Journal of Assisted Reproduction and Genetics (2022) 39:1917–1926
https://doi.org/10.1007/s10815-022-02546-5

ASSISTED REPRODUCTION TECHNOLOGIES

Basal serum level of Δ4-androstenedione reflects the ovaries' ability to respond to stimulation in IVF cycles: setting up a new reliable index of both ovarian reserve and response

Emanuele Garzia[1,2] · Valentina Galiano[1,3] · Laura Guarnaccia[4,5] · Giovanni Marfia[2,4,5] · Giulia Murru[6] · Ellade Guermandi[1] · Jennifer Riparini[1] · Patrizia Sulpizio[1] · Anna Maria Marconi[7]

Received: 24 January 2022 / Accepted: 9 June 2022 / Published online: 27 June 2022

Abstract

Purpose Adequate androgen levels are necessary for regular follicular growth, progression beyond the pre-antral stage, and prevention of follicular atresia. The main purpose of this study was to investigate whether baseline androgen levels had a predictive value on stimulation outcomes in IVF cycles. The secondary purpose was to compare the possible predictive value of androgens with that of already known markers.

Methods The study included 91 infertile patients aged 30–45 years awaiting the first IVF cycle. All women underwent the same stimulation protocol and the same starting dose of recombinant FSH. As stimulation outcomes, the number of follicles recruited, estradiol and progesterone levels on the day of trigger, the total dose of gonadotropins administered, and the number of oocytes collected were recorded. Multiple linear regression and multivariate logistic regression were used to evaluate the significant predictive value of the variables for response to controlled ovarian stimulation (COS). By studying the reliability of different markers, an attempt was made to develop a single index with the highest predictive value.

Results Pearson's correlation revealed a statistically significant inverse correlation between oocytes collected and age ($r = -0.333$, $p < 0.001$) and a positive correlation with AMH (anti-müllerian hormone) ($r = 0.360$, $p < 0.001$), antral follicle count (AFC) ($r = 0.639$, $p < 0.001$), and androstenedione (Δ4-A) ($r = 0.359$, $p < 0.001$). No significant correlation was reported with FSH ($r = -0.133$, $p = 0.207$) and total testosterone ($r = 0.180$, $p = 0.088$). In COS good responders, the G-index (= AMH ng/mL*AFC/Δ4-A ng/dL) revealed a significantly higher level ($p < 0.001$) than AMH, AFC, and Δ4-A alone.

Conclusion Baseline serum Δ4-A, presumably crucial for ensuring a regular follicular growth, is a reliable marker of ovarian response to stimulation. Since the ovarian capacity to respond to gonadotropins does not depend exclusively on the presence of follicles, we suggest a new index, the G-index, able to contemplate both the ovarian reserve and the Δ4-A level.

Keywords Androgens · IVF · Markers of ovarian stimulation response · Controlled ovarian stimulation · Δ4-Androstenedione

Introduction

Appropriate recruitment and development of multiple follicles are key factors for the success of assisted reproductive techniques (ART). Broad evidence shows, particularly when fresh embryo transfer is planned, that optimal — rather than maximal — oocyte recruitment is the preferred outcome of controlled ovarian stimulation (COS) [1]. Indeed, an optimal number of oocytes collected correlates with high live-birth rates, while a low-response and a hyper-response are associated with higher cancellation rates, significant decrease in implantation rates, increase in obstetric risks, and increase in the risk of ovarian hyperstimulation syndrome (OHSS) [1–3]. The ovarian response to COS is certainly function of the ovarian reserve and the patient's age [4, 5], and has a great inter-individual variability [6]. Predicting ovarian response before embarking on the in vitro fertilization (IVF) cycle can not only help physicians determine the most appropriate clinical protocols and predict cycle outcomes [7], but it can also help couples to decide whether to undergo to expensive and often challenging and disappointing IVF treatments.

