; FET = frozen embryo transfer; hCG = human chorionic gonadotropin; IVF = in vitro fertilization; rFSH = recombinant follicle stimulating hormone; SD = standard deviation.



Model I



Model II



Model III

Figure 4: Calibration plot.

N = 2600, live birth = 1020, 150 subjects per group. Resampling optimism added using bootstrap method 300 times.

Dots: predicted live birth rate compared with fraction livebirth rate; Crosses: resampling optimism added using bootstrap method 300 times;

Grey line: ideal where predicted live birth equals to fraction live birth; Mean standard error: 0.066.

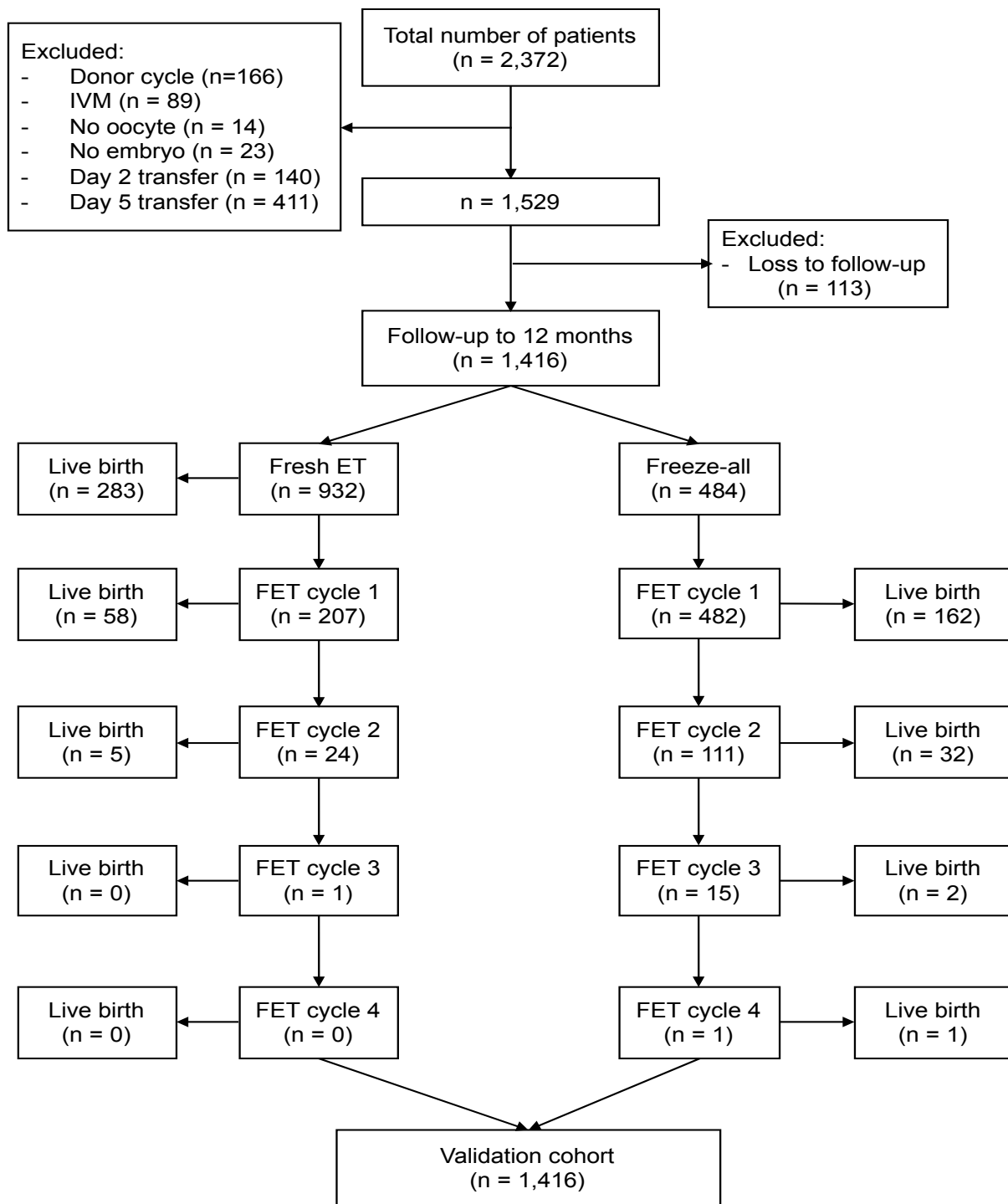
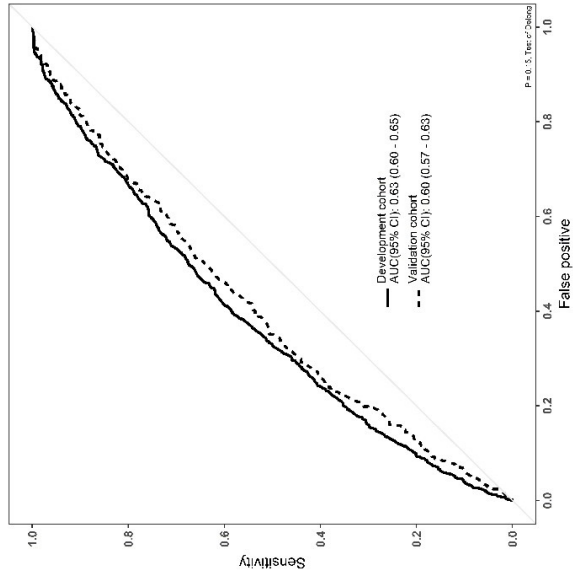
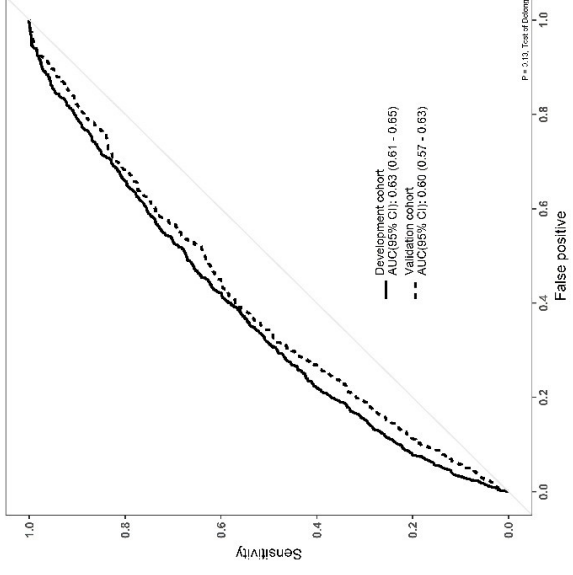


