



OPEN ACCESS

EDITED BY

Tom Kelsey,
University of St Andrews, United Kingdom

REVIEWED BY

Natali Schachter Safrai,
Rabin Medical Center, Israel
Joao Cunha Filho,
Federal University of Rio Grande do Sul,
Brazil

*CORRESPONDENCE

Jie Qiao

✉ jie.qiao@263.net

Huiyu Xu

✉ huiyu56@bjmu.edu.cn

†These authors have contributed
equally to this work and share
first authorship

SPECIALTY SECTION

This article was submitted to
Reproduction,
a section of the journal
Frontiers in Endocrinology

RECEIVED 19 October 2022

ACCEPTED 05 January 2023

PUBLISHED 20 January 2023

CITATION

Ma C, Xu H, Wang H, Feng G, Han Y,
Alpadi K, Li R and Qiao J (2023) An online
tool for predicting ovarian responses in
unselected patients using dynamic inhibin
B and basal antimüllerian hormone levels.
Front. Endocrinol. 14:1074347.
doi: 10.3389/fendo.2023.1074347

COPYRIGHT

© 2023 Ma, Xu, Wang, Feng, Han, Alpadi, Li
and Qiao. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

An online tool for predicting ovarian responses in unselected patients using dynamic inhibin B and basal antimüllerian hormone levels

Congcong Ma^{1,2,3,4†}, Huiyu Xu^{1,2,3,4†*}, Haiyan Wang^{1,2,3,4†},
Guoshuang Feng⁵, Yong Han⁶, Kannan Alpadi⁷,
Rong Li^{1,2,3,4} and Jie Qiao^{1,2,3,4,8,9*}

¹Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing, China, ²National Clinical Research Center for Obstetrics and Gynecology (Peking University Third Hospital), Beijing, China, ³Key Laboratory of Assisted Reproduction (Peking University), Ministry of Education, Beijing, China, ⁴Beijing Key Laboratory of Reproductive Endocrinology and Assisted Reproductive Technology, Beijing, China, ⁵Big Data Center, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China, ⁶Hangzhou Qingguo Medical Technology Co. Ltd., Hangzhou, Zhejiang, China, ⁷The Predict Health, Houston, TX, United States, ⁸Beijing Advanced Innovation Center for Genomics, Beijing, China, ⁹Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China

Background: Reliable predictive models for predicting excessive and poor ovarian response in controlled ovarian stimulation (COS) is currently lacking. The dynamic (Δ) inhibin B, which refers to increment of inhibin B responding to exogenous gonadotropin, has been indicated as a potential predictor of ovarian response.

Objective: To establish mathematical models to predict ovarian response at the early phase of COS using Δ inhibin B and other biomarkers.

Materials and methods: Prospective cohort study in a tertiary teaching hospital, including 669 cycles underwent standard gonadotropin releasing hormone (GnRH) antagonist ovarian stimulation between April 2020 and September 2020. Early Δ inhibin B was defined as an increment in inhibin B from menstrual day 2 to day 6 through to the day of COS. Least Absolute Shrinkage and Selection Operator (LASSO) logistic regression with 5-fold cross-validation was applied to construct ovarian response prediction models. The area under the receiver operating characteristic curve (AUC), prevalence, sensitivity, and specificity were used for evaluating model performance.

Results: Early Δ inhibin B and basal antimüllerian hormone (AMH) levels were the best measures in building models for predicting ovarian hypo- or hyper-responses, with AUCs and ranges of 0.948 (0.887–0.976) and 0.904 (0.836–0.945) in the validation set, respectively. The contribution of the early Δ inhibin B was 67.7% in the poor response prediction model and 56.4% in the excessive response prediction model. The basal AMH level contributed 16.0% in the poor response prediction model and 25.0% in the excessive response prediction model. An online website-

based tool (<http://121.43.113.123:8001/>) has been developed to make these complex algorithms available in clinical practice.

Conclusion: Early Δ inhibin B might be a novel biomarker for predicting ovarian response in IVF cycles. Limiting the two prediction models to the high and the very-low risk groups would achieve satisfactory performances and clinical significance. These novel models might help in counseling patients on their estimated ovarian response and reduce iatrogenic poor or excessive ovarian responses.

KEYWORDS

inhibin B, antimüllerian hormone, ovarian response, controlled ovarian stimulation, predictive model

Introduction

Ovarian response, defined as the number of oocytes retrieved during controlled ovarian stimulation (COS), is a keystone for success of *in vitro* fertilization (IVF) (1, 2). Excessive response during ovarian stimulation indicates the potential risk of ovarian hyperstimulation syndrome (OHSS), which is a serious iatrogenic complication of COS (3, 4). Predicting excessive ovarian response to early identify those hyper-responders would allow early interventions thus minimizing potential complications. Likewise, failure to respond adequately with few oocytes retrieved during COS is called poor ovarian response (POR), resulting in a high risk of cycle cancellation and a significantly diminished probability of pregnancy (1). Fortunately, there are some options, such as dose adjustment of follicle-stimulating hormone (FSH), for an identified abnormal ovarian response in the early stage of COS. Therefore, an accurate predictive system to predict excessive or poor ovarian response in the IVF cycle is critical to both optimize outcomes and reduce complications.

Several factors are widely used to help clinicians early identify ovarian response, including age, antral follicle counts (AFC), AMH, basal FSH levels, basal estradiol levels and luteinizing hormone (LH)/FSH ratio (5–14). Although evidence indicates that AMH and AFC may be superior to others in predicting ovarian responses (10, 13, 15–17), a reliable predictive model in all clients is currently lacking. The above facts promote the need for the establishment of sophisticated models using multiple factors to predict ovarian response.

Serum inhibin B has been advocated for decades as one of the potential novel biomarkers in predicting ovarian response. It is produced by the cohort of developing preantral and early antral follicles under the stimulation of FSH, and it peaks at the mid-follicular phases (18, 19). Early follicular inhibin B levels decrease during reproductive aging, leading to increased FSH concentrations (20). Although the absolute value of inhibin B at the start of COS remains controversial to reflect ovarian reserve, the dynamic (Δ) inhibin B, which refers to increment of inhibin B responding to

exogenous gonadotropin, has been indicated as a better predictor of ovarian response than AFC and AMH (21, 22).