✉ Emanuele Garzia
emanuele.garzia@asst-santipaolocarlo.it

Extended author information available on the last page of the article

Several clinical, serological, and ultrasound markers discovered under baseline conditions have been proposed for predictive value, both as single and combined tests, such as age, serum follicle-stimulating hormone (FSH), 17β estradiol (E2), FSH/LH (luteinizing hormone) ratio, serum anti-müllerian hormone (AMH), serum inhibin B, and antral follicle count (AFC). Among these, AFC and AMH have demonstrated the best sensitivity and specificity in predicting hyper- and hypo-response, although they did not correlate closely with the number and quality of responsive follicles [8–12].

In recent decades, growing interest has been focused on the role of androgens in regulating follicular development and influencing female fertility. Firstly, the theca cell-derived androgens are converted into estradiol under the aromatase activity of granulosa cells. Furthermore, adequate androgenic activity carried out through androgen receptors (ARs) is required for regular follicular growth, for progression beyond the pre-antral stage, and for the prevention of follicular atresia. These findings are the result of different AR-knockout mouse models, along with various in vitro and in vivo studies [13–16]. Finally, it is plausible that androgens play a facilitating role on the follicle response to FSH, since there is a correlation between the expression levels of FSH receptors and ARs in women receiving androgen treatment [17]. These findings suggest a role for androgens in both the poor response to FSH stimulation in patients with diminished ovarian reserve (DOR) and the FSH hypersensitivity of women with polycystic ovary syndrome (PCOS). Indeed, serum concentrations of dehydroepiandrosterone sulfate (DHEA-S), dehydroepiandrosterone (DHEA), and total testosterone (T) tend to decrease in women with advanced age along with decline in reproductive capacity [18, 19], and similar evidence can be found in primary ovarian failure (POF) [20] and DOR [21]. On the other hand, PCOS is characterized by the ovarian overproduction of androgens [22], whose circulating levels are positively correlated with the antral follicle count [23, 24].

The possibility of correlating the serum androgen level with the stimulation parameters of IVF has attracted some research groups. Low baseline T levels were initially associated with poor pregnancy rate after IVF in women with normal ovarian reserve by Frattarelli et al. [25], who subsequently denied this predictive value [26]. Qin and colleagues proposed basal T levels as a marker of ovarian response in DOR women [27] and other researchers correlated them with the number of follicles > 14 mm on human chorionic gonadotropin (HCG) day and total gonadotropin dose [28].

A significant weakness of previous studies comparing baseline androgen levels with ovarian response, stimulation outcomes, and pregnancy rates is the lack of a uniform stimulation protocol. In fact, the use of different stimulation procedures necessarily affects IVF cycle results. In addition, the inclusion of pregnancy rate among the study outcomes increased the number of variables that could influence the results, such as semen parameters and uterine cavity characteristics.

The aim of this study was to investigate whether baseline androgen levels had a predictive value for ovarian response and stimulation outcomes in IVF cycles in patients undergoing the same stimulation protocol and starting dose of gonadotropin. The secondary outcome was to compare the possible predictive value of androgens with that of the known markers of ovarian reserve and response and to assess whether a more reliable marker could be developed.

Material and methods

Study population

103 infertile patients undergoing the first IVF cycle at the Reproductive Medicine Unit of the San Paolo University Hospital in Milano were prospectively recruited and 91 were eligible according to the inclusion and exclusion criteria.

The inclusion criterion was the age of the patients ranging from 30 to 45 years: since the same stimulation protocol and starting dose were used in all patients, we decided to exclude younger patients to contain the risk of OHSS. The exclusion criteria were the history of ovarian or adnexal surgery, the diagnosis of pelvic endometriosis, the suspicious findings of ovarian malignancy, and the presence of endocrine disorders such as diabetes mellitus, hyper-prolactinemia, thyroid dysfunction, congenital adrenal hyperplasia, Cushing's syndrome, and adrenal insufficiency.

None of the participants had any systemic disease and/or reported the use of lipid-modulating drugs or other substances that could interfere with the normal function of the hypothalamus-pituitary–gonadal axis.

The 91 patients presenting the inclusion criteria respected a confidence level of 99%, with a confidence interval of 5%. We divided the population by age into a group below and a group above 35 years of age to highlight possible differences in the values of main markers of ovarian reserve and stimulation outcomes. Regarding the subclassification of patients for age, the distribution of the two groups presented an alpha value of 0.05, a power of 80%, and effect size of 0.65, considered medium/large by Cohen's indications [29].

