CERM

Nomogram to predict the number of oocytes retrieved in controlled ovarian stimulation

Kyoung Yong Moon[1,2], Hoon Kim[1,3], Joong Yeup Lee[4], Jung Ryeol Lee[1,5], Byung Chul Jee[1,5], Chang Suk Suh[1,3], Ki Chul Kim[4], Won Don Lee[2], Jin Ho Lim[2], Seok Hyun Kim[1,3]

[1]Department of Obstetrics and Gynecology, Seoul National University College of Medicine, Seoul; [2]Seoul Maria Fertility Hospital, Seoul; [3]Department of Obstetrics and Gynecology, Seoul National University Hospital, Seoul; [4]Hamchoon Women's Clinic, Seoul; [5]Department of Obstetrics and Gynecology, Seoul National University Bundang Hospital, Seongnam, Korea

Objective: Ovarian reserve tests are commonly used to predict ovarian response in infertile patients undergoing ovarian stimulation. Although serum markers such as basal follicle-stimulating hormone (FSH) or random anti-Müllerian hormone (AMH) level and ultrasonographic markers (antral follicle count, AFC) are good predictors, no single test has proven to be the best predictor. In this study, we developed appropriate equations and novel nomograms to predict the number of oocytes that will be retrieved using patients' age, serum levels of basal FSH and AMH, and AFC.

Methods: We analyzed a database containing clinical and laboratory information of 141 stimulated *in vitro* fertilization (IVF) cycles performed at a university-based hospital between September 2009 and December 2013. We used generalized linear models for prediction of the number of oocytes.

Results: Age, basal serum FSH level, serum AMH level, and AFC were significantly related to the number of oocytes retrieved according to the univariate and multivariate analyses. The equations that predicted the number of oocytes retrieved (log scale) were as follows: model (1) $3.21-0.036\times(\text{age})+0.089\times(\text{AMH})$, model (2) $3.422-0.03\times(\text{age})-0.049\times(\text{FSH})+0.08\times(\text{AMH})$, model (3) $2.32-0.017\times(\text{age})+0.039\times(\text{AMH})+0.03\times(\text{AFC})$, model (4) $2.584-0.015\times(\text{age})-0.035\times(\text{FSH})+0.038\times(\text{AMH})+0.026\times(\text{AFC})$. model 4 showed the best performance. On the basis of these variables, we developed nomograms to predict the number of oocytes that can be retrieved.

Conclusion: Our nomograms helped predict the number of oocytes retrieved in stimulated IVF cycles.

Keywords: Anti-Müllerian hormone; Fertilization *in vitro*; Nomograms; Oocytes; Ovarian reserve

Introduction

Prediction of individual ovarian response to exogenous gonadotropin is one of the most important strategies for successful and safe *in vitro* fertilization (IVF) treatment. Basal serum follicle-stimulating hor-

mone (FSH) concentration is traditionally the most commonly used ovarian reserve test (ORT). A menstrual cycle day 3 serum FSH concentration >10–20 IU/L indicates diminished ovarian function, which shows high specificity (80%–100%) for the prediction of a poor response to ovarian stimulation but low sensitivity (10%–30%) [1]; however, it is difficult to predict a high response.

Anti-Müllerian hormone (AMH) is another serum marker that can be used as an ORT. The AMH level is associated with low oocyte retrieval and poor embryo quality, but this factor has low sensitivity and specificity for predicting successful achievement of pregnancy [2-5]. Serum AMH level reflects the overall amount of granulosa cells in the follicular pool; thus, the AMH level is generally known to be proportional to the number of follicles. However, the serum AMH level depends on follicle size, granulosa cell volume or maturity, and

KY Moon et al. Nomographic estimate of oocyte retrieval

CERM

intrafollicular environment, which are inconsistent among follicles and affected by individual genetic characteristics [6]. A previous study noted the presence of inter-test discrepancies that occurred as frequently as one in every five subjects when serum values of AMH <0.8 ng/mL and FSH >10 IU/L were defined as risk factors for poor ovarian response, and this was more common with an increase in age [7].

