Reproductive BioMedicine Online (2014) xxx, xxx–xxx

www.sciencedirect.com
www.rbmonline.com

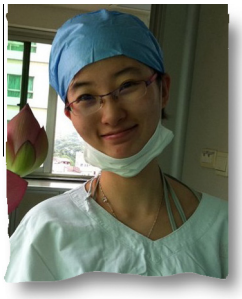
ARTICLE

Predictive value of androgens and multivariate model for poor ovarian response

Jing Guo, Qingxue Zhang, Yu Li, Jia Huang, Wenjun Wang, Lili Huang, Xiaomiao Zhao *,[1], Dongzi Yang *,[1]

Department of Obstetrics and Gynecology, Sun Yat–Sen Memorial Hospital of Sun Yat–Sen University, Guangzhou, Guangdong, China

* Corresponding authors. *E-mail addresses:* zhaoxmiao@163.com (X Zhao), yangdz@mail.sysu.edu.cn (D Yang). [1] Dongzi Yang and Xiaomiao Zhao contributed equally as the co-corresponding authors.



Dr Guo is a PhD student of reproductive endocrinology and infertility at the Memorial Hospital of Sun Yat-sen University of China, with a special focus on assisted reproduction technology. She is specialized in female reproductive endocrinology. Dr Guo's current research includes strategies for predicting poor ovarian response, investigating the basic mechanisms involved in the effects of vitamin D on oocyte development and maturation and describing the abnormal metabolism with polycystic ovarian syndrome.

Abstract No single or multivariate model is effective for predicting poor ovarian response (POR) with satisfactory sensitivity and specificity. This study investigated whether dehydroepiandrosterone sulphate (DHEAS) or basal testosterone concentrations could be effective predictors of POR defined by the Bologna criteria. This retrospective study included 79 poor responders and 128 normal responders. Serum FSH, LH, oestradiol, DHEAS and testosterone concentrations on day 3 of the menstrual cycle before the treatment cycle were measured. All patients received standard ovarian stimulation with FSH under pituitary suppression with gonadotrophin-releasing hormone agonist. DHEAS concentration was not significantly different between poor and normal responders or between pregnant and nonpregnant women. Basal testosterone, unlike DHEAS concentration, was predictive, but with limited ability as a single predictor, for POR. The multivariate model composed of age, AFC, FSH, FSH/LH and testosterone was reliably predictive for POR (ROC_{AUC} = 0.976, cut-off point >0.51, sensitivity 88.6%, specificity 98.3%) and clinical pregnancy (ROC_{AUC} = 0.716, cut-off point ≤−0.22, sensitivity 75%, specificity 62.5%) and was better than antral follicle count for predicting both POR and clinical pregnancy. This multivariate model might be useful for identifying patients at risk of poor response in order to optimize the stimulation regimens.

KEYWORDS: dehydroepiandrosterone sulphate, IVF, multivariate model, poor ovarian response, testosterone, pregnancy outcome

http://dx.doi.org/10.1016/j.rbmo.2014.02.009
1472-6483/

Please cite this article in press as: Guo, J et al. Predictive value of androgens and multivariate model for poor ovarian response. Reproductive BioMedicine Online (2014), http://dx.doi.org/10.1016/j.rbmo.2014.02.009

Introduction

Delayed childbearing in women has been a significant demographic trend, with the consequence of a marked increase in the numbers of older women who often fail to respond satisfactorily to ovarian stimulation (Sunkara et al., 2012). A proportion of women (2–30%) undergoing ovarian stimulation experience poor response (Hendriks et al., 2005), which results in treatment cancellation and reduced live birth rate. A recent international survey involving 196 IVF centres in 45 countries reported an increase in the burden of poor ovarian response (POR) over the last decade (IVF Worldwide Survey, 2010). Identification of women at increased risk for POR prior to IVF could be useful, as this could either prevent unnecessary continuation of treatment (Klinkert et al., 2004) or help to make individual interventions in order to maximize ovarian response (Klinkert et al., 2005).

Several reviews have showed the predictive value of single and combined tests performed in basal conditions. Of all the tests, antral follicle count (AFC) and anti-Müllerian homone (AMH) has the best sensitivity and specificity for predicting POR (Broekmans et al., 2006; La Marca et al., 2010; Verhagen et al., 2008). However, even the best ovarian reserve marker at the best cut-off value is associated with a false positive rate of 10–20% (Broekmans et al., 2006; Ferraretti et al., 2011), which may have negative consequences on the couple's life since false positive results might incorrectly prohibit these women from undergoing IVF (La Marca et al., 2010). Besides, the accuracy of predicting the occurrence of pregnancy is very limited for all tests (Broekmans et al., 2006; Broer et al., 2009). In addition, more than 35–41 definitions for POR were used in these studies, implying a troublesome issue in clinical application. The Bologna criteria developed by European Society for Human Reproduction and Embryology consensus in 2011 for the first time reached an agreement on universal definition of POR (Ferraretti et al., 2011). A single and simple test demonstrating a better performance for predicting POR than the currently available tests would be preferable.

The potential stimulating role of androgens on folliculogenesis has been suggested by a number of basic research studies and illustrated by some pathophysiological conditions and clinical models (Fanchin et al., 2011). To assess the possible action of androgens on human ovary, some investigators focused on the potential effect of androgen pretreatment before ovarian stimulation. Whether androgen supplementation is an effective treatment for POR remains highly controversial (Urman and Yakin, 2012; Yakin and Urman, 2011); however, some studies have reported encouraging outcomes, including improved ovarian response and live birth rate, with systemic administration of either dehydroepiandrosterone (DHEA) or testosterone (Gleicher et al., 2010; Kim et al., 2011). However, it is worth noting that none of the studies has characterized the androgen status of the participating women prior to treatment (Sunkara et al., 2012).

According to the 'androgen hypothesis' for treating ovarian function defects (Fanchin et al., 2011), the current study investigated whether DHEAS or testosterone concentrations could be effective predictors for POR. As a second target, the accuracy of a multivariate model for predicting POR following the Bologna criteria was estimated.

Materials and methods

Study population

All patient information was obtained from the database of the Centre for Reproductive Medicine and fertility. Study patients were recruited consecutively in this retrospective study. Study patients fulfilling the inclusion criteria from March 2011 to March 2013 were defined as either poor or normal responders.