Pre-stimulation assays

Between 3rd and 5th day of a spontaneous menstrual cycle within 3 months before the fresh IVF cycle, overnight fasting blood samples were collected. The blood samples were obtained through an intravenous catheter placed in the forearm for the determination of fasting blood glucose, insulin, gonadotropins, E2, T, DHEA-S, Δ4-androstenedione (Δ4-A), AMH,

progesterone, and sex hormone binding globulin (SHBG). On the same day, transvaginal sonography was performed to obtain AFC. As recommended [30], 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 anthropometric, endocrine, and metabolic characteristics of the enrolled women are shown in Table 1.

The Institutional Review Board of the San Paolo Hospital Medical School of Milan approved the treatment protocol and signed informed consents were obtained from all patients before commencing data collection.

Ovarian stimulation protocol

All women underwent COS with gonadotropin agonist (GnRHa) long protocol. All participants received folic acid 400 µg/day before initiation the induction cycle. The standard protocol started with the daily subcutaneous injection of a short-acting GnRH analogue (Triptorelin 0.1 mg, Fertipeptil; Ferring, Switzerland) from the previous midluteal phase until the day of HCG administration. When satisfactory pituitary desensitization was achieved, stimulation with exogenous gonadotropins (Gonal-F; Serono, Switzerland) at a dose of 150 IU was started regardless of age and other ovarian response markers. At the 6th day of stimulation, the dose could be adjusted according to ovarian response which was assessed by transvaginal ultrasound (TV-US) and serum total E2 assay. Each dose change was carefully recorded and serial ovarian response monitoring was performed and recorded over the following days.

The criteria for HCG (Gonasi 5000 Serono, Switzerland) administration were the presence of at least one follicle > 18 mm in diameter with a consistent rise in serum E2. Oocyte aspiration was performed 36–38 h later under transvaginal sonography guide.

According to the protocols used in the Reproductive Medicine Unit of San Paolo Hospital in Milan, we did not proceed to fresh embryo transfer, resorting to embryo freezing, in case of plasma progesterone higher than 2 ng/mL on the day of HCG administration or in case of OHSS risk. The risk of OHSS was derived from the combined assessment of several factors such as the number of follicles recruited, the number of oocytes collected, the level of estradiol on the day of HCG administration, and the somatic characteristics of the women.

Table 1 Clinical, biochemical, and biological parameters of study population

Parameters	Median (IQR)
Age (years)	37 (34–40)
BMI (kg/m^2)	22.35 (20.19–24.91)
FSH (mIU/mL)	7.30 (5.90–8.60)
LH (mIU/mL)	5.30 (4–6.50)
E2 (pg/mL)	50 (34–74)
AMH (ng/mL)	1.60 (0.83–2.40)
Δ4-A (ng/mL)	1.80 (1.40–2.67)
DHEA-S (µg/dL)	148 (119–221)
T (ng/mL)	0.34 (0.24–0.50)
SHBG (nmol/L)	63 (47.60–78)
FAI (%)	2.03 (1.03–3.22)
AFC (n°)	11 (8–14)
Total IU administered	1950 (1500–2500)
Duration of stimulation (days)	12 (11–13)
N° follicles > 16 mm	7 (5–10)
N° follicles between 10 and 16 mm	6 (3–10)
N° total follicles	15 (10–21)
17β-Estradiol at induction (pg/mL)	1389 (833–2154)
Progesterone at induction (ng/mL)	1.30 (0.94–1.78)
N° oocytes	9 (5–12)

Data are listed as median (IQR). The following parameters were assessed under basal conditions: *BMI*, body mass index; *FSH*, follicle-stimulating hormone; *LH*, luteinizing hormone; *E2*, 17β-estradiol; *AMH*, anti-müllerian hormone; *Δ4-A*, Δ4-androstenedione; *DHEA-S*, dehydroepiandrosterone sulfate; *T*, total testosterone; *SHBG*, sex hormone binding globulin; *FAI*, free androgen index; *AFC*, antral follicle count

Stimulation outcomes

The primary outcome of the study, in order to evaluate the response to ovarian stimulation, was the number of oocytes retrieved. We also considered the number of follicles > 16 mm [31], the total number of follicles, the E2 and P level on the day of HCG, and the total dose of gonadotropins administered.