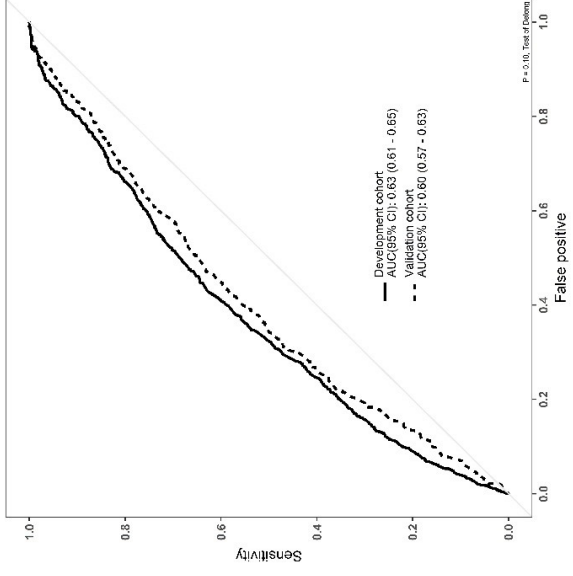
Figure 5. Definition of eligible validation cohort. ET = embryo transfer; FET = frozen embryo transfer; IVM = in vitro maturation.



Model I



Model II



Model III

Figure 6. Area under the receiver operating characteristics (AUROC) curve of the validation versus development cohort.

AUC = area under the curve; CI, confidence interval.