To take early measures to reduce iatrogenic POR thus reducing the unnecessary ovarian puncture, as well as to reduce the excessive ovarian response thus reducing the incidence of OHSS, we aimed to establish multivariate mathematical models to predict excessive and poor ovarian response at the early phase of COS, using early Δ inhibin B (defined as an increase in the inhibin B level on menstrual day 6 from menstrual day 2) and other biomarkers.

Materials and methods

Subjects

This was a prospective observational cohort study involving a continuous series of women seeking IVF from April 2020 to September 2020 in Peking University Third Hospital. Patients undergoing standard GnRH antagonist ovarian stimulation were included. There were no exclusion criteria for the participant recruitment. Participants all provided written informed consent. Approval from the Human Reproductive Ethics Committee of Peking University Third Hospital was obtained (2015sz-017).

Controlled ovarian stimulation

The COS protocol was as described (5, 6). Briefly, recombinant human FSH (rFSH) treatment was initiated on menstrual day 2. On menstrual day 6, the follicle growth was monitored by ultrasound and the dosage of rFSH was adjusted according to follicle growth and E_2 levels. The GnRH antagonist treatment was initiated on stimulation day 5–7 when the growing follicles were 10–12 mm in diameter. When at least 2 dominant follicles (diameter ≥ 18 mm) were observed by ultrasound, 5000–10000 IU hCG was administered to trigger final

oocyte maturation. Thirty-six hours after hCG administration, oocyte retrieval was performed. Then, embryos were transferred to the mother after culture or cryopreserved for future use.

Endocrine assays

On menstrual days 2 and day 6 of the COS treatment cycle, intravenous blood was collected for measuring LH, progesterone, estradiol (E_2), FSH, testosterone (TES), androstenedione (AND), AMH and inhibin B concentrations. Blood samples were collected, and the collection tubes immediately inverted five times and centrifuged (1800g, 10min) for further endocrine assessments.

Serum levels of FSH, LH, E_2 , progesterone, TES, and AND were tested using a Siemens Immulite 2000 immunoassay system (Siemens Healthcare Diagnostics, Shanghai, P. R. China). Quality control samples were supplied by Bio-RAD Laboratories (Hercules, CA, USA; Lyphochek Immunoassay Plus Control, Trilevel, catalog number 370, lot number 40370). Serum AMH and inhibin B concentrations were measured using an ultrasensitive two-site ELISA (Ansh Labs LLC; Webster, TX, USA), using quality controls supplied with the kit. For the Trilevel controls, the coefficients of variation for the assays were <6% for AMH, inhibin B, FSH and LH, and <10% for E_2 , progesterone, AND, and TES.

Clinical outcomes

The outcome variables were POR or excessive ovarian response. POR was defined as <5 oocytes retrieved, and the excessive ovarian response was defined as >20 oocytes retrieved.

Development of predictive models

Sixteen potential candidate variables were considered for inclusion in poor and excessive ovarian response prediction models, including age, basal hormone levels on menstrual day 2 (inhibin B, AMH, LH, FSH, E_2 , progesterone, TES and AND) and changed hormone levels on menstrual day 6 from day 2 (Δ inhibin B, Δ AMH, Δ LH, ΔE_2 , Δ progesterone, Δ TES and Δ AND). Prior to modeling, we first explored the distribution of the independent variables. Because the distribution of Δ inhibin B, basal AMH, and ΔE_2 were obviously skewed, logarithmic transformation of these three variables was performed. The Bayesian Information Criterion (BIC) value was calculated through univariate logistic regression to evaluate whether the transformed data were better than the original data for constructing the prediction models.

Seventy percent of the data were selected randomly as the training set, which was used for model establishment, and the rest 30% of the