Antral follicle count (AFC) is an ultrasonographic marker used as an ORT that is proportional to the number of primordial follicles remaining in the ovary [8]. AFC is known to have a high specificity (73%–100%) for predicting a poor response, but exhibits low sensitivity (9%–73%) for predicting a poor response and low sensitivity (8%–33%) for predicting a failure to achieve pregnancy [9-12]. Large inter-observer variation is an inherent shortcoming because of the subjective nature of ultrasonography for measuring follicle counts.

Thus far, no single ORT has shown superior accuracy. Therefore, the present study aimed to investigate the combined relationship of age and various ORTs with the final number of oocytes retrieved in order to develop appropriate equations and novel nomograms to predict the final number of oocytes retrieved. This could help ensure more objective and quantitative prediction of ovarian response and counsel for patients requiring ovarian stimulation.

Methods

1. Study subjects

We selected 141 infertile patients who were undergoing the first IVF cycle at a fertility clinic of the Seoul National University Hospital (October 2009 to December 2013). This study was approved by the Institutional Review Board of Seoul National University Hospital. Clinical data of treatment cycles were collected from the medical database. Basal serum FSH level and random serum level of AMH were recorded if they were measured within 6 months prior to ovarian stimulation. AFC through transvaginal ultrasonography was recorded if it was measured on menstrual day 2 to 4 within 6 months prior to ovarian stimulation.

Patients with polycystic ovary syndrome (PCOS), in accordance with the diagnosis defined by the 2003 ASRM/ESHRE Rotterdam consensus, were excluded from this study. Other exclusion criteria were a history of surgical correction of ovarian endometriosis or current endometriosis, irregular menstruation, ovarian cysts >2 cm at the time of ovarian stimulation, a history of oral contraceptive use for 3 months before the start of ovarian stimulation, abnormal findings on thyroid function test or elevated prolactin level, and current medication for chronic diseases.

2. Blood test

Basal serum FSH level was measured on menstrual cycle day 2 to 3 using a Coat-A-Count FSH immunoradiometric assay procedure (Siemens Healthcare Diagnostics, Los Angeles, CA, USA). The sensitivity of the FSH measurement was 0.06 IU/L, inter-assay and the intra-assay coefficients of variation were 6.8% and 4.9%, respectively. AMH level was measured using a kit *via* enzyme-linked immunosorbent assay, which was measured by an Immunotech version (ref A11893, Beckman Coulter Inc., Marseilles, France) until September 2012 and Gen II (ref A79765, Beckman Coulter, Brea, CA, USA) starting in October 2012. Their sensitivity was 0.14 ng/mL and 0.08 ng/mL, inter-assay coefficients of variation were 5.6% and 4.5%, and intra-assay coefficients of variation were 5.4% and 3.6%, respectively.

3. Controlled ovarian stimulation and oocyte retrieval

Ovarian stimulation was performed using recombinant FSH (rFSH; Gonal-F, Merck-Serono, Geneva, Switzerland). The starting dose of rFSH varied from 75 IU to 450 IU according to five clinicians' judgment. Usually, the starting dose for women expected to be normal responders was 225 IU. For expected high responders, it was 150 IU and for expected poor responders, it was 300 IU. For expected extremely high and poor responders, the starting dose was 75–112.5 IU and 375–450 IU, respectively. The pituitary was suppressed using a flexible multiple-dose protocol of gonadotropin-releasing hormone (GnRH) antagonist (Cetrotide 0.25 mg/d, Merck-Serono) (n = 75) or a mid-luteal long protocol of GnRH agonist (Decapeptyl 0.1 mg/d, Ferring, Malmo, Sweden) (n = 66). Follicular growth was monitored from menstrual cycle day 7 *via* transvaginal ultrasonography. Dose adjustment was performed once at that time if necessary, and follicular growth was then observed every 1 to 2 days. When the mean diameter of a dominant follicle reached 18 mm in diameter or two follicles reached 17 mm in diameter, recombinant human chorionic gonadotropin (rhCG; Ovidrel 250 μg, Merck-Serono) was administered to trigger final follicular maturation. Oocyte retrieval was performed 35 to 36 hours after rhCG administration. Oocyte retrieval was implemented by experienced clinicians using a 17-gauge single lumen ovum aspiration needle (Cook Medical, Queensland, Australia), and all follicles with a diameter >10 mm were aspirated. MI and MII oocytes were counted as the total number of oocytes considered to have the potential for fertilization.