According to the Bologna criteria (Ferraretti et al., 2011), POR in this study was defined if one of the following four features was present: (i) AFC $\leq$5 follicles and age $\geq$40 years; (ii) age $\geq$40 years and a previous POR ($\leq$3 oocytes collected with a conventional stimulation protocol in which at least 150 IU FSH was consumed per day); (iii) AFC $\leq$5 follicles and a previous POR; and (iv) two episodes of POR after maximal stimulation (cycle was cancelled for following the development of less than three growing follicles).

Normal responders who were considered as the control population satisfied all the following conditions simultaneously: (i) their first IVF–embryo transfer (fresh) cycle; (ii) $\geq$5 oocytes with a conventional stimulation protocol; and (iii) age $\leq$35 years, AFC $\geq$7 follicles and basal FSH $<$13 IU/l.

Exclusion criteria were patients who received androgen supplementation at any time before enrolment. Patients with endocrine disorders or anatomical abnormalities were excluded, including polycystic ovarian syndrome (PCOS), abnormal thyroid function and hyperprolactinaemia, as well as uterine malformation, submucous myoma and multiple myomata. PCOS was diagnosed according to the Rotterdam criteria (Rotterdam ESHRE/ASRM–Sponsored PCOS Consensus Workshop Group, 2004). Hyperandrogenaemia was defined as serum DHEAS $>$4.92 μmol/l or testosterone $>$2.39 nmol/l (Zhao et al., 2011). Thyroid function was screened by serum thyroid-stimulating hormone (0.55–4.78 mU/l), triiodothyronine (0.92–2.79 nmol/l), thyroxine (58.1–140.6 nmol/l), free triiodothyronine (3.5–6.5 pmol/l) and free thyroxine (11.5–22.7 pmol/l) combined with the clinical symptoms and signs. Anatomical abnormalities were discovered by abdominal ultrasound scanning and transvaginal sonography scanning.

On day 3 of a spontaneous menstrual cycle within 3 months before fresh IVF cycle, a blood sample was taken in the morning to evaluate basal hormone (FSH, LH, oestradiol, testosterone, DHEAS) concentrations. On the same day, transvaginal sonography was performed to obtain AFC and ovarian volume. As recommended (Broekmans et al., 2010), the follicles visualized and counted were 2–10 mm in size, and the numbers of follicles in both ovaries were added to obtain the total AFC. The volume of each ovary was calculated by measuring the ovarian diameters in three perpendicular directions and the final result was calculated automatically. The volumes of both ovaries were added to obtain a mean value which was defined as mean ovarian volume.

Please cite this article in press as: Guo, J et al. Predictive value of androgens and multivariate model for poor ovarian response. Reproductive BioMedicine Online (2014), http://dx.doi.org/10.1016/j.rbmo.2014.02.009

Ethical approval

Ethical approval was not required for this study as it was a retrospective study in which data were collected from the study centre's database. No potential intervention that had any unusual hazards inherent was applied. Written informed consent was obtained from all study subjects.

Treatment protocol

All patients received standard ovarian stimulation with FSH under pituitary suppression with gonadotrophin-releasing hormone (GnRH) agonist according to a protocol used routinely (Penarrubia et al., 2010). None of the patients used traditional Chinese medicine or oral contraceptives prior to stimulation start.

Dosage of GnRH agonist was chosen flexibly according to patient age, AFC and basal endocrine status. For all of the normal responders and eight (10.1%) poor responders, controlled-release GnRH analogue (1.25 mg or 0.85 mg/ampoule; triptorelin; Ipsen, France) was administered once in mid-luteal phase. For majority of poor responders, short-acting GnRH analogue (0.1 mg or 0.05 mg/ampoule; triptorelin) was used daily from the midluteal phase till the day of human chorionic gonadotrophin (HCG) administration.

When a satisfactory pituitary desensitization was achieved, stimulation with exogenous gonadotrophins (Gonal-F; Serono, Switzerland) was started. Serial monitoring of ovarian response was assessed by transvaginal ultrasound and serum total oestradiol assays. The criteria for HCG (Profasi; Serono, Switzerland) administration were the presence of at least one follicle $\geq$18 mm in diameter with a consistent rise in serum oestradiol. Oocyte aspiration was performed 36–38 h later with transvaginal sonography.

Standard laboratory protocols were followed. Embryo transfer took place 2 or 3 days after oocyte retrieval, depending on patient age and the number and quality of embryos available. As a rule, and if available, two embryos were transferred in young women (<35 years) and three embryos were transferred in older women ($\geq$35 years). The luteal phase was supported with progesterone in oil (40 mg i.m. daily) starting on the day of oocyte aspiration and continuing either up to menstruation or at least the first 8 weeks of pregnancy if the patient became pregnant (Li et al., 2010). Clinical pregnancy was diagnosed by increasing serum βHCG on days 10–14 after embryo transfer and the subsequent demonstration of an intrauterine gestational sac by transvaginal sonography.

Hormone assays

All blood samples were immediately processed to separate the serum. Hormone assays were performed on the serum samples by two experienced technicians working in the gynaecological endocrine laboratory of SunYat–Sen Memorial Hospital. Concentrations of FSH, LH, oestradiol and testosterone were measured by chemiluminescence using a ACS180.SE autoanalyser (Bayer Diagnostics, Fernwald, Germany) and DHEAS concentration was measured by enzyme-linked immunosorbent assay (DRG Instruments, Marburg, Germany), according to the manufacturer's instructions. The intraassay and interassay coefficients of variance for hormones measured were as follows: FSH (1.5% and 4.3%), LH (1.7% and 3.9%), oestradiol (2.5% and 4.1%), testosterone (2.6% and 3.9%) and DHEAS (4.6% and 5.4%), respectively.

Statistical analysis

Statistical analysis was performed using Statistical Package for Social Sciences version 13.0 (SPSS, Chicago, IL, USA). Normality was tested by the Kolmogorov–Smirnov test. Quantitative variables were presented as mean ± standard deviation (SD) or median (interquartile range). Qualitative variables were presented as *n* (%). Normally distributed variables were compared by the Student's t-test. Due to non-normality of quantitative variables, the Mann–Whitney test was used to compare two independent groups. $P < 0.05$ was considered statistically significant. Comparison of causes of infertility in different groups was assessed using Fisher's exact test. Spearman correlation was used to assess the association between two quantitative variables.

Considering that the study variables (predictors for POR) were relative with each other, principal component analysis (PCA) was used to produce independent factors (principal components) which were appropriate for multivariate logistic regression analysis. Meantime, different combinations of study variables that would reflect POR to the highest degree were screened and selected. First, the univariate logistic regression analysis was used to estimate the risk of each principal component on the dependent variable. Secondly, the multivariate logistic regression analysis was used to assess the independent effect of these components after controlling for confounders.