Assay method

FSH, LH, E2, T, Δ4-A, DHEA-S, and progesterone levels were measured using chemiluminescence immunoassay (Vitros 5600; Ortho Clinical Diagnostics, New Jersey, USA). Serum SHBG and AMH levels were measured using electrochemiluminescence (Immulite 2000 Xpi; ORM Immunoassay System, Siemens Medical Solutions-Diagnostics-USA—ex DPC Instrument Systems Division—New Jersey, USA). The intra-assay and inter-assay coefficients of variation (CVs) for each aforementioned biochemical or hormonal parameter were evaluated, and the values of the CVs were in any case respectively lower than 6 and 11%.

Statistical analysis

Statistical analysis was carried out using IBM SPSS Statistics 26.0 software. Continuous variables are presented

as median and interquartile range (IQR). Categorical variables are presented as counts and percentages. Demographic, clinical, and biochemical parameters were tested for normality using the Kolmogorov–Smirnov and Shapiro–Wilk tests. When normally distributed, a two-way ANOVA has been used; otherwise, when not normally distributed, the variables were compared by the U Mann–Whitney test for independent samples. The Pearson correlation test was used to assess the univariate association between variables. The primary clinical endpoint of our study was the production of more than 10 oocytes after stimulation, clinically considered a good ovarian stimulation response [32]. The visual binning was used to create the 3 equal width categories of analysis and evaluate the trend of our outcome. Multiple linear regression analysis was used to evaluate the predictive value of variables for response to COS. To evaluate the significant independent predictors for good responders, multivariate logistic regression with stepwise procedure was performed. To compare the predictive abilities of different biomarkers, receiver operating characteristic (ROC) curves were constructed to identify the global accuracy (area under the curve [AUC]) of our parameters of interest. The optimal cut-off was calculated according to the maximum value of Youden index and the hazard ratio was calculated using the Cox proportional hazard model. Data acquisition was performed blindly. The tests were considered statistically significant when $p < 0.05$.

Results

A total of 91 women with a median age of 37 years (IQR: 34–40) were recruited for this study. The main characteristics of the study population, including their demographic, clinical, and biochemical parameters, are shown in Table 1. After the first 5 days of stimulation, the gonadotropin dose was reduced in 19 cases and increased in 42. Thirty of 91 women did not require drug dose adjustment. As a result of the ovarian stimulation, all patients were submitted to the oocytes retrieval. In one case, no oocyte was picked up. Overall, 843 oocytes were collected with a median of 9 oocytes per woman (IQR: 5–12). Fourteen women did not continue with fresh embryo transfer and the embryos were cryo-preserved: 9 of them due to OHSS risk, 8 due to increased serum progesterone in the late follicular phase, and 3 women experienced both these events. Dividing the population by age into a group below and a group above 35 years, we reported no significant differences in the main ovarian reserve markers. On the other hand, total number of follicles, number of follicles larger than 16 mm, progesterone at induction, and number of oocytes retrieved were statistically significant between the two groups (Table 2).

Considering the number of oocytes as the most reliable indicator of COS response, Pearson's correlation revealed a statistically significant inverse correlation with age ($r = -0.333$, $p < 0.001$) and a positive correlation with AMH ($r = 0.360$, $p < 0.001$), AFC ($r = 0.639$, $p < 0.001$), and Δ4-A ($r = 0.359$, $p < 0.001$). No significant correlation was demonstrated with FSH ($r = -0.133$, $p = 0.207$) and testosterone ($r = 0.180$, $p = 0.088$) (Fig. 1). Similar results were found when comparing the other stimulation outcomes, particularly the number of follicles > 16 mm, the total number of follicles, the E2 level on the day of HCG, and the total dose of gonadotropins administered, with the markers (data not shown). AMH and AFC levels are largely reported as reliable biomarkers in predicting both good and poor ovarian response in women undergoing IVF, while the reliability of FSH levels has been reconsidered over time and is now certainly lower. Since AMH, AFC, and Δ4-A have different origins, it is reasonable that their predictive significance reflects different aspects of ovarian reserve and response to stimulation. Figure 2 shows the correspondence between the values of the markers in the different ranges and the number of oocytes collected.