4. Statistical analysis

To predict the number of oocytes retrieved, a generalized linear model (GLM) including Poisson distribution and log link function was applied with the response variable of the number of oocytes and the prediction variables of age, basal FSH level, AMH level, and AFC. Cycles with GnRH antagonist and agonist long protocol were not sepa-

rated because a similar predictive performance of each ORT and their combinations was produced in both protocols. Since the AMH values obtained from the Gen I and Gen II version were almost same, they were analyzed together [13]. SAS ver. 9.2 (SAS Institute Inc., Cary, NC, USA) was used for all statistical analyses and the production of prediction equations and nomograms (R-3.0.3). Akaike information criterion (AIC) and Bayesian information criterion (BIC) were employed to compare goodness of fit among the prediction models. The predictive performance of each prediction model was evaluated *via* differences between the values observed by calibration plots and those predicted in the models. For discrimination, the linear relationship between the observed and the predicted values was assessed. The results were considered statistically significant at *p*-values of < 0.05.

Results

1. Clinical characteristics and univariate analysis

The statistical characteristics of woman's age, ORTs and stimulation outcomes from 141 IVF cycles are summarized in Table 1.

Univariate analysis revealed that the number of oocytes retrieved had a significant negative relationship with age (Spearman correlation coefficient $\rho = -0.433$, $p < 0.001$) and basal serum FSH level ($\rho = -0.404$, $p < 0.001$), but showed a significant positive relationship with serum AMH level ($\rho = 0.718$, $p < 0.001$) and AFC ($\rho = 0.757$, $p < 0.001$) (Figure 1). The number of oocytes retrieved was significantly different between cycles with high basal FSH level (> 10 IU/L) and normal FSH level (< 10 IU/L) (relative risk, 0.46; 95% confidence interval, 0.40–0.54, $p < 0.001$).

Table 1. Statistical characteristics of several ovarian reserve markers and stimulation outcomes from 141 IVF cycles

	Mean ± SD	Median	Range	Normal distribution
Age (yr)	36.3 ± 4.6	36	26–49	Yes
Basal serum FSH (IU/L)	8.8 ± 4.0	7.7	2.9–26.6	No
Serum AMH (ng/mL)	3.30 ± 3.28	2.42	0.08–17.25	No
Antral follicle count	12.2 ± 8.2	10	1–55	No
No. of oocytes retrieved	9.8 ± 7.7	8	0–36	No
Starting dose of rFSH (IU)	259.6 ± 93.6	225	75–450	No
Total dose of rFSH (IU)	2,328.2 ± 931.7	2,100	600–5,850	No