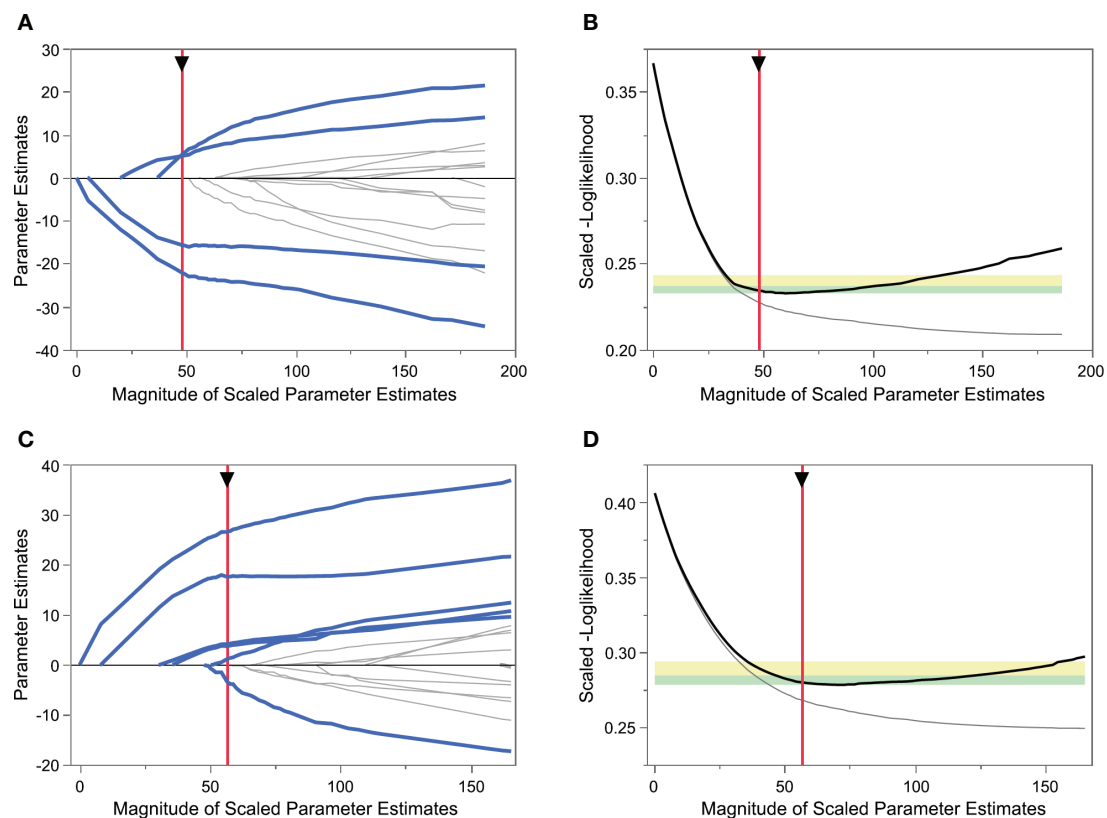


FIGURE 1

Variable selection using the least absolute shrinkage and selection operator (LASSO) logistic regression model. (A) LASSO model showing coefficient profiles of the candidate variables for predicting poor ovarian response (POR). Each trace represents a regression parameter, and the penalty decreases from left to right. (B) Tuning parameter selection of POR prediction model by 5-fold cross-validation. The vertical red axis indicates the optimal values on the basis of the minimum scaled(-log(likelihood)). (C) LASSO model coefficient profiles of the candidate variables for excessive ovarian response prediction. (D) Tuning parameter selection of excessive ovarian response prediction model by 5-fold cross-validation.

data were used as the validation set, which was used for model evaluation. For variable selection, the sixteen variables were entered into the selection process. There is a strong correlation between the variables, as shown in Supplemental Figure 1. To minimize the potential collinearity of variables and overfitting, Least Absolute Shrinkage and Selection Operator (LASSO) regression was used to filter the variables. The optimal model was established *via* 5-fold cross-validation. The criterion for variable selection was the scaled negative log-likelihood in the training set, which considered the one with the smallest scaled negative log-likelihood as the best. The number of predictors selected for constructing the final model was also adjusted to balance both the accuracy and simplicity of the model. The main effect of each included variable reflects the relative contribution of the variable alone, and the total effect represents the relative contribution of that variable both alone and in combination with other variables.

Evaluation of model performance

The performance of our model was assessed by calculating the area under the receiver-operator characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). When the model was evaluated as being satisfactory, we aimed to develop an easy-to-use clinician-friendly online tool to stratify patients as having a certain risk of poor or excessive ovarian response.

Statistical analysis

Normally distributed variables are shown as the mean and standard deviation, while variables not normally distributed are presented as the median and interquartile range (IQR). An

TABLE 1 Clinical characteristics of 642 IVF cycles of standard GnRH antagonist ovarian stimulation.

Variable		
Age (y), mean \pm SD		32.77 \pm 4.32
BMI (kg/m ²), mean \pm SD		22.48 \pm 3.42
Main cause of infertility, n (%)		
Male factors		190 (29.60%)
Tubal factor		173 (26.95%)
PCOS		105 (16.36%)
Uterine cavity factors		53 (8.26%)
Endometriosis		41 (6.39%)
Unexplained cause		80 (12.46%)
Serum hormones	Basal levels	Δ levels
AMH (ng/mL), median (IQR)	3.00 (1.60, 5.30)	-0.51 (-1.22, -0.16)
FSH (IU/L), median (IQR)	6.30 (5.20, 7.90)	-
LH, median (IQR)	3.40 (2.40, 4.80)	-1.97 (-3.01, -1.15)
E ₂ (pmol/L), median (IQR)	152.00 (122.00, 176.00)	1126.00 (484.00, 2127.75)
Progesterone, median (IQR)	1.20 (1.00, 1.40)	-0.22 (-0.49, 0.00)
TES, median (IQR)	0.70 (0.70, 0.80)	0.00 (0.00, 0.15)
AND, median (IQR)	6.60 (5.00, 9.20)	0.89 (-0.77, 2.92)
Inhibin B, median (IQR)	88.50 (63.40, 114.23)	649.14 (325.75, 1187.69)
Peak E ₂ (pg/ml), median (IQR)	2480.00 (1577.00, 3857.00)	
Total rFSH dose (IU), median (IQR)	2325 (1725, 3150)	
Obtained oocytes, median (IQR)	11.00 (7.00, 17.00)	
Ovarian response, n (%)		
Normal ovarian response	467 (72.74%)	
Poor ovarian response	83 (12.93%)	
Excessive ovarian response	92 (14.33%)	

BMI, body mass index; PCOS, polycystic ovary syndrome; AMH, antimüllerian hormone; FSH, follicle stimulating hormone; LH, luteinizing hormone; E₂, estradiol; TES, testosterone; AND, androstenedione; rFSH, recombinant follicle stimulating hormone.
-, not available; Δ , dynamic.

independent samples t-test or non-parametric test was used to analyze continuous variables. All the analyses in this study were performed using SAS JMP Pro (version 14.2; SAS Institute, Cary, NC, USA), and $p < 0.05$ was considered statistically significant.

Results

Patient demographics

We collected 669 standard GnRH antagonist COS cycles prospectively without data selection. After excluding 27 cycles in which the numbers of oocytes retrieved were not recorded, 642 cycles were finally included for analysis. Participants had a mean age of 32.8 years (range 22-45 years) in this study. The basic characteristics of the cycle data are shown in [Table 1](#). A total of 467 patients (72.7%) were normal responders and the median number of obtained oocytes was 11 (IQR 8-16). Ninety-two patients (14.3%) were diagnosed as excessive responders and the median number of obtained oocytes was 25 (IQR 22-28). The remaining 83 patients (12.9%) were diagnosed as poor responders and the median number of obtained oocytes was 2 (IQR 2-3). In the poor responder group, three patients had failed oocyte retrieval.