Receiver operating characteristic (ROC) curves were constructed to examine the predictive value of variables for POR and clinical pregnancy. The area under the curve (AUC) and the cut-off point were computed. Finally, the best test was selected out with the biggest likelihood ratio of a positive test (LR+) and the smallest likelihood ratio of a negative test (LR−).

Results

Baseline and cycle characteristics of POR

A total of 79 poor ovarian responders and 128 normal ovarian responders were eligible for this study. For the poor ovarian responders, 58 cases underwent a fresh IVF–embryo transfer cycle, seven cases were cancelled due to poor ovarian response following ovarian stimulation and 14 cases had cycles cancelled as no suitable embryos were available for transfer. Types of infertility (primary, secondary) and causes of infertility (male factor, tubal factor, endometriosis, ovulatory obstruction, unexplained infertility, combined factors) were comparable between poor and normal ovarian responders. As shown in Table 1, poor ovarian responders were significantly older than normal ovarian responders ($P < 0.001$), with higher body mass index (BMI; $P = 0.001$), basal FSH ($P < 0.001$), FSH/LH ($P < 0.001$) and oestradiol ($P < 0.001$) and lower AFC ($P < 0.001$), mean ovarian volume ($P < 0.001$) and testosterone ($P = 0.022$). A

significantly higher total gonadotrophin dosage was consumed and lower peak oestradiol, number of mature oocytes, fertilization rate, cleavage rate, implantation rate and clinical pregnancy rate (all $P < 0.001$) were achieved by poor ovarian responders. DHEAS concentration was not significantly different between the two groups.

Relationship between study variables and clinical pregnancy

Patients achieving clinical pregnancy were significantly younger ($P < 0.001$) and had a higher AFC ($P < 0.001$), greater mean ovarian volume ($P = 0.038$), lower basal FSH ($P = 0.002$), lower basal FSH/LH ($P = 0.001$) and higher basal testosterone ($P = 0.044$) than patients who did not achieve clinical pregnancy. As shown in Table 2, no significant differences were observed in BMI or oestradiol or DHEAS concentrations between the two groups, and number of embryo transfers was comparable. Patients achieving clinical pregnancy consumed a significantly lower total gonadotrophin dose ($P = 0.009$) and had significantly higher peak oestradiol ($P = 0.021$), number of oocytes retrieved ($P = 0.003$) and number of mature oocytes ($P = 0.002$). Spearman correlation suggested that age ($R = -0.58$, $P < 0.001$), BMI ($R = -0.241$, $P = 0.001$), AFC ($R = 0.598$, $P < 0.001$), basal FSH ($R = -0.309$, $P < 0.001$), basal FSH/LH ($R = 0.196$, $P = 0.006$) and basal tes-

Table 1 Baseline, treatment and outcome characteristics of poor responders compared with normal responders.

Characteristic	*Poor responders* (n = 79)	*Normal responders* (n = 128)	P-*value*
Age (years)	39.3 ± 3.8	29.7 ± 3.3	<0.001
BMI (kg/m^2)	21.4 ± 3.67	20.2 ± 2.28	0.001
AFC	10 (7–13)	13 (10–18)	<0.001
Mean ovarian volume (ml)	3.63 (2.62–5.71)	6.22 (4.33–8.52)	<0.001
Basal FSH (IU/l)	10.8 ± 5.9	7.77 ± 2.11	<0.001
Basal FSH/LH	2.96 ± 3.87	1.98 ± 0.93	<0.001
Basal oestradiol (pg/ml)	60.59 ± 108.32	40.56 ± 20.85	<0.001
Basal testosterone (nmol/l)	0.88 ± 0.55	1.05 ± 0.56	0.022
DHEAS (μmol/l)	3.18 ± 1.0	3.44 ± 1.00	NS
Total gonadotrophin dose (IU)	3253 ± 932	2103 ± 673	<0.001
Peak oestradiol (pg/ml)	1257 ± 1153	3289 ± 1376	<0.001
No. of mature oocytes	2 (1–5)	11 (8–15)	<0.001
Fertilization rate (%)	18.4 ± 36.9	82.5 ± 19.6	<0.001
Cleavage rate (%)	20.7 ± 40.2	96.5 ± 10.8	<0.001
Implantation rate (%)	16.1 ± 41.9	42.5 ± 59.8	<0.001
Clinical pregnancies	12 (20.7)	71 (55.5)	<0.001

Values are mean ± SD, median (interquartile range) or *n* (%). AFC = antral follicle count; BMI = body mass index; DHEAS = dehydroepiandrosterone sulphate; NS = not statistically significant.

Table 2 Patient characteristics according to clinical pregnancy status.

Characteristic	*Pregnant* (n = 83)	*Not pregnant* (n = 103)	P-*value*
Age (years)	30.57 ± 4.57	34.65 ± 6.07	<0.001
BMI (kg/m^2)	20.54 ± 3.34	20.48 ± 2.43	NS
AFC	14 (10–19)	10 (5–15)	<0.001
Mean ovarian volume (ml)	5.83 (4.16–8.86)	4.89 (3.09–7.3)	0.038
Basal FSH (IU/l)	8.21 ± 4.3	9.48 ± 4.46	0.002
Basal FSH/LH	1.91 ± 0.9	2.68 ± 3.43	0.001
Basal oestradiol (pg/ml)	39.07 ± 22.17	53.17 ± 92.6	NS
Basal testosterone (nmol/l)	1.1 ± 0.6	0.9 ± 0.5	0.044
DHEAS (μmol/l)	3.27 ± 1.08	3.47 ± 0.95	NS
Total gonadotrophin dose (IU)	2237 ± 679	2645 ± 1046	0.009
Peak oestradiol (pg/ml)	2963 ± 1458	2363 ± 1659	0.021
No. of oocytes retrieved	12 (7–16)	9 (3–14)	0.003
No. of mature oocytes	10 (6–15)	8 (3–12)	0.002
No. of embryos transferred	2 (2–2)	2 (2–2)	NS

Values are mean ± SD or median (interquartile range).
AFC = antral follicle count; BMI = body mass index; DHEAS = dehydroepiandrosterone sulphate; NS = not statistically significant.