IVF, *in vitro* fertilization; SD, standard deviation; FSH, follicle-stimulating hormone; AMH, anti-Müllerian hormone; rFSH, recombinant human FSH.
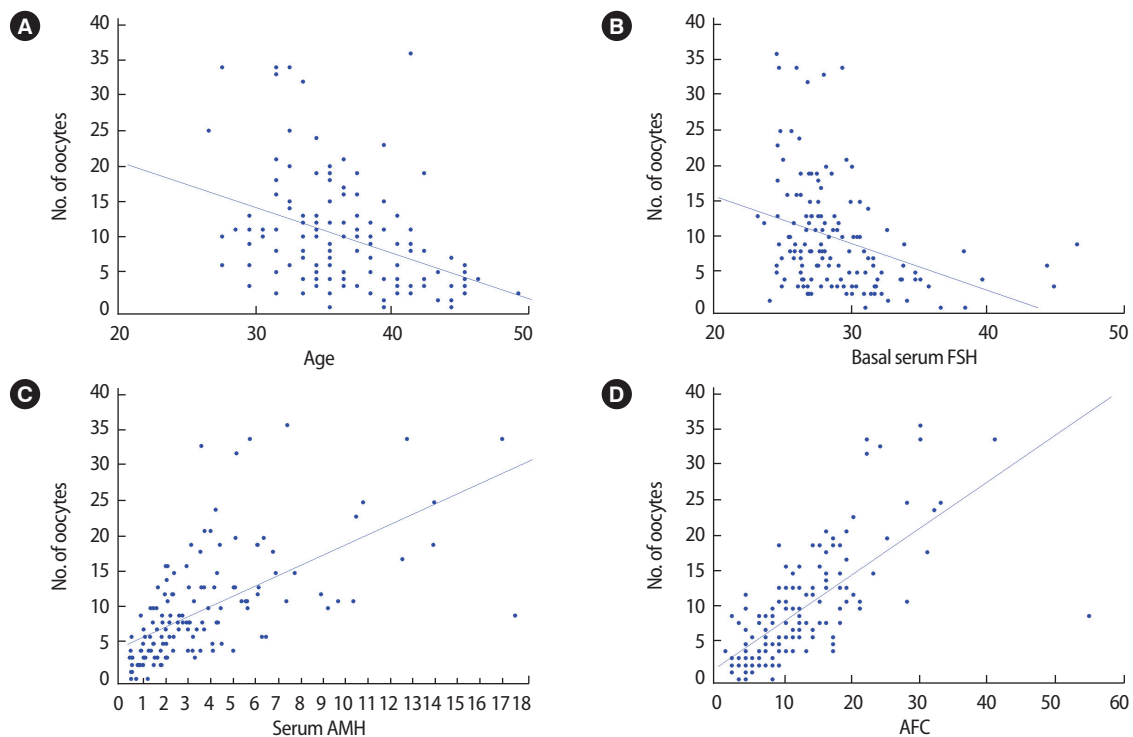

Figure 1. Correlation between the number of oocytes retrieved and women's age (A), basal serum follicle-stimulating hormone (FSH) level (B), serum anti-Müllerian hormone (AMH) level (C), and antral follicle count (AFC) (D).

KY Moon et al.   Nomographic estimate of oocyte retrieval

CERM

2. Multivariate analysis and prediction equations

Multivariate analysis was performed using significant predictive factors for the number of oocytes retrieved (age, basal serum FSH level, serum AMH level, and AFC) (Table 2). Four prediction models were constructed using several combinations of ORTs, and all the models showed statistical significance:

Model 1: Ln (number of oocytes retrieved) = 3.21−0.036 × (age) + 0.089 × (AMH)

Model 2: Ln (number of oocytes retrieved) = 3.422−0.03 × (age)−0.049 × (FSH) + 0.08 × (AMH)

Model 3: Ln (number of oocytes retrieved) = 2.32−0.017 × (age) + 0.039 × (AMH) + 0.03 × (AFC)

Model 4: Ln (number of oocytes retrieved) = 2.584−0.015 × (age)−0.035 × (FSH) + 0.038 × (AMH) + 0.026 × (AFC)

Table 2. Beta coefficient and 95% confidence interval by generalized linear model in each prediction model