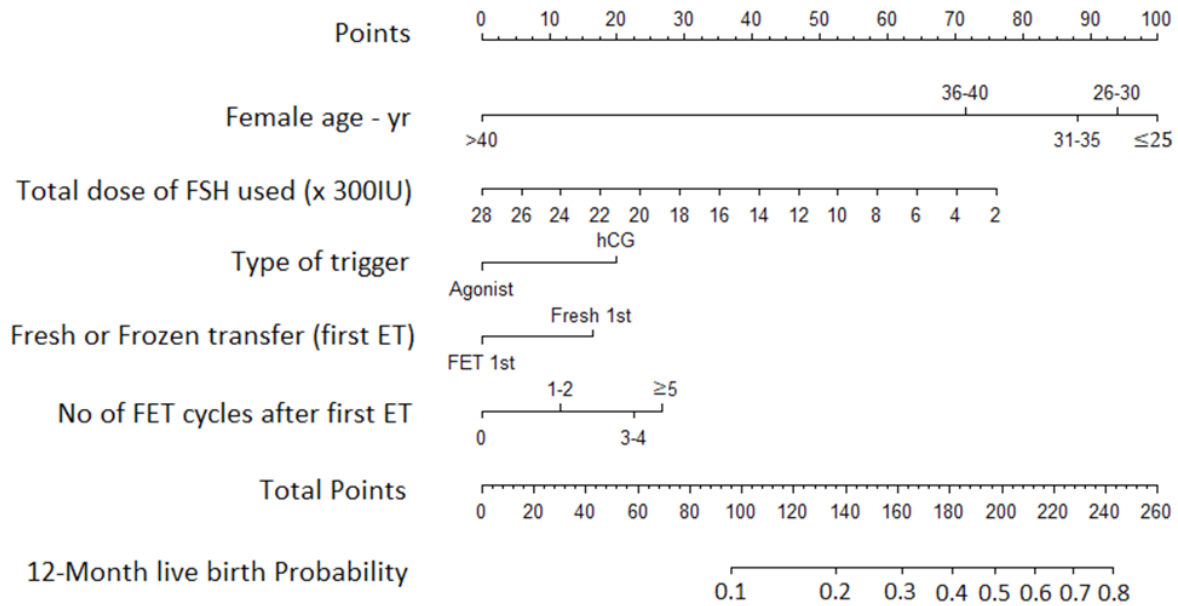


Figure 7. Nomogram for predicting live birth at 12 months after starting IVF/ICSI treatment based on Model I.

Instruction for use: Mark the patient's age at the "Female age" axis and draw a vertical line to the "Points" axis.

The "Points" axis indicates how many points towards the probability of live birth for the patient's age. Note the points obtained for each predictor and sum them up to get the total points. Locate the total points on the "Total points" axis. Draw a vertical line from the 'Total points' axis to the "12-month live birth probability" axis. The value obtained on the "12-month live birth probability" indicates the patient's probability for a favourable outcome following all fresh/frozen embryo transfers in a IVF cycle.

*ET = embryo transfer; FET = frozen embryo transfer; FSH = follicle stimulating hormone;
hCG = human chorionic gonadotrophin; IU = international unit; yr = year.*