Please cite this article in press as: Guo, J et al. Predictive value of androgens and multivariate model for poor ovarian response. Reproductive BioMedicine Online (2014), http://dx.doi.org/10.1016/j.rbmo.2014.02.009

Table 3 Spearman correlations for patient characteristics according to number of oocytes and clinical pregnancy in women undergoing ovarian stimulation and embryo transfer.

Characteristic	*No. of oocytes*		*Clinical pregnancy*[a]	
	R	P-*value*	R	P-*value*
Age (years)	−0.58	<0.001	−0.333	<0.001
BMI (kg/m^2)	−0.241	0.001	0.061	NS
AFC	0.598	<0.001	0.27	<0.001
Mean ovarian volume (ml)	−0.018	NS	0.169	0.038
Basal FSH (IU/l)	−0.309	<0.001	−0.224	0.002
Basal FSH/LH	0.196	0.006	−0.252	0.001
Basal oestradiol (pg/ml)	−0.077	NS	0.112	NS
Basal testosterone (nmol/l)	0.149	0.038	0.15	0.043
DHEAS (μmol/l)	−0.033	NS	−0.079	NS

Values are mean ± SD or median (interquartile range).
AFC = antral follicle count; BMI = body mass index; DHEAS = dehydroepiandrosterone sulphate; NS = not statistically significant.
[a]Dichotomous variables for pregnancy: 0 = no pregnancy; 1 = pregnancy.

tosterone ($R = 0.149$, $P = 0.038$) were significantly correlated with the number of retrieved oocytes, while age ($R = -0.333$, $P < 0.001$), AFC ($R = 0.27$, $P < 0.001$), mean ovarian volume ($R = 0.169$, $P = 0.038$), basal FSH ($R = -0.224$, $P = 0.002$), basal FSH/LH ($R = -0.252$, $P = 0.001$) and basal testosterone ($R = 0.15$, $P = 0.043$) were significantly correlated with clinical pregnancy occurrence (Table 3). In summary, among all study variables, age, AFC, basal FSH, basal FSH/LH and basal testosterone were significantly associated with both ovarian response and clinical pregnancy.

Predictive value of study variables for POR or clinical pregnancy

The PCA of predictors for POR is shown in Table 4. The results of the Kaiser–Meyer–Olkin and Bartlett's test suggested that it was a proper method for undergoing PCA. According to the Eigen value (considering >1) and total variance explained (considering cumulative variance explained >80%) for each component, the first three components were finally selected. The component score functions were as follows: $X_1 = (0.563 \times Z_{age}) - (0.505 \times Z_{AFC}) + (0.506 \times Z_{FSH}) + (0.34 \times Z_{FSH/LH}) - (0.24 \times Z_{testosterone})$; $X_2 = (-0.223 \times Z_{age}) + (0.374 \times Z_{AFC}) + (0.384 \times Z_{FSH}) + (0.676 \times Z_{FSH/LH}) + (0.455 \times Z_{testosterone})$; $X_3 = (0.288 \times Z_{age}) - (0.368 \times Z_{AFC}) - (0.166 \times Z_{FSH}) - (0.178 \times Z_{FSH/LH}) + (0.849 \times Z_{testosterone})$, where Z_{Xi} = standardized values of X_i. Then, logistic regression analysis was used to estimate the principal components achieved for predicting POR and the outcome of no clinical pregnancy (Table 5). Both univariate and multivariate logistic regression analysis showed that X_1, X_2 and X_3 were predictors for POR, while only X_1 was predictive for no clinical pregnancy.

As shown in Figure 1, ROC curves indicated that a X_1 value of 0.51 produced a maximum sensitivity of 88.6% and a specificity of 98.3% for predicting POR and that a X_1 value of −0.22 produced a maximum sensitivity of 75% and a specificity of 62.5% with a ROC_{AUC} >0.7 for predicting clinical pregnancy. A AFC value of 9 produced a maximum sensitivity of 84.2% and a specificity of 88.6% for predicting POR and a AFC value of 12 produced a maximum sensitivity of 68.7% and specificity of 58.6% with a ROC_{AUC} <0.7 for predicting clinical pregnancy. Predicting either POR or clinical pregnancy, X_1 had a larger ROC_{AUC}, a greater LR + and a smaller LR− than AFC (Table 6). ROC curves also indicated that basal testosterone was predictive for POR ($ROC_{AUC} = 0.644$, $P = 0.003$) but did not reach statistical significance for predicting clinical pregnancy ($ROC_{AUC} = 0.583$, $P > 0.05$).

Discussion

This study examined the role of serum androgen concentrations in Chinese women without known factors affecting androgen concentrations undergoing IVF/intracytoplasmic sperm injection in the prediction of POR following Bologna criteria as the primary outcome. This study suggested that basal testosterone, instead of DHEAS concentration, was predictive, but with limited ability as a single predictor, for POR. However, the multivariate model composed of age, AFC, basal FSH, basal FSH/LH and basal testosterone performed better than AFC both for predicting POR and clinical pregnancy.

The study by Frattarelli and Peterson (2004) first indicated a correlation between DHEAS concentration and ovarian response to gonadotrophins, and day-3 testosterone concentration <20 ng/dl were associated with poor IVF success rates. Nevertheless, in a subsequent study (Frattarelli and Gerber, 2006), only concentrations of ovarian androgens (i.e. testosterone and androstenedione) were found to be correlated with many IVF stimulation parameters, most significantly with peak oestradiol , number of follicles and number of oocytes retrieved, but none could predict pregnancy. In accordance with that study, the data in this study cast doubt on the importance of adrenal androgen production during IVF. This study demonstrates that testosterone, most of which is of ovarian origin, is important for IVF success. This is partly in accordance with the study by Qin et al. (2011), which suggested that basal testosterone

Please cite this article in press as: Guo, J et al. Predictive value of androgens and multivariate model for poor ovarian response. Reproductive BioMedicine Online (2014), http://dx.doi.org/10.1016/j.rbmo.2014.02.009

Table 4 Principal component analysis for predictors of poor ovarian response.