	Model 1	Model 2	Model 3	Model 4
Intercept	3.210 (2.732, 3.689)[a]	3.422 (2.937, 3.908)[a]	2.320 (1.779, 2.862)[a]	2.584 (2.026, 3.141)[a]
Age	−0.036 (−0.049, −0.023)[a]	−0.030 (−0.043, −0.017)[a]	−0.017 (−0.031, −0.004)[b]	−0.015 (−0.029, −0.001)[b]
Basal serum FSH	-	−0.049 (−0.066, −0.031)[a]	-	−0.035 (−0.053, −0.017)[a]
Serum AMH	0.089 (0.076, 0.102)[a]	0.080 (0.066, 0.093)[a]	0.039 (0.020, 0.058)[a]	0.038 (0.019, 0.057)[a]
AFC	-	-	0.030 (0.022, 0.038)[a]	0.026 (0.018, 0.034)[a]

FSH, follicle-stimulating hormone; AMH, anti-Müllerian hormone; AFC, antral follicle count.
[a]$p < 0.001$; [b]$p < 0.05$.

Table 3. Comparison of performance among the prediction models for the number of oocytes

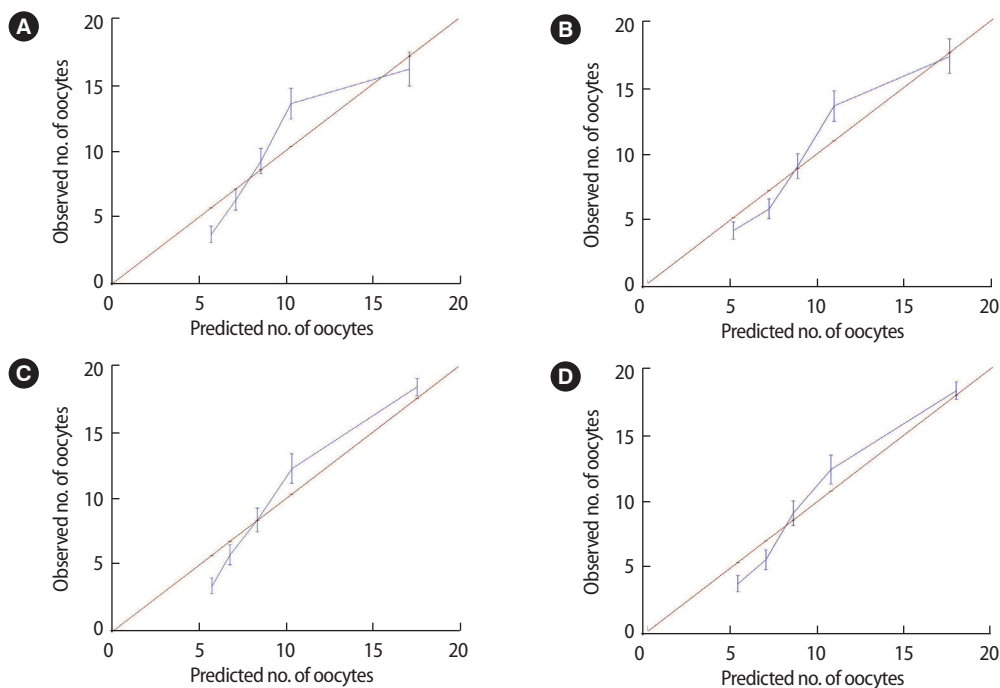|	Model 1	Model 2	Model 3	Model 4
Model fitness				
The Akaike information criterion	1,026.78	996.25	972.57	958.36
The Bayesian information criterion	1,035.62	1,008.05	984.37	973.11
Discrimination[a]	0.6822	0.6881	0.759	0.7543
Leave-one-out cross-validation[a]	0.6645	0.6704	0.7469	0.7082

[a]The Spearman correlation coefficient.