Establishment of ovarian response prediction models

After logarithmic transformation of the Δ inhibin B, basal AMH and ΔE_2 measures, the BIC of each predictor before and after transformation was decreased significantly, from 387.90 to 332.65 for Δ inhibin B, from 401.31 to 351.28 for basal AMH and from 445.31 to 424.74 for ΔE_2 , indicating the advantages of data transformation.

The process of variable selection using LASSO and fivefold cross-validation is shown in [Figure 1](#). When predicting POR, the normal responders and excessive responders were grouped as non-poor ovarian responders. As shown in [Figures 1A, B](#), the scaled $[-\log(\text{likelihood})]$ value in the validation set was smallest when four variables were included, thus including more variables did not induce a decrease of the scaled $[-\log(\text{likelihood})]$. Based on Occam's razor principle (23), four variables were finally included in the POR prediction model (Model 1), namely, $\log[\Delta$ inhibin B], $\log[\text{basal AMH}]$, basal FSH, and Δ TES.

Similarly, when predicting excessive ovarian response, the normal responders and poor responders were grouped as non-excessive ovarian responders. Using the same method, six variables including $\log[\Delta$ inhibin B], $\log[\text{basal AMH}]$, basal FSH, Δ AND, basal AND, and basal LH were selected into the excessive ovarian response prediction model (Model 2) ([Figures 1C, D](#)).

The main and total effects of each variable in Model 1 predicting POR and Model 2 predicting excessive ovarian response were shown in [Supplemental Table 1](#). In both models, the most important variable was $\log[\Delta$ inhibin B], with a main effect of 67.7% and total effect of 77.7% in Model 1, and with main effect of 56.6% and total effect of 69.2% in Model 2. $\log[\text{basal AMH}]$ was the second important

variable in both two models, with main effect of 16.0% and total effect of 25.0% in Model 1, and with main effect of 22.1% and total effect of 33.3% in Model 2. In order to find a simpler model for predicting excessive ovarian response, we tried to build another model using only these two most important predictors (Model 2s). The main and total effects of $\log[\Delta$ inhibin B] in Model 2s were 56.4% and 75.0%, and of $\log[\text{basal AMH}]$ 25.0% and 43.5%, respectively.

Performances of established prediction models

The performances of Models 1, 2, and 2s are shown in [Supplemental table 2](#). The AUC values of Model 1, Model 2 and Model 2s in the training set were 0.910, 0.896 and 0.875; in the validation set they were 0.948, 0.882 and 0.904, respectively. The sensitivity and specificity of Model 1 were 0.456 and 0.978, of Model 2 0.431 and 0.980 and of Model 2s 0.389 and 0.984, respectively. The PPV and NPV of Model 1 were 0.756 and 0.922, of Model 2 0.775 and 0.914 and of Model 2s 0.800 and 0.908, respectively. The ROC curves of Model 1, Model 2 and Model 2s are shown in [Figures 2A-C](#), respectively. The AUC values of Model 2 and Model 2s were compared using $1000 \times$ bootstrap sampling in the validation set. The distributions of AUCs of Models 2 and 2s are shown in [Figure 2D](#). There were little differences in the AUCs between Models 2 and 2s. Considering its simplicity, we recommend using Model 2s for predicting excessive ovarian responses.

Interpretation of prediction models and development of the website-based tool

To explore the relationship between the predicted probability and the incidence of poor or excessive ovarian response, we constructed a histogram to compare them ([Figure 3](#)). As shown in [Figure 3A](#), the predicted probability of Model 1 was classified into eight groups. The actual incidence of POR continuously increased with the increase in predicted probability. When the predicted probability was $<5\%$, the actual incidence of POR was zero. And when the predicted probability was $\geq 70\%$, the actual incidence of POR was over 90%. Similarly, as shown in [Figure 3B](#), the actual incidence of excessive ovarian response gradually increased with the increase in predicted probability. When predicted probability was $<5\%$, the actual incidence of excessive ovarian response was 1%. The incidence of excessive ovarian response was 86% when the predicted probability was $\geq 70\%$.

The risk groups of poor and excessive ovarian response were classified based on the incidence and predicted probability in all participants enrolled ([Figures 3C, D](#)). When limited to the high and the very-low risk groups, both models achieved satisfactory performances. To be practical, the algorithms of Model 1 predicting POR and Model 2s predicting excessive ovarian responses were developed into a website-based tool (<http://121.43.113.123:8001/>). When the user inputs the value of required indicators and clicks 'calculate', the predicted probability and risk group of poor or excessive ovarian response in a certain patient are displayed.

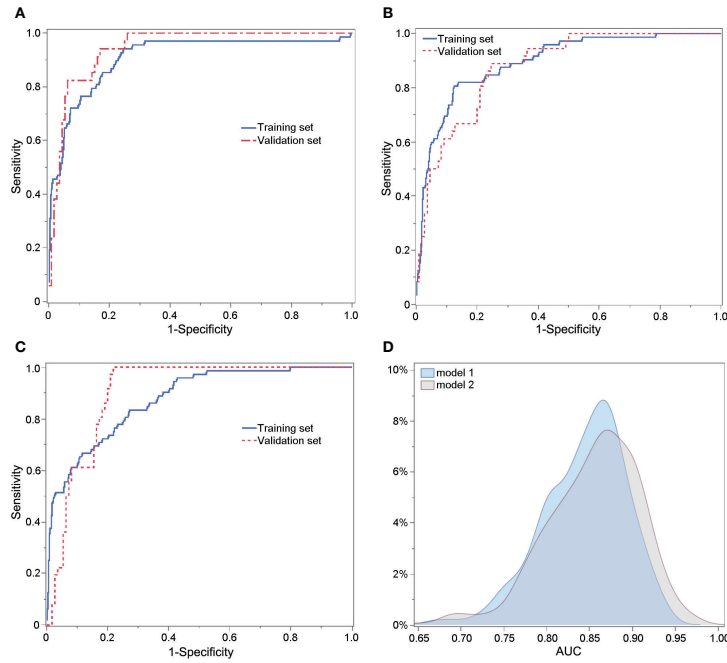


FIGURE 2 ROC curves of Model 1 predicting poor ovarian response (A), Model 2 (B) and Model 2s (C) for predicting excessive ovarian response. (D) Comparison of the AUCs of Models 2 and 2s using 1000 x bootstrap sampling in the validation set.