Kaiser–Meyer–Olkin and Bartlett's test			
Kaiser–Meyer–Olkin measure of sampling adequacy	0.706		
Bartlett's test of sphericity			
Approximate chi-squared	181.304		
Degrees of freedom	10		
P-value	<0.001		
Total variance explained			
Component	Initial Eigen value		
	Total	Variance (%)	Cumulative variance (%)
1	2.099	41.979	41.979
2	1.193	23.857	65.836
3	0.868	17.357	83.193
4	0.47	9.394	92.588
5	0.371	7.412	100
Component matrix			
Variable	Component		
	X_1	X_2	X_3
Age	0.815	−0.244	0.268
AFC	−0.731	0.409	−0.343
Basal FSH	0.733	0.419	−0.155
Basal FSH/LH	0.492	0.738	−0.166
Basal testosterone	−0.347	0.497	0.791

X_1 = first component: $(0.563 \times Z_{age}) - (0.505 \times Z_{AFC}) + (0.506 \times Z_{FSH}) + (0.34 \times Z_{FSH/LH}) - (0.24 \times Z_{testosterone})$.
X_2 = second component: $(-0.223 \times Z_{age}) + (0.374 \times Z_{AFC}) + (0.384 \times Z_{FSH}) + (0.676 \times Z_{FSH/LH}) + (0.455 \times Z_{testosterone})$.
X_3 = third component: $(0.288 \times Z_{age}) - (0.368 \times Z_{AFC}) - (0.166 \times Z_{FSH}) - (0.178 \times Z_{FSH/LH}) + (0.849 \times Z_{testosterone})$.
AFC = antral follicle count.

Table 5 Principal components achieved for poor ovarian response and no clinical pregnancy using univariate and multivariate logistic regression analysis.

Factor	*Odds ratio (95% CI)*	*P-value*
Poor ovarian response		
Univariate analysis		
X_1	45.666 (13.688–152.349)	<0.001
X_2	0.396 (0.25–0.629)	<0.001
X_3	1.845 (1.312–2.596)	<0.001
Multivariate analysis		
X_1	469.369 (23.531–9362.564)	<0.001
X_2	0.081 (0.015–0.421)	0.003
X_3	32.995 (5.087–214.021)	<0.001
No pregnancy		
Univariate analysis		
X_1	0.502 (0.373–0.676)	<0.001
X_2	1.195 (0.865–1.651)	NS
X_3	1.059 (0.762–1.472)	NS
Multivariate analysis		
X_1	0.497 (0.344–0.717)	<0.001
X_2	0.991 (0.543–1.811)	NS
X_3	1.102 (0.758–1.602)	NS

X_1 = first component; X_2 = second component; X_3 = third component (see note for Table 4 for calculation).
NS = not statistically significant.

Please cite this article in press as: Guo, J et al. Predictive value of androgens and multivariate model for poor ovarian response. Reproductive BioMedicine Online (2014), http://dx.doi.org/10.1016/j.rbmo.2014.02.009

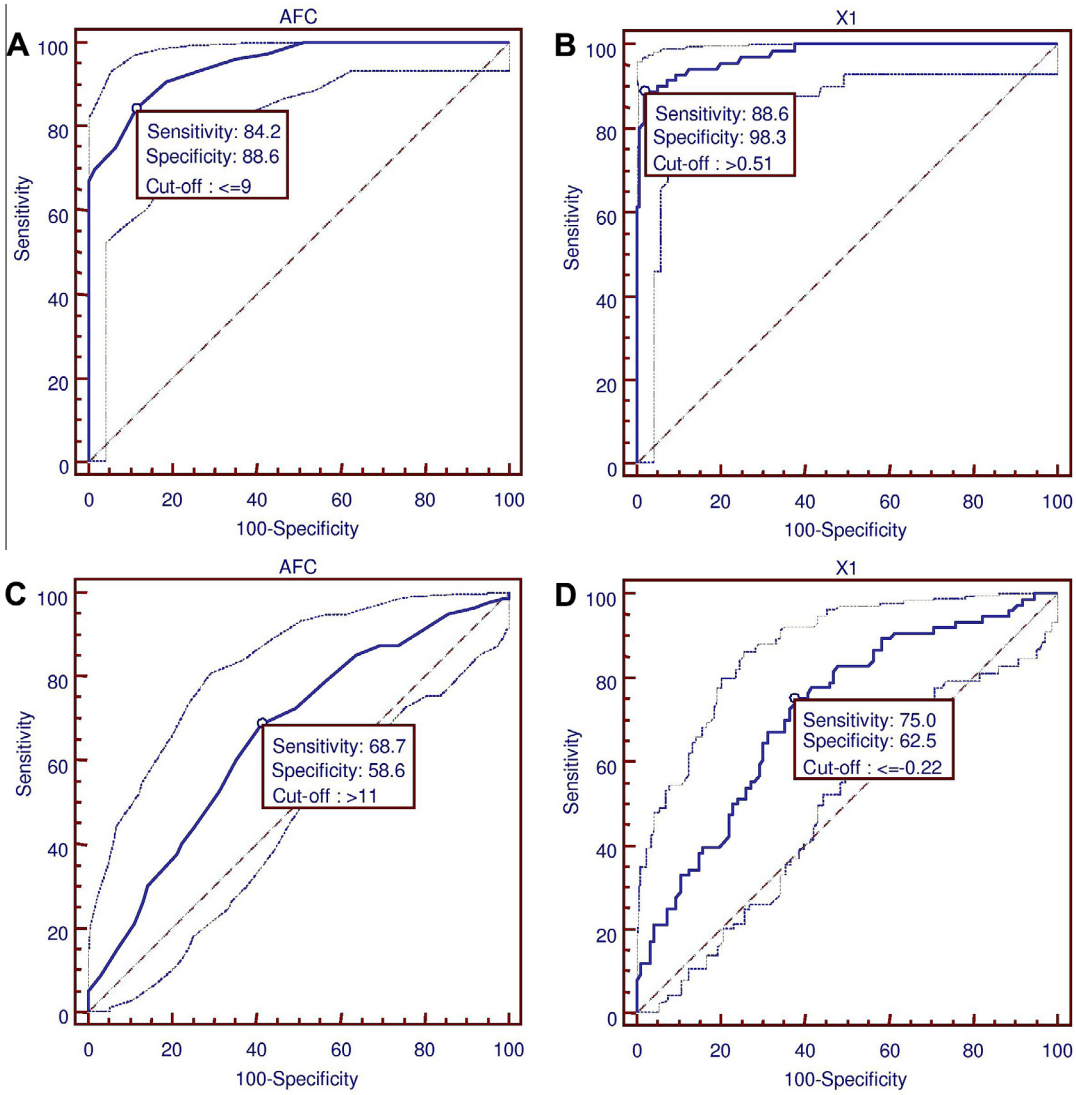


Figure 1 Receiver operating characteristic curves of AFC (A, C) and X_1 (B, D) for poor ovarian response (A, B) and clinical pregnancy (C, D) in women undergoing IVF.