Figure 2. Calibration plots showing the difference between the predicted and observed number of oocytes according to model 1 (A), model 2 (B), model 3 (C), and model 4 (D).
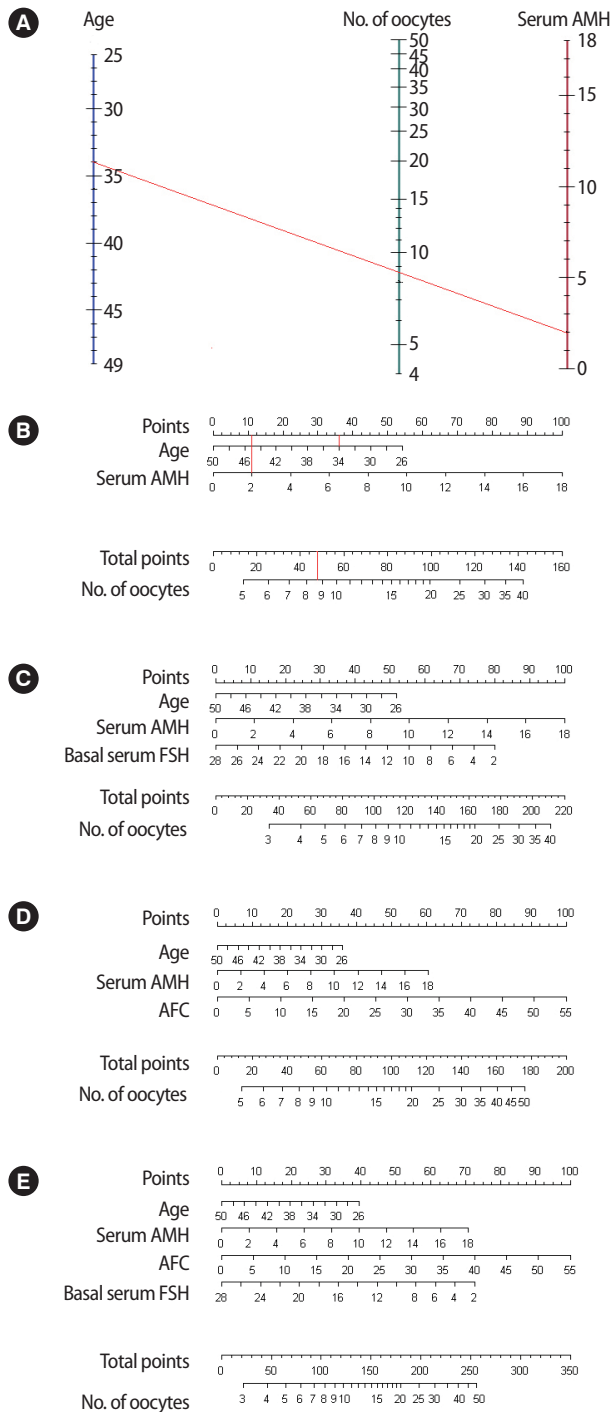
Figure 3. Nomograms to predict the number of oocytes retrieved using each prediction model: model 1 (A, B), model 2 (C), model 3 (D), and model 4 (E). For example, in a 34-year-old woman who has a serum AMH value of 2.0 ng/mL, approximately 8.7 oocytes could be expected by nomogram (A). The same result could be obtained by nomogram (B) (age, 34 years; gains score, 36.5; and AMH 2.0 gains score, 11.5; thus their sum score 48 corresponds to approximately 8.7 oocytes). AMH, anti-Müllerian hormone; FSH, follicle-stimulating hormone; AFC, antral follicle count.

On AIC and BIC analysis, model 4 showed the lowest values, indicating that this model had the best goodness of fit (Table 3). In the evaluation of discrimination, the Spearman correlation coefficient was the highest for model 3, followed by models 4, 2, and 1; however, models 3 and 4 showed a non-significant difference. Leave-one-out cross-validation performed as an internal validation also showed the same pattern of the order.

On a calibration plot of the predicted number of oocytes and the actual number of oocytes observed (Figure 2), models 3 and 4 showed a similar trend; that is, these models were the closest to the straight line that intercepts the origin (0, 0), indicating a perfect match between the observed and predicted number of oocytes.

According to the four prediction models, four nomograms were generated (Figure 3).

Discussion

This is the first study to establish nomograms that predict the final number of oocytes in stimulated IVF cycles by using several combinations including age, basal serum FSH level, serum AMH level, and AFC. Our developed nomograms may be clinically useful because the number of oocytes can be predicted in women with basal serum FSH or serum AMH level alone.