Discussion

The basal level of serum AMH is frequently used to predict ovarian response and guide the starting dose of rFSH for COS in assisted reproductive technology (ART) clinics. However, basal assays tend to reflect the size of the resting ovarian follicle pool (24, 25). In contrast, dynamic assays of hormonal parameters during ovarian

stimulation might better reflect the number of rFSH-sensitive follicles. Thus, dynamic changes (Δ) in hormone levels could be used to guide the adjustment of the rFSH dosage during COS. Our results show that the early Δ inhibin B combined with basal AMH levels can better predict ovarian response. Using these indicators, we have established models for POR (Model 1) and excessive ovarian response (Model 2s), respectively. The contribution of the early Δ inhibin B is 67.7% in

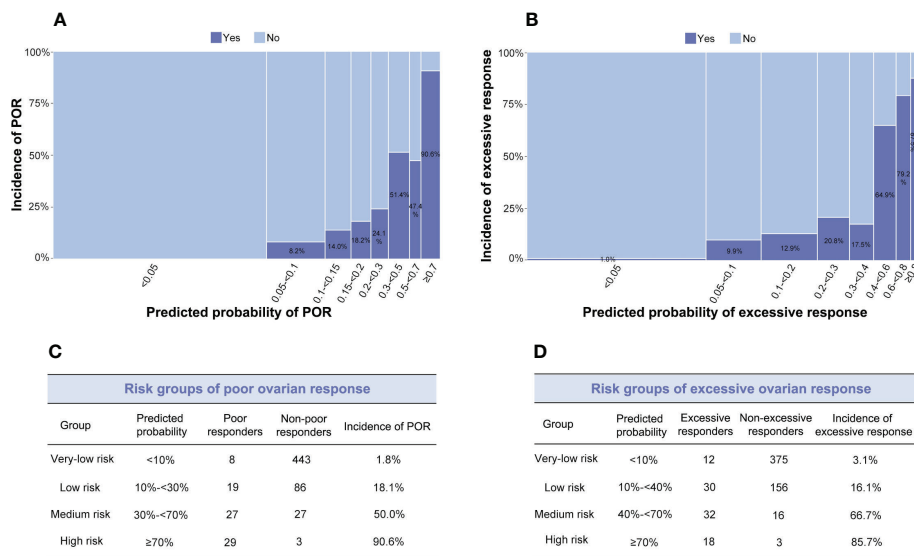


FIGURE 3 (A, B) Distribution between observed and predicted incidences of poor ovarian response (POR) (Model 1) and excessive ovarian response (Model 2s). The ordinate signifies the incidence, and the abscissa indicates the predicted probability; the wider the abscissa, the larger the sample size in each sub-group. (A) Light blue indicates the POR group; dark blue signifies the non-POR group. (B) Light blue indicates the excessive ovarian response group; dark blue signifies the non-excessive ovarian response group. (C) Risk groups of POR. (D) Risk groups of excessive ovarian response.

Model 1 and 56.4% in Model 2s, while the contribution of the basal AMH level is 16.0% in Model 1 and 25.0% in Model 2s. We have built an online website according to these models to facilitate in counseling patients on their estimated ovarian response and to adjust the dose of rFSH during COS protocols, thus reducing iatrogenic poor or excessive responses.

Predicting ovarian response is not easy clinically. Theoretically, for use in everyday ART practice, a simple approach would have an advantage over any complicated or unwieldy system. Ovarian reserve markers, including basal AMH, FSH and AFC reflect the size of primordial follicle pool. Generally, the better the ovarian reserve, the better the ovarian response. However, although ovarian reserve is related to the ovarian response to external stimuli, it is still unclear whether there is a linear relationship between ovarian response and ovarian reserve. In addition, there is heterogeneity between individuals with the same ovarian reserve, probably because of inter-individual variability such as the number of FSH receptors expressed. Therefore, the ovarian response to external gonadotropins tends to be a complex biological process. Actually, Δ inhibin B reflects FSH-dependent follicular growth, thus, addition of early Δ inhibin B greatly improves the performance of the model for predicting poor ovarian response, from AUC of around 0.838–0.862 (5) to 0.887–0.976. Therefore, combination of early Δ inhibin B and ovarian reserve markers may better predict ovarian responses and may even have the potential to replace the ultrasound tests during ovarian stimulation.

Based on the above, we established two models for predicting poor and excessive ovarian response using both basal and dynamic biomarkers. Although age was one of the potential parameters considered for inclusion in the prediction models for both poor and excessive response, it was not included finally. Explanations may as follow (1): Age is one of the most common used indicators of ovarian reserve. As mentioned above, better ovarian reserve usually leads to better ovarian response. However, a strong inter-individual variability for ovarian reserve exists within the same chronological age group (2). In both poor and excessive response prediction models, basal AMH was included finally. There are strong correlations between age and AMH, the role of age may be replaced by basal AMH, which suggests that AMH may be a better predictor than age for ovarian reserve. The AUCs and specificities of the two models were satisfactory. However, the sensitivities of Models 1 and 2s were low at 41.2% and 50.0%, respectively. The main reason for this is that both sensitivity and specificity are calculated according to a cut-off point, which divides the outcome into binary groups, i.e., poor/non-poor response or excessive/non-excessive response. However, as shown in Figure 3, we further grouped the cohort into subgroups according to the distribution between predicted probability and the incidence of poor or excessive ovarian response, thereby avoiding the misclassification of poor/non-poor response or excessive/non-excessive response predicted by a probability of 50%. Thus, our models can be of better clinical significance.