Table 6 Comparison of performance of AFC and X_1 for poor ovarian response and clinical pregnancy.

Variable	*ROC_{AUC}*	*P-value*	*Cut off*	*Sensitivity (%)*	*Specificity (%)*	*LR+*	*LR−*
Poor ovarian response							
AFC	0.943	<0.001	≤9	84.2	88.6	7.39	0.18
X_1	0.976	<0.001	>0.51	88.6	98.3	52.12	0.12
Clinical pregnancy							
AFC	0.657	<0.001	>11	68.7	58.6	1.66	0.53
X_1	0.716	<0.001	≤−0.22	75.0	62.5	2.00	0.40

LR+ = likelihood ratio of a positive test: sensitivity/(1 − specificity); LR− = likelihood ratio of a negative test: (1 − sensitivity)/specificity.

concentration was a predictor for the number of large follicles on HCG day and pregnancy in Chinese women with diminished ovarian reserve, indicating that lower concentration of testosterone might relate with potential ovarian poor response.

According to the two-cell two-gonadotrophin theory, androgens play an essential role in ensuring adequate follicular steroidogenesis in humans (Hillier et al., 1994). They are produced primarily by the theca cells and believed to act as a substrate for aromatase in the granulosa cells,

Please cite this article in press as: Guo, J et al. Predictive value of androgens and multivariate model for poor ovarian response. Reproductive BioMedicine Online (2014), http://dx.doi.org/10.1016/j.rbmo.2014.02.009

which converts androgens to oestrogens (Sunkara et al., 2012). Besides, androgens exert a direct autocrine and/or paracrine effect to regulate follicular function (Suzuki et al., 1994). In a rhesus monkey model, androgen receptors are abundant in the granulosa cells of preantral and antral follicles (Vendola et al., 1998), and androgens augment FSH receptor expression in the granulosa cells to promote follicular growth and oestrogen biosynthesis by amplifying the effects of FSH (Weil et al., 1999). Similarly, a study in humans (Nielsen et al., 2011) has shown that androgen receptor mRNA and androgen concentrations in follicular fluid are correlated with FSH receptor mRNA expression in granulosa cells from small antral follicles. Gleicher and Barad (2011) suggested a new concept of ovarian ageing, where ovarian environments, but not oocytes themselves, age, therefore implying that normal androgenic ovarian endocrine microenvironments would positively influence pregnancy chances with IVF (Weghofer et al., 2012).

At the other extreme, high concentrations of androgens are capable of inducing the histopathological changes of PCOS in the ovary and bringing about follicular maturation arrest (Hugues and Durnerin, 2005). A meta-analysis has shown that there is a lower fertilization rate in PCOS patients undergoing IVF (Heijnen et al., 2006). Androgen excess might negatively impact oocyte quality and the current study excluded patients with hyperandrogenaemia. DHEAS, the adrenal original, is the most abundant pre-androgen in the circulation. No research has indicated the androgen concentrations necessary for promoting human follicular growth. Considering that DHEAS concentration is 100–500-times higher than testosterone concentration (Haning et al., 1991), it is deduced that follicle development is not so sensitive to changes of DHEAS as it is to changes in testosterone in local ovarian tissue in women of reproductive age.

A large number of clinical parameters might predict POR and they have been introduced to clinical practice. These include age, basal FSH (Scott et al., 1989, 1990), basal FSH/LH (Liu and Greenblatt, 2008; Mukherjee et al., 1996), basal oestradiol, inhibin B (Penarrubia et al., 2010; Seifer et al., 1997), AFC, ovarian volume (Jayaprakasan et al., 2009), a number of dynamic tests (Broekmans et al., 2006) and AMH (Lekamge et al., 2007). According to the hypothesis of Kol and Homburg (2008) of a theory of relativity, 2008), hormonal ratios, rather than absolute hormone concentrations, are more important for representing ovarian function. The interplay of LH and FSH represents a good example (Weghofer and Gleicher, 2009). Lenton et al. (1988) concluded that an increase in FSH several years before the elevation in LH and the first sign of decreased ovarian reserve might be due to an increased FSH/ LH ratio (Lenton et al., 1988). Of all tests, both AMH and AFC must be considered as the most reliable and accurate markers for predicting POR. This study did not measure AMH because it was retrospective, which represents a significant weakness. However, as a meta-analysis has shown (Broer et al., 2009), AMH has at least the same concentration of accuracy and clinical value for the prediction of poor response and no pregnancy as AFC. Therefore, this study performed a multivariate model analysis for predicting ovarian reserve and clinical pregnancy, compared with AFC.

Neither DHEAS or testosterone concentration, as a single predictor, was qualified for predicting POR and clinical pregnancy in IVF cycles. Therefore, the value of the multivariate model containing age, BMI, AFC, FSH, FSH/LH, oestradiol and testosterone was further evaluated. Logistic regression is most used to evaluate multivariables; however, it is not a suitable method when multicolinearity exists. As the data suggested, most study variables were highly correlated with each other. PCA was therefore recommended because it is a useful statistical technique for feature extraction and can help a classifier produce more accurate predictive performance (Avci and Turkoglu, 2009). PCA is on the assumption that the most information about classification is contained in the directions along which the feature values are the largest (Li and Sun, 2011). PCA is considered to be one of the best methods to identify unobservable 'latent' factors that underlie or explain a set of observed variables that are ordinal or interval scaled (Coste et al., 2005). PCA combined with logistic regression was chosen in this study following steps described in literature (Su et al., 2009). In the current work, the first component, X_1, might be considered as an integrative index of ovarian reserve. The component score was reliably predictive for POR and pregnancy outcome, and its power was better than AFC.

There are several limitations in this study. First, two different stimulation protocols were used, which represents a significant confounding variable. It would be appropriate to treat all patients with a single protocol in further studies. Secondly, poor responders were significantly older than normal responders, which was also a clear significant confounding variable given the fact that it has been shown that testosterone and DHEAS concentrations decline with age (Colakoglu, 1986). However, previous work by the current study group showed no significant differences were observed in testosterone and DHEAS concentrations in Chinese women aged <30 years, 30–34 years, 35–40 years and >40 years (Guo et al., 2014). A population of age-matched patients would be more precise.