It may therefore be useful to combine this information and evaluate it at the same time. With the intention of creating a single index with the highest predictive value, we developed the G-index. In order to enhance a well-known datum as ovarian reserve, expressed through AMH and AFC, with another datum, the ability of the ovary to determine follicular growth after stimulation, here expressed through Δ4-A, we thought of dividing the product of the reserve markers by the response marker:

$$G - index = AMH * AFC / \Delta 4 - A$$

To assess good response to COS, we selected an ovarian response of > 10 retrieved oocytes as a criterion [32]. The analysis of G-index distribution revealed a statistically significant higher level in COS good responders ($p < 0.001$, Fig. 3). ROC analyses showed that, in women with an age > 35 years, the G-index has the highest sensitivity and specificity in predicting a good ovarian response (AUC: 0.753) compared to AMH, AFC, and Δ4-A alone (Fig. 4). Furthermore, multiple linear regression with stepwise method revealed that the G-index is the only significant predictive parameter in patients aged higher than 35 years ($p = 0.001$, Table 3).

Interestingly, the Cox proportional hazard model (days of stimulation used as time) showed that, dividing the G-index into three groups (33.3%), women produced a progressive significantly higher number of oocytes (Fig. 5A), with a significant higher hazard ratio of producing more than 10 oocytes [Exp(B) = 5.104 (95% CI: 1.416–18.400, $p = 0.013$)], when G-index > 13.62 (Fig. 5B).

Table 2 Demographic and clinical characteristics of the patients classified by age

Parameters	Age < 35 years ($n = 24$)	Age > 35 years ($n = 67$)	p-value
BMI (kg/m^2)	22.64 (20.96–24.53)	22.20 (20.08–26.03)	0.825
FSH (mIU/mL)	7.63 (6.80–9.20)	7.11 (5.50–8.60)	0.096
LH (mIU/mL)	5.35 (4.55–7.15)	5.30 (4.00–6.20)	0.326
17β-Estradiol (pg/mL)	48.5 (27.7–74.0)	50.0 (34.0–79.0)	0.801
AMH (ng/mL)	2.09 (1.21–3.04)	1.42 (0.73–2.20)	0.092
Δ4-Androstenedione (ng/mL)	1.90 (1.55–3.05)	1.80 (1.40–2.48)	0.148
DHEA-S (μg/dL)	157 (105–248)	147 (122–219)	0.808
Total testosterone (ng/mL)	0.33 (0.19–0.51)	0.37 (0.24–0.50)	0.871
SHBG (nmol/L)	64 (46–94)	61 (48–77)	0.885
FAI (%)	2.03 (0.91–3.29)	2.03 (1.17–3.17)	0.709
AFC (n°)	12 (9–16)	10 (7–14)	0.130
Total IU administered	1813 (1406–2200)	1950 (1650–2625)	0.149
Duration of stimulation (days)	13 (10–13)	12 (11–13)	0.984
N° follicles > 16 mm	9 (7–13)	6 (5–10)	**0.032***
N° follicles between 10 and 16 mm	7 (5–13)	6 (3–10)	0.132
N° total follicles	18 (12–24)	14 (8–20)	**0.013***
17β-Estradiol at induction (pg/mL)	1584 (1062–2239)	1329 (798–2126)	0.216
Progesterone at induction (ng/mL)	1.60 (1.21–1.86)	1.20 (0.85–1.64)	**0.050***
N° oocytes	11 (8–17)	8 (4–11)	**0.011***

Data are listed as median (IQR). *BMI*, body mass index; *FSH*, follicle-stimulating hormone; *LH*, luteinizing hormone; *AMH*, anti-müllerian hormone; *DHEA-S*, dehydroepiandrosterone sulfate; *SHBG*, sex hormone binding globulin; *FAI*, free androgen index; *AFC*, antral follicle count