The role of inhibin B levels in predicting ovarian response is unclear (26–33). Previous studies indicated that inhibin B does not have a better predictive effect than other indicators in predicting ovarian response, mainly because of its variation between individuals

and fluctuations within menstrual cycles (27, 29, 32). However, dynamic levels in an ART cycle could minimize these influences. Peñarrubia et al. found that day 5 inhibin B levels were the best predictor of POR in ART cycles stimulated with gonadotropins after pituitary suppression, suggesting that the exogenous FSH-stimulated inhibin B level is better than baseline determinations (34). Li et al. reported the correlation of serum inhibin B levels on basal, day 5, the day of giving hCG and the day of oocytes retrieval, respectively, in women with diminished ovarian response (35). They found that inhibin B levels on day 5 and day-hCG had more predictive value than that on day 2 or day 3. However, they did not investigate the predictive value of Δ inhibin B.

Decanter et al. used only Δ inhibin B, defined as day 8 inhibin B minus day 6 inhibin B, to predict POR. They discovered that a minimal increase of 300 pg/mL is needed to rule out a POR, with the sensitivity of 70% and specificity of 94% (30). Similarly, we defined the early Δ inhibin B value as serum inhibin levels of day 6 minus day 2. Compared with basal levels of AMH, FSH and inhibin B, early Δ inhibin B was the most effective in predicting both excessive and poor responses. Instead of identifying a cut-off value of Δ inhibin B, first, we established mathematical models combining basal AMH and Δ inhibin B and other variables to predict both poor and excessive ovarian responses, with AUCs and ranges of 0.948 (0.887–0.976) in the validation data of Model 1 and 0.904 (0.836–0.945) in the validation data of Model 2s. Second, in the modeling process, we applied cross-validation to identify the ideal model, indicating more stable of our models. Third, we also applied LASSO regression to minimize the potential collinearity and over-fitting of variables. In addition, we have made the algorithms into software to make the complex algorithms available in ART practice. In addition, the Δ inhibin B in our study is earlier (day 6 minus day 2) than the previous study (day 8 minus day 6). From the perspective of clinical significance, our earlier Δ inhibin B measured enables the clinician to adjust gonadotropin doses better.

In addition to the main predictors (Δ inhibin B and basal AMH), we also found that Δ TES and basal FSH levels contributed to predicting POR. It is reasonable that high basal FSH together with low basal AMH can better reflect poor ovarian reserve, which generally results in POR. As for the variable Δ TES, a previous study demonstrated that the serum level of TES increases slightly during COS (36), which was also confirmed here. The Δ TES was selected into the prediction model of POR and its parameter estimation is positive, which means that a high value of Δ TES would predict a high risk of POR. The increased Δ TES in POR might be due to the blocking conversion of testosterone into estradiol by reduced FSH sensitivity (37). However, the contribution of Δ TES to the POR model is relatively low (3.2%), and its predictive value needs to be validated in the future.

The stimulated E_2 level during COS is often used to predict excessive ovarian response or OHSS to take appropriate management and to minimize potential complications (38, 39). However, in this study, we discovered that the early ΔE_2 is not of significance compared with early Δ inhibin B and basal AMH in predicting excessive ovarian response. The underlying mechanism may be that there is a strong correlation between early Δ inhibin B and early ΔE_2 (Pearson's $r=0.74$ in our study, data not shown), the role of early ΔE_2

could be replaced by early Δ inhibin B, which suggests that early dynamic changes in inhibin B may be a better predictor than E_2 for adjusting rFSH dosing during COS.

Although it needs to be validated, the biological rationale of models established in this study is that basal AMH (with or without basal FSH) and Δ inhibin B (with or without Δ TES) could be complementary: the former for reflecting ovarian reserve and the latter for indicating the ovarian sensitivity in a COS treatment cycle. We aim to perform further basic research to determine the mechanism of Δ inhibin B in predicting ovarian response.

The strength of this study is that we established ovarian response models based on a relatively large sample size. The original data in this study are reliable because of its prospective nature. Notably, we included all patients who underwent the most common IVF treatment without selection, thus our models can be applied to most patients.

The main limitation of this study is that we included patients in a single center and further internal and external validation of the models need to be performed. Another limitation is the unstandardized inhibin B assay kits. The variations in inhibin B values when tested by different kits would limit the applicability of our model. Standardization of inhibin B measures from different test kits might solve this problem in the future to better realize the clinical application of our models.

Conclusions

In conclusion, the early Δ inhibin B levels during COS proved very informative in predicting ovarian response. We have established sophisticated models to predict excessive and poor ovarian response using novel and traditional biomarkers. Limiting the two prediction models to the high and the very-low risk groups would achieve satisfactory performances and clinical significance. The online tools based on the models we established have good potential to facilitate ART clinicians in counseling patients on their estimated ovarian response and make better decisions on treatment adjustment or cycle cancellation. Today, faster, more sensitive, and repeatable inhibin B kits based on chemiluminescence methods are available, making it possible to use this measure as a valuable additional predictor of ovarian response during COS in the future.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Approval from the Human Reproductive Ethics Committee of Peking University Third Hospital was obtained (2015sz-017). The patients/participants provided their written informed consent to participate in this study.

Author contributions

Conceptualization: JQ, HX; Cohort enrollment and clinical data collection: JQ, RL, HW; Methodology: HX, CM, GF, YH, KA; Study organization assistance: RL; Writing original draft: CM, HX, HW; Writing review and editing: all authors; Supervision: JQ. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China for Distinguished Young Scholars (Grant No. 81925013); the National Key Research and Development Program of China (Grant No. 2018YFC1002104, 2018YFC1002106); the Innovation & Transfer Fund of Peking University Third Hospital (Grant No. BYSYZHZB2020102); the Major National R&D Projects of China (Grant No. 2017ZX09304012-012); National Natural Science Foundation of China (Grant No. 81771650); the Capital Health Research and Development of Special Project (Grant No. 2018-1-4091); China Postdoctoral Science Foundation (2022M710261) and China Postdoctoral Innovation Talent Support Program (BX20220020).