As far as is known, this is the first study evaluating variables for POR following the Bologna criteria. In summary, serum basal testosterone concentration was a moderate predictor for POR. The multivariate model composed of age, AFC, basal FSH, basal FSH/LH and basal testosterone performed better than AFC for predicting both POR and pregnancy outcome. The equation could be easily calculated in clinical practice. This might be meaningful for most of reproductive medicine centres in China, where AMH measurement is costly and not in routine use, whereas basal detection of hypothalamic–pituitary–ovarian axis hormones and transvaginal sonography are available. Well-designed prospective research is needed to investigate the predictive value of the multivariate model for IVF success. Thereafter, clinicians would be able to identify patients at risk of poor response in order to optimize the stimulation regimens and to counsel them on probability of pregnancy as accurately as possible.

Acknowledgements

This work was supported by the Key Project of Research of the National Ministry of Health (2010 439), the specialized research fund for the doctoral programme of the Chinese Ministry of Education (Grant No. 20100171110085) and the

National Natural Science Foundation of China, China (Grant No. 30700899). The authors gratefully acknowledge the hormone analysis work of Tao Du and Yabo Yang in gynaecological endocrine laboratory.

References

Avci, E., Turkoglu, I., 2009. An intelligent diagnosis system based on principle component analysis and ANFIS for the heart valve diseases. Expert Syst. Appl. 36, 2873–2878.

Broekmans, F.J., Kwee, J., Hendriks, D.J., Mol, B.W., Lambalk, C.B., 2006. A systematic review of tests predicting ovarian reserve and IVF outcome. Hum. Reprod. Update 12, 685–718.

Broekmans, F.J., de Ziegler, D., Howles, C.M., Gougeon, A., Trew, G., Olivennes, F., 2010. The antral follicle count: practical recommendations for better standardization. Fertil. Steril. 94, 1044–1051.

Broer, S.L., Mol, B.W., Hendriks, D., Broekmans, F.J., 2009. The role of antimullerian hormone in prediction of outcome after IVF: comparison with the antral follicle count. Fertil. Steril. 91, 705–714.

Colakoglu, M., 1986. The effect of dehydroepiandrosterone sulfate prolactin and testosterone hormones in female fertility and hirsutism. Clin. Exp. Obstet. Gynecol. 13, 32–34.

Coste, J., Bouee, S., Ecosse, E., Leplege, A., Pouchot, J., 2005. Methodological issues in determining the dimensionality of composite health measures using principal component analysis: case illustration and suggestions for practice. Qual. Life Res. 14, 641–654.

Fanchin, R., Frydman, N., Even, M., Berwanger, D.S.A., Grynberg, M., Ayoubi, J.M., 2011. Androgens and poor responders: are we ready to take the plunge into clinical therapy? Fertil. Steril. 96, 1062–1065.

Ferraretti, A.P., La Marca, A., Fauser, B.C., Tarlatzis, B., Nargund, G., Gianaroli, L., 2011. ESHRE consensus on the definition of 'poor response' to ovarian stimulation for in vitro fertilization: the Bologna criteria. Hum. Reprod. 26, 1616–1624.

Frattarelli, J.L., Peterson, E.H., 2004. Effect of androgen levels on in vitro fertilization cycles. Fertil. Steril. 81, 1713–1714.

Frattarelli, J.L., Gerber, M.D., 2006. Basal and cycle androgen levels correlate with in vitro fertilization stimulation parameters but do not predict pregnancy outcome. Fertil. Steril. 86, 51–57.

Gleicher, N., Barad, D.H., 2011. Dehydroepiandrosterone (DHEA) supplementation in diminished ovarian reserve (DOR). Reprod. Biol. Endocrinol. 9.

Gleicher, N., Weghofer, A., Barad, D.H., 2010. Improvement in diminished ovarian reserve after dehydroepiandrosterone supplementation. Reprod. Biomed. Online 21, 360–365.

Guo, J., Zhang, Q.X., Li, Y., Huang, J., Wang, W.J., Huang, L.L., Zhao, X.M., Yang, D.Z., 2014. Predictive value of serum DHEAS level for poor ovarian response and pregnant outcome in IVF–ET cycles. J. Sun Yat-Sen Univ. 35, 156–160.

Haning, R.J., Flood, C.A., Hackett, R.J., Loughlin, J.S., McClure, N., Longcope, C., 1991. Metabolic clearance rate of dehydroepiandrosterone sulfate, its metabolism to testosterone, and its intrafollicular metabolism to dehydroepiandrosterone, androstenedione, testosterone, and dihydrotestosterone in vivo. J. Clin. Endocrinol. Metab. 72, 1088–1095.

Heijnen, E.M., Eijkemans, M.J., Hughes, E.G., Laven, J.S., Macklon, N.S., Fauser, B.C., 2006. A meta-analysis of outcomes of conventional IVF in women with polycystic ovary syndrome. Hum. Reprod. Update 12, 13–21.

Hendriks, D.J., Mol, B.W., Bancsi, L.F., Te, V.E., Broekmans, F.J., 2005. 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. 83, 291–301.

Hillier, S.G., Whitelaw, P.F., Smyth, C.D., 1994. Follicular oestrogen synthesis: the 'two-cell, two-gonadotrophin' model revisited. Mol. Cell. Endocrinol. 100, 51–54.

Hugues, J.N., Durnerin, I.C., 2005. Impact of androgens on fertility–physiological, clinical and therapeutic aspects. Reprod. Biomed. Online 11, 570–580.

IVF Worldwide Survey. Poor responders: how to define, diagnose and treat? <www.IVF-Worldwide.com> (8 August 2010, date last accessed).

Jayaprakasan, K., Al-Hasie, H., Jayaprakasan, R., Campbell, B., Hopkisson, J., Johnson, I., Raine-Fenning, N., 2009. The three-dimensional ultrasonographic ovarian vascularity of women developing poor ovarian response during assisted reproduction treatment and its predictive value. Fertil. Steril. 92, 1862–1869.

Kim, C.H., Howles, C.M., Lee, H.A., 2011. The effect of transdermal testosterone gel pretreatment on controlled ovarian stimulation and IVF outcome in low responders. Fertil. Steril. 95, 679–683.

Klinkert, E.R., Broekmans, F.J., Looman, C.W., Te, V.E., 2004. A poor response in the first in vitro fertilization cycle is not necessarily related to a poor prognosis in subsequent cycles. Fertil. Steril. 81, 1247–1253.

Klinkert, E.R., Broekmans, F.J., Looman, C.W., Habbema, J.D., Te, V.E., 2005. Expected poor responders on the basis of an antral follicle count do not benefit from a higher starting dose of gonadotrophins in IVF treatment: a randomized controlled trial. Hum. Reprod. 20, 611–615.