REFERENCES

- European Society of Human Reproduction and Embryology (ESHRE) Helsinki Highlights, September 2016. ICMART puts global ART baby total at 6.5 million. Focus on Reproduction. Available at: <https://www.eshre.eu/Publications/Focus-on-Reproduction/Focus-in-2016.aspx>. Last accessed 11 September 2017.
- V.A. Moragianni, A.S. Penzias. Cumulative live-birth rates after assisted reproductive technology. *Curr Opin Obstet Gynecol*. 2010;3:189–92.
- D.J. McLernon, E.W. Steyerberg, E.R. te Velde, A.J. Lee, S. Bhattacharya. Predicting the chances of a live birth after one or more complete cycles of in vitro fertilisation: population based study of linked cycle data from 113873 women. *Br Med J*. 2016;355:i5735.
- K. Lukaszuk, M. Kunicki, J. Liss, M. Lukaszuk, G. Jakiel. Use of ovarian reserve parameters for predicting live births in women undergoing in vitro fertilization. *Eur J Obstet Gynecol Reprod Biol*. 2013;168:173–7.
- S.E. Elizur, L. Lerner-Geva, J. Levron, A. Shulman, D. Bider, J. Dor. Factors predicting IVF treatment outcome: a multivariate analysis of 5310 cycles. *Reprod Biomed Online*. 2005;10:645–9.
- J. Carrera-Rotllan, L. Estrada-Garcia, J. Sarquella-Ventura. Prediction of pregnancy in IVF cycles on the fourth day of ovarian stimulation. *J Assist Reprod Genet*. 2007;24:387–94.
- S.S. Srouji, A. Mark, Z. Levine, R.A. Betensky, M.D. Hornstein, E.S. Ginsburg. Predicting in vitro fertilization live birth using stimulation day 6 estradiol, age, and follicle-stimulating hormone. *Fertil Steril*. 2005;84:795–7.
- L. van Loendersloot, S. Repping, P.M.M. Bossuyt, F. van der Veen, M. van Wely. Prediction models in in vitro fertilization; where are we? A mini review. *J Adv Res*. 2014;5:295–301.
- E. Leushuis, J.W. van der Steeg, P. Steures, P.M. Bossuyt, M.J. Eijkemans, F. van der Veen F, et al. Prediction models in reproductive medicine: a critical appraisal. *Hum. Reprod. Update*. 2009;15:537–52.
- R.K. Dhillon, D.J. McLernon, P.P. Smith, S. Fishel, K. Dowell, J.J. Deeks, et al. Predicting the chance of live birth for women undergoing IVF: a novel pretreatment counselling tool. *Hum Reprod*. 2016;31:84–92.
- S.M. Nelson, D.A. Lawlor. Predicting live birth, preterm delivery, and low birth weight in infants born from in vitro fertilisation: a prospective study of 144,018 treatment cycles. *PLoS Med*. 2011;8:e1000386.
- D. Minaretzis, D. Harris, M.M. Alper, J.F. Mortola, M.J. Berger, D. Power. Multivariate analysis of factors predictive of successful live births in in vitro fertilization (IVF) suggests strategies to improve IVF outcome. *J Assist Reprod Genet*. 1998;15:365–71.
- A. Templeton, J.K. Morris, W. Parshlow. Factors that affect outcome of in-vitro fertilisation treatment. *Lancet*. 1996;348:1402–6.
- L.A. Humphries, O. Chang, K. Humm, D. Sakkas, M.R. Hacker. Influence of race and ethnicity on in vitro fertilization outcomes: systematic review. *Am J Obstet Gynecol*. 2016;214:212.e1–17.
- R.K. Dhillon, P.P. Smith, R. Malhas, H.M. Harb, I.D. Gallos, K. Dowell, et al. Investigating the effect of ethnicity on IVF outcome. *Reprod Biomed Online*. 2015;31:356–63.
- H.G. Al-Inany, M.A. Youssef, R.O. Aweleke, J. Brown, W.S. Lam, F.J. Broekmans. Gonadotrophin-releasing hormone antagonists for assisted reproductive technology. *Cochrane Database Syst. Rev*. 2016;4:CD001750.
- S. Dyer, G.M. Chambers, J. de Mouzon, K.G. Nygren, F. Zegers-Hochschild, R. Mansour, et al. International Committee for Monitoring Assisted Reproductive Technologies world report: assisted reproductive technology 2008, 2009 and 2010. *Hum Reprod*. 2016;31:1588–609.