Acknowledgments

We thank James Cummins, PhD, from Liwen Bianji, Edanz Editing China (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript.

Conflict of interest

YH was employed by Hangzhou Qingguo Medical Technology Co. Ltd. and KA is employed by The Predict Health.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2023.1074347/full#supplementary-material>

References

1. Abu-Musa A, Haahr T, Humaidan P. Novel physiology and definition of poor ovarian response; clinical recommendations. *Int J Mol Sci* (2020) 21(6):2110. doi: 10.3390/ijms21062110
2. Sunkara SK, Rittenberg V, Raine-Fenning N, Bhattacharya S, Zamora J, Coomarasamy A. Association between the number of eggs and live birth in ivf treatment: An analysis of 400 135 treatment cycles. *Hum Reprod* (2011) 26(7):1768–74. doi: 10.1093/humrep/der106
3. Fernandez-Sanchez M, Visnova H, Yuzpe A, Klein BM, Mannaerts B, Arce JC, et al. Individualization of the starting dose of follitropin delta reduces the overall ohss risk and/or the need for additional preventive interventions: Cumulative data over three stimulation cycles. *Reprod BioMed Online* (2019) 38(4):528–37. doi: 10.1016/j.rbmo.2018.12.032
4. Smith V, Osianlis T, Vollenhoven B. Prevention of ovarian hyperstimulation syndrome: A review. *Obstet Gynecol Int* (2015) 2015:514159. doi: 10.1155/2015/514159
5. Xu H, Shi L, Feng G, Xiao Z, Chen L, Li R, et al. An ovarian reserve assessment model based on anti-mullerian hormone levels, follicle-stimulating hormone levels, and age: Retrospective cohort study. *J Med Internet Res* (2020) 22(9):e19096. doi: 10.2196/19096
6. Xu H, Feng G, Wang H, Han Y, Yang R, Song Y, et al. A novel mathematical model of true ovarian reserve assessment based on predicted probability of poor ovarian response: A retrospective cohort study. *J Assist Reprod Genet* (2020) 37(4):963–72. doi: 10.1007/s10815-020-01700-1
7. Xu H, Zhang M, Zhang H, Alpadi K, Wang L, Li R, et al. Clinical applications of serum anti-mullerian hormone measurements in both males and females: An update. *Innovation (N Y)* (2021) 2(1):100091. doi: 10.1016/j.xinn.2021.100091
8. Xu H, Zeng L, Yang R, Feng Y, Li R, Qiao J. Retrospective cohort study: Amh is the best ovarian reserve markers in predicting ovarian response but has unfavorable value in predicting clinical pregnancy in grh antagonist protocol. *Arch gynecol obstetrics* (2017) 295(3):763–70. doi: 10.1007/s00404-016-4274-8
9. Qiao J, Wang ZB, Feng HL, Miao YL, Wang Q, Yu Y, et al. The root of reduced fertility in aged women and possible therapeutic options: Current status and future prospects. *Mol Aspects Med* (2014) 38:54–85. doi: 10.1016/j.mam.2013.06.001
10. Broer SL, Dolleman M, Opmeer BC, Fauser BC, Mol BW, Broekmans FJ. Amh and afc as predictors of excessive response in controlled ovarian hyperstimulation: A meta-analysis. *Hum Reprod Update* (2011) 17(1):46–54. doi: 10.1093/humupd/dmq034
11. Broer SL, Mol BW, Hendriks D, Broekmans FJ. The role of antimullerian hormone in prediction of outcome after ivf: Comparison with the antral follicle count. *Fertil Steril* (2009) 91(3):705–14. doi: 10.1016/j.fertnstert.2007.12.013
12. 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(4):979–87. doi: 10.1016/j.fertnstert.2004.11.029
13. Ng EH, Tang OS, Ho PC. The significance of the number of antral follicles prior to stimulation in predicting ovarian responses in an ivf programme. *Hum Reprod* (2000) 15(9):1937–42. doi: 10.1093/humrep/15.9.1937
14. Lenton EA, Sexton L, Lee S, Cooke ID. Progressive changes in lh and fsh and lh: Fsh ratio in women throughout reproductive life. *Maturitas* (1988) 10(1):35–43. doi: 10.1016/0378-5122(88)90129-6
15. Fanchin R, Schonauer LM, Righini C, Guibourdenche J, Frydman R, Taieb J. Serum anti-mullerian hormone is more strongly related to ovarian follicular status than serum inhibin b, estradiol, fsh and lh on day 3. *Hum Reprod* (2003) 18(2):323–7. doi: 10.1093/humrep/deg042
16. Ferraretti AP, La Marca A, Fauser BC, Tarlatzis B, Nargund G, Gianaroli L. Eshre consensus on the definition of 'Poor response' to ovarian stimulation for in vitro fertilization: The Bologna criteria. *Hum Reprod* (2011) 26(7):1616–24. doi: 10.1093/humrep/der092
17. Alviggi C, Andersen CY, Buehler K, Conforti A, De Placido G, Esteves SC, et al. A new more detailed stratification of low responders to ovarian stimulation: From a poor ovarian response to a low prognosis concept. *Fertil Steril* (2016) 105(6):1452–3. doi: 10.1016/j.fertnstert.2016.02.005
18. Groome NP, Illingworth PJ, O'Brien M, Pai R, Rodger FE, Mather JP, et al. Measurement of dimeric inhibin b throughout the human menstrual cycle. *J Clin Endocrinol Metab* (1996) 81(4):1401–5. doi: 10.1210/jcem.81.4.8636341
19. Welt CK, Schneyer AL. Differential regulation of inhibin b and inhibin a by follicle-stimulating hormone and local growth factors in human granulosa cells from small antral follicles. *J Clin Endocrinol Metab* (2001) 86(1):330–6. doi: 10.1210/jcem.86.1.7107