Among prediction models in the present study, model 4, which used four variables, showed the best ability to predict the number of oocytes. However, model 3 (leaving out basal serum FSH level) also showed superior results. This suggests a relatively weak predictive performance of basal serum FSH level to determine ovarian response to exogenous stimulation. This finding was in accordance with the results of Broer et al. [14]. They reported that predictive power for poor responders (number of oocytes < 4) was superior by using a combination of age, serum AMH level, and AFC, when compared to the use of one or two variables, but no marked difference was seen when four variables including FSH level were combined.

High basal serum FSH level usually predicts a poor response, but in this case, the serum AMH levels were helpful to determine outcome of stimulation. In one study, among women with elevated basal serum FSH levels, the subgroup with serum AMH levels ≥ 0.6 ng/mL yielded twice the number of oocytes compared with the subgroup with AMH levels < 0.6 ng/mL [15].

Previous studies have usually reported the predictive performance of each ORT or their combination for the prediction of poor responders; however, our study focused on the performance of several combinations of ORTs to predict the number of oocytes to quantitatively measure ovarian reserve.

As serum AMH measurement was introduced in clinics since 2009, previous studies have usually used age and basal serum FSH level to

KY Moon et al.   Nomographic estimate of oocyte retrieval

CERM

determine the starting dose of rFSH [16-18]. However, a recent study included serum AMH level, and nomograms were developed for determining the rFSH starting dose by using a combination of age, basal serum FSH level, and serum AMH level [19]. The nomograms were generated from the prediction model of the rFSH starting dose to retrieve nine oocytes; therefore, many high (women with polycystic ovary on ultrasonography) and poor responders (basal serum FSH level > 15 IU/L) were excluded.

We also excluded women with PCOS, but included high, normal, and poor responders in the data set. Therefore, our developed nomograms can be applied to both substantially high and poor responders.

A higher starting dose of gonadotropin would usually result in a greater number of oocytes in stimulated IVF cycles. However, in the present study, the number of oocytes retrieved was inversely proportional to the starting and total dose of rFSH. This appears to be mainly due to the inclusion of poor responders in the study; in poor responders, few oocytes were retrieved even after the administration of a high dose of gonadotropin. In the GLM model, the starting dose and the total dose of rFSH were excluded because of their low predictive performance. The number of oocytes retrieved might depend on the ORT or combination of ORTs used, rather than the starting or total dose of gonadotropin.

We included all cycles with the GnRH antagonist or agonist long protocol. Although the GnRH antagonist protocol was more frequently used in poor responders in the present study, a similar predictive performance of ORTs and their combinations was produced with both protocols.

In conclusion, the number of oocytes was accurately predicted by a combination of age, basal serum FSH level, serum AMH level, and AFC. From several prediction models, we developed nomograms to predict the number of oocytes retrieved. This would greatly improve the area of assisted conception. The nomogram needs further validation to improve individualized prediction.

Conflict of interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

References

1. Broekmans FJ, Kwee J, Hendriks DJ, Mol BW, Lambalk CB. A systematic review of tests predicting ovarian reserve and IVF outcome. Hum Reprod Update 2006;12:685-718.

2. Ebner T, Sommergruber M, Moser M, Shebl O, Schreier-Lechner E, Tews G. Basal level of anti-Mullerian hormone is associated with oocyte quality in stimulated cycles. Hum Reprod 2006;21:2022-6.

3. Ficicioglu C, Kutlu T, Baglam E, Bakacak Z. Early follicular antimullerian hormone as an indicator of ovarian reserve. Fertil Steril 2006;85:592-6.

4. Gnoth C, Schuring AN, Friol K, Tigges J, Mallmann P, Godehardt E. Relevance of anti-Mullerian hormone measurement in a routine IVF program. Hum Reprod 2008;23:1359-65.