- Alpha Scientists in Reproductive Medicine ESHRE Special Interest Group of Embryology. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Hum Reprod*. 2011;26:1270–83.
- M. Goswami, D. Nikolaou. Is AMH level, independent of age, a predictor of live birth in IVF? *J Hum Reprod Sci*. 2017;10:24–30.
- A. La Marca, S.M. Nelson, G. Sighinolfi, M. Manno, E. Baraldi, L. Roli, et al. Anti-Müllerian hormone-based prediction model for a live birth in assisted reproduction. *Reprod Biomed Online*. 2011;22:341–9.
- 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*. 2009;91:1595–7.
- H.W. Li, V.C. Lee, E.Y. Lau, W.S. Yeung, P.C. Ho, E.H. Ng. Role of baseline antral follicle count and anti-Müllerian hormone in prediction of cumulative live birth in the first in vitro fertilisation cycle: a retrospective cohort analysis. *PLoS One*. 2013;8:e61095.
- A. Maheshwari, D. McLernon, S. Bhattacharya. Cumulative live birth rate: time for a consensus? *Hum Reprod*. 2015;30:2703–7.
- M. Toftager, J. Bogstad, K. Løssl, L. Praetorius, A. Zedeler, T. Bryndorf, et al. Cumulative live birth rates after one ART cycle including all subsequent frozen-thaw cycles in 1050 women: secondary outcome of an RCT comparing GnRH-antagonist and GnRH-agonist protocols. *Hum Reprod*. 2017;32:556–67.
- Y. Sudewo, B. Halin. Comparative study between long protocol with antagonist protocol on IVF cycle. The 6th Congress of the Asia Pacific Initiative on Reproduction (ASPIRE 2016). *KnE Medicine*. 2016;32–42.
- V. Viardot-Foucault, S. Nadarajah, W.K. Lye, H.H. Tan. GnRH antagonist pre-treatment: one centre's experience for IVF ICSI cycle scheduling. *Reprod Biomed Online*. 2015;30:366–72.
- S. Yang, X.N. Chen, J. Qiao, P. Liu, R. Li, G.A. Chen, C.H. Ma. Comparison of GnRH antagonist fixed protocol and GnRH agonists long protocol in infertile patients with normal ovarian reserve function in their first in vitro fertilization-embryo transfer cycle. *Zhonghua Fu Chan Ke Za Zhi*. 2012;47:245–9.
- M.R. Rashid, F.B. Ong, M.H. Omar, S.P. Ng, A. Nurshaireen, N.S.M.N. Sharifah-The, et al. GnRH agonist and GnRH antagonist in intracytoplasmic injection cycles. *Med J Malaysia*. 2008;63:113–7.
- S.Y. Huang, H.Y. Huang, H.T. Yu, H.S. Wang, C.K. Chen, C.L. Lee, et al. Low-dose GnRH antagonist protocol is as effective as the long GnRH agonist protocol in unselected patients undergoing in vitro fertilization and embryo transfer. *Taiwan J Obstet Gynecol*. 2011;50:432–5.
- G. Pettersson, A.N. Andersen, P. Broberg, J.C. Arce. Pre-stimulation parameters predicting live birth after IVF in the long GnRH agonist protocol. *Reprod Biomed Online*. 2010;20:572–81.
- K.K. Vaegter, T.G. Lakic, M. Olovsson, L. Berglund, T. Brodin, J. Holte. 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;107:641–8.
- P. Drakopoulos, C. Blockeel, D. Stoop, M. Camus, M. de Vos, H. Tournaye, et al. Conventional ovarian stimulation and single embryo transfer for IVF/ICSI. How many oocytes to we need to maximize cumulative live birth rates after utilization of all fresh and frozen embryos? *Hum Reprod*. 2016;3:370–6.
- L. Saha, S.M. Fook-Chong, H. Rajesh, D.S. Chia, S.L. Yu. Use of in vitro fertilisation prediction model in an Asian population – experience in Singapore. *Ann Acad Med Singapore*. 2015;44:524–9.
- J.M. Smeenk, A.M. Stolwijk, J.A. Kremer, D.D. Braat. External validation of the Templeton model for predicting success after IVF. *Hum. Reprod*. 2000;15:1065–8.
- S.F.F.J. Coppus, F. van der Veen, B.C. Opmeer, B.W.J. Mol, P.M.M. Bossuyt. Evaluating prediction models in reproductive medicine. *Hum Reprod*. 2009;24:1774–8.
- C.J. Pannucci, E.G. Wilkins. Identifying and avoiding bias in research. *Plast Reconstr Surg*. 2010;126:619–25.