Kol, S., Homburg, R., 2008. Change, change, change: hormonal actions depend on changes in blood levels. Hum. Reprod. 23, 1004–1006.

La Marca, A., Sighinolfi, G., Radi, D., Argento, C., Baraldi, E., Artenisio, A.C., Stabile, G., Volpe, A., 2010. Anti-Mullerian hormone (AMH) as a predictive marker in assisted reproductive technology (ART). Hum. Reprod. Update 16, 113–130.

Lekamge, D.N., Barry, M., Kolo, M., Lane, M., Gilchrist, R.B., Tremellen, K.P., 2007. Anti-Mullerian hormone as a predictor of IVF outcome. Reprod. Biomed. Online 14, 602–610.

Lenton, E.A., Sexton, L., Lee, S., Cooke, I.D., 1988. Progressive changes in LH and FSH and LH:FSH ratio in women throughout reproductive life. Maturitas 10, 35–43.

Li, H., Sun, J., 2011. Empirical research of hybridizing principal component analysis with multivariate discriminant analysis and logistic regression for business failure prediction. Expert Syst. Appl. 38, 6244–6253.

Li, Y., Yang, D., Zhang, Q., 2010. Impact of overweight and underweight on IVF treatment in Chinese women. Gynecol. Endocrinol. 26, 416–422.

Liu, K.E., Greenblatt, E.M., 2008. Elevated day 3 follicle-stimulating hormone/luteinizing hormone ratio > or = 2 is associated with higher rates of cancellation in in vitro fertilization–embryo transfer cycles. Fertil. Steril. 90, 297–301.

Mukherjee, T., Copperman, A.B., Lapinski, R., Sandler, B., Bustillo, M., Grunfeld, L., 1996. An elevated day three follicle-stimulating hormone:luteinizing hormone ratio (FSH:LH) in the presence of a normal day 3 FSH predicts a poor response to controlled ovarian hyperstimulation. Fertil. Steril. 65, 588–593.

Nielsen, M.E., Rasmussen, I.A., Kristensen, S.G., Christensen, S.T., Mollgard, K., Wreford, A.E., Byskov, A.G., Yding, A.C., 2011. In human granulosa cells from small antral follicles, androgen receptor mRNA and androgen levels in follicular fluid correlate with FSH receptor mRNA. Mol. Hum. Reprod. 17, 63–70.

Penarrubia, J., Peralta, S., Fabregues, F., Carmona, F., Casamitjana, R., Balasch, J., 2010. 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. Fertil. Steril. 94, 2590–2595.

Qin, Y., Zhao, Z., Sun, M., Geng, L., Che, L., Chen, Z.J., 2011. Association of basal serum testosterone levels with ovarian response and in vitro fertilization outcome. Reprod. Biol. Endocrinol. 9, 9.

Rotterdam ESHRE/ASRM—Sponsored PCOS Consensus Workshop Group, 2004. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). Hum. Reprod. 19, 41—47.

Scott, R.J., Hofmann, G.E., Oehninger, S., Muasher, S.J., 1990. Intercycle variability of day 3 follicle-stimulating hormone levels and its effect on stimulation quality in in vitro fertilization. Fertil. Steril. 54, 297—302.

Scott, R.T., Toner, J.P., Muasher, S.J., Oehninger, S., Robinson, S., Rosenwaks, Z., 1989. Follicle-stimulating hormone levels on cycle day 3 are predictive of in vitro fertilization outcome. Fertil. Steril. 51, 651—654.

Seifer, D.B., Lambert-Messerlian, G., Hogan, J.W., Gardiner, A.C., Blazar, A.S., Berk, C.A., 1997. Day 3 serum inhibin-B is predictive of assisted reproductive technologies outcome. Fertil. Steril. 67, 110—114.

Su, Q.J., Zang, N., Xiao, X., Deng, W., Bi, Z.Y., Cen, P., Wu, S.Z., Liang, H., 2009. Zhu cheng fen logistic hui gui mo xing zai xiao chu shu ju duo chong gong xian xing zhong de ying yong. Chin. J. Health Stat. 26, 206—208.

Sunkara, S.K., Coomarasamy, A., Arlt, W., Bhattacharya, S., 2012. Should androgen supplementation be used for poor ovarian response in IVF? Hum. Reprod. 27, 637—640.

Suzuki, T., Sasano, H., Kimura, N., Tamura, M., Fukaya, T., Yajima, A., Nagura, H., 1994. Immunohistochemical distribution of progesterone, androgen and oestrogen receptors in the human ovary during the menstrual cycle: relationship to expression of steroidogenic enzymes. Hum. Reprod. 9, 1589—1595.

Urman, B., Yakin, K., 2012. Does dehydroepiandrosterone have any benefit in fertility treatment? Curr. Opin. Obstet. Gynecol. 24, 132—135.

Vendola, K.A., Zhou, J., Adesanya, O.O., Weil, S.J., Bondy, C.A., 1998. Androgens stimulate early stages of follicular growth in the primate ovary. J. Clin. Invest. 101, 2622—2629.

Verhagen, T.E., Hendriks, D.J., Bancsi, L.F., Mol, B.W., Broekmans, F.J., 2008. The accuracy of multivariate models predicting ovarian reserve and pregnancy after in vitro fertilization: a meta-analysis. Hum. Reprod. Update 14, 95—100.

Weghofer, A., Gleicher, N., 2009. Ovarian function: a theory of relativity. Hum. Reprod. 24, 17—19.

Weghofer, A., Kim, A., Barad, D.H., Gleicher, N., 2012. The impact of androgen metabolism and FMR1 genotypes on pregnancy potential in women with dehydroepiandrosterone (DHEA) supplementation. Hum. Reprod. 27, 3287—3293.

Weil, S., Vendola, K., Zhou, J., Bondy, C.A., 1999. Androgen and follicle-stimulating hormone interactions in primate ovarian follicle development. J. Clin. Endocrinol. Metab. 84, 2951—2956.

Yakin, K., Urman, B., 2011. DHEA as a miracle drug in the treatment of poor responders; hype or hope? Hum. Reprod. 26, 1941—1944.

Zhao, X., He, Z., Mo, Y., Chen, X., Chen, Y., Yang, D., 2011. Determining the normal cut-off levels for hyperandrogenemia in Chinese women of reproductive age. Eur. J. Obstet. Gynecol. Reprod. Biol. 154, 187—191.

Declaration: The authors report no financial or commercial conflicts of interest.

Received 4 October 2013; refereed 11 February 2014; accepted 11 February 2014.