5. Silberstein T, MacLaughlin DT, Shai I, Trimarchi JR, Lambert-Messerlian G, Seifer DB, et al. Mullerian inhibiting substance levels at the time of HCG administration in IVF cycles predict both ovarian reserve and embryo morphology. Hum Reprod 2006;21:159-63.

6. Nelson SM, Telfer EE, Anderson RA. The ageing ovary and uterus: new biological insights. Hum Reprod Update 2013;19:67-83.

7. Leader B, Hegde A, Baca Q, Stone K, Lannon B, Seifer DB, et al. High frequency of discordance between antimullerian hormone and follicle-stimulating hormone levels in serum from estradiol-confirmed days 2 to 4 of the menstrual cycle from 5,354 women in U.S. fertility centers. Fertil Steril 2012;98:1037-42.

8. Scheffer GJ, Broekmans FJ, Dorland M, Habbema JD, Looman CW, te Velde ER. Antral follicle counts by transvaginal ultrasonography are related to age in women with proven natural fertility. Fertil Steril 1999;72:845-51.

9. Frattarelli JL, Levi AJ, Miller BT, Segars JH. A prospective assessment of the predictive value of basal antral follicles in in vitro fertilization cycles. Fertil Steril 2003;80:350-5.

10. Hendriks DJ, Mol BW, Bancsi LF, Te Velde ER, Broekmans FJ. Antral follicle count in the prediction of poor ovarian response and pregnancy after in vitro fertilization: a meta-analysis and comparison with basal follicle-stimulating hormone level. Fertil Steril 2005;83:291-301.

11. Kupesic S, Kurjak A, Bjelos D, Vujisic S. Three-dimensional ultrasonographic ovarian measurements and in vitro fertilization outcome are related to age. Fertil Steril 2003;79:190-7.

12. McIlveen M, Skull JD, Ledger WL. Evaluation of the utility of multiple endocrine and ultrasound measures of ovarian reserve in the prediction of cycle cancellation in a high-risk IVF population. Hum Reprod 2007;22:778-85.

13. Kumar A, Kalra B, Patel A, McDavid L, Roudebush WE. Development of a second generation anti-Mullerian hormone (AMH) ELISA. J Immunol Methods 2010;362:51-9.

14. Broer SL, van Disseldorp J, Broeze KA, Dolleman M, Opmeer BC,

Bossuyt P, et al. Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: an individual patient data approach. Hum Reprod Update 2013;19:26-36.

15. Buyuk E, Seifer DB, Younger J, Grazi RV, Lieman H. Random anti-Mullerian hormone (AMH) is a predictor of ovarian response in women with elevated baseline early follicular follicle-stimulating hormone levels. Fertil Steril 2011;95:2369-72.

16. Howles CM, Saunders H, Alam V, Engrand P; FSH Treatment Guidelines Clinical Panel. Predictive factors and a corresponding treatment algorithm for controlled ovarian stimulation in patients treated with recombinant human follicle stimulating hormone (follitropin alfa) during assisted reproduction technology (ART) procedures: an analysis of 1378 patients. Curr Med Res Opin 2006;22:907-18.

17. Olivennes F, Howles CM, Borini A, Germond M, Trew G, Wikland M, et al. Individualizing FSH dose for assisted reproduction using a novel algorithm: the CONSORT study. Reprod Biomed Online 2009;18:195-204.

18. Popovic-Todorovic B, Loft A, Lindhard A, Bangsboll S, Andersson AM, Andersen AN. A prospective study of predictive factors of ovarian response in 'standard' IVF/ICSI patients treated with recombinant FSH: a suggestion for a recombinant FSH dosage normogram. Hum Reprod 2003;18:781-7.

19. La Marca A, Papaleo E, Grisendi V, Argento C, Giulini S, Volpe A. Development of a nomogram based on markers of ovarian reserve for the individualisation of the follicle-stimulating hormone starting dose in in vitro fertilisation cycles. BJOG 2012; 119:1171-9.