20. Wen J, Huang K, Du X, Zhang H, Ding T, Zhang C, et al. Can inhibin b reflect ovarian reserve of healthy reproductive age women effectively? *Front Endocrinol* (2021) 12:626534. doi: 10.3389/fendo.2021.626534
21. Danis RB, Sriprasert I, Ho JR, McGinnis LK, Kumar A, Stanczyk FZ. Association of bioavailable inhibin b and oocyte yield in controlled ovarian stimulation. *FS Rep* (2021) 2(2):189–94. doi: 10.1016/j.xfre.2021.03.005
22. Eldar-Geva T, Ben-Chetrit A, Spitz IM, Rabinowitz R, Markowitz E, Mimoni T, et al. Dynamic assays of inhibin b, anti-mullerian hormone and estradiol following fsh stimulation and ovarian ultrasonography as predictors of ivf outcome. *Hum Reprod* (2005) 20(11):3178–83. doi: 10.1093/humrep/dei203
23. Van Den Berg HA. Occam's razor: From ockham's *Via moderna* to modern data science. *Sci Prog* (2018) 101(3):261–72. doi: 10.3184/003685018X15295002645082
24. Findlay JK, Hutt KJ, Hickey M, Anderson RA. What is the "Ovarian reserve"? *Fertil Steril* (2015) 103(3):628–30. doi: 10.1016/j.fertnstert.2014.10.037
25. Broer SL, Broekmans FJ, Laven JS, Fauser BC. Anti-mullerian hormone: Ovarian reserve testing and its potential clinical implications. *Hum Reprod Update* (2014) 20(5):688–701. doi: 10.1093/humupd/dmu020
26. Seifer DB, Lambert-Messerlian G, Hogan JW, Gardiner AC, Blazar AS, Berk CA. Day 3 serum inhibin-b is predictive of assisted reproductive technologies outcome. *Fertil Steril* (1997) 67(1):110–4. doi: 10.1016/s0015-0282(97)81865-1
27. Hall JE, Welt CK, Cramer DW. Inhibin a and inhibin b reflect ovarian function in assisted reproduction but are less useful at predicting outcome. *Hum Reprod* (1999) 14(2):409–15. doi: 10.1093/humrep/14.2.409
28. Creus M, Penarrubia J, Fabregues F, Vidal E, Carmona F, Casamitjana R, et al. Day 3 serum inhibin b and fsh and age as predictors of assisted reproduction treatment outcome. *Hum Reprod* (2000) 15(11):2341–6. doi: 10.1093/humrep/15.11.2341
29. Bancsi LF, Broekmans FJ, Eijkemans MJ, de Jong FH, Habbema JD, te Velde ER. Predictors of poor ovarian response in in vitro fertilization: A prospective study comparing basal markers of ovarian reserve. *Fertil Steril* (2002) 77(2):328–36. doi: 10.1016/s0015-0282(01)02983-1
30. Decanter C, Pigny P, Lefebvre C, Thomas P, Leroy M, Dewailly D. Serum inhibin b during controlled ovarian hyperstimulation: An additional criterion for deciding whether to proceed with egg retrieval. *Fertil Steril* (2009) 91(6):2419–25. doi: 10.1016/j.fertnstert.2008.03.037
31. Fawzy M, Lambert A, Harrison RF, Knight PG, Groome N, Hennelly B, et al. Day 5 inhibin b levels in a treatment cycle are predictive of ivf outcome. *Hum Reprod* (2002) 17(6):1535–43. doi: 10.1093/humrep/17.6.1535
32. Muttukrishna S, Suharjo H, McGarrigle H, Sathanandan M. Inhibin b and anti-mullerian hormone: Markers of ovarian response in Ivf/Icsi patients? *BJOG* (2004) 111(11):1248–53. doi: 10.1111/j.1471-0528.2004.00452.x
33. Tan R, Pu D, Liu L, Liu J, Wu J. Comparisons of inhibin b versus antimullerian hormone in poor ovarian responders undergoing in vitro fertilization. *Fertil Steril* (2011) 96(4):905–11. doi: 10.1016/j.fertnstert.2011.07.1102
34. Penarrubia J, Peralta S, Fabregues F, Carmona F, Casamitjana R, Balasch J. Day-5 inhibin b serum concentrations and antral follicle count as predictors of ovarian response and live birth in assisted reproduction cycles stimulated with gonadotropin after pituitary suppression. *Fertility sterility* (2010) 94(7):2590–5. doi: 10.1016/j.fertnstert.2010.03.009
35. Li Y, Nie M, Liu Y, Zhang W, Yang X. The dynamic changes of anti-mullerian hormone and inhibin b during controlled ovarian hyperstimulation in decreased ovarian reserve women and the effect on clinical outcome. *Gynecol Endocrinol* (2015) 31(6):450–3. doi: 10.3109/09513590.2014.998187
36. Fanchin R, Schonauer LM, Righini C, Frydman N, Frydman R, Taieb J. Serum anti-mullerian hormone dynamics during controlled ovarian hyperstimulation. *Hum Reprod* (2003) 18(2):328–32. doi: 10.1093/humrep/deg043
37. Henriquez S, Kohen P, Xu X, Villarrol C, Munoz A, Godoy A, et al. Significance of pro-angiogenic estrogen metabolites in normal follicular development and follicular growth arrest in polycystic ovary syndrome. *Hum Reprod* (2020) 35(7):1655–65. doi: 10.1093/humrep/deaa098
38. Ho HY, Lee RKK, Lin MH, Hwu YM. Estradiol level on day 9 as a predictor of risk for ovarian hyperresponse during controlled ovarian hyperstimulation. *J Assist Reprod Gen* (2003) 20(6):222–6. doi: 10.1023/A:1024155411444
39. Hendriks DJ, Klinkert ER, Bancsi LFJMM, Looman CWN, Habbema JDF, te Velde ER, et al. Use of stimulated serum estradiol measurements for the prediction of hyperresponse to ovarian stimulation in in vitro fertilization (Ivf). *J Assist Reprod Gen* (2004) 21(3):65–72. doi: 10.1023/B:JARG.0000027016.65749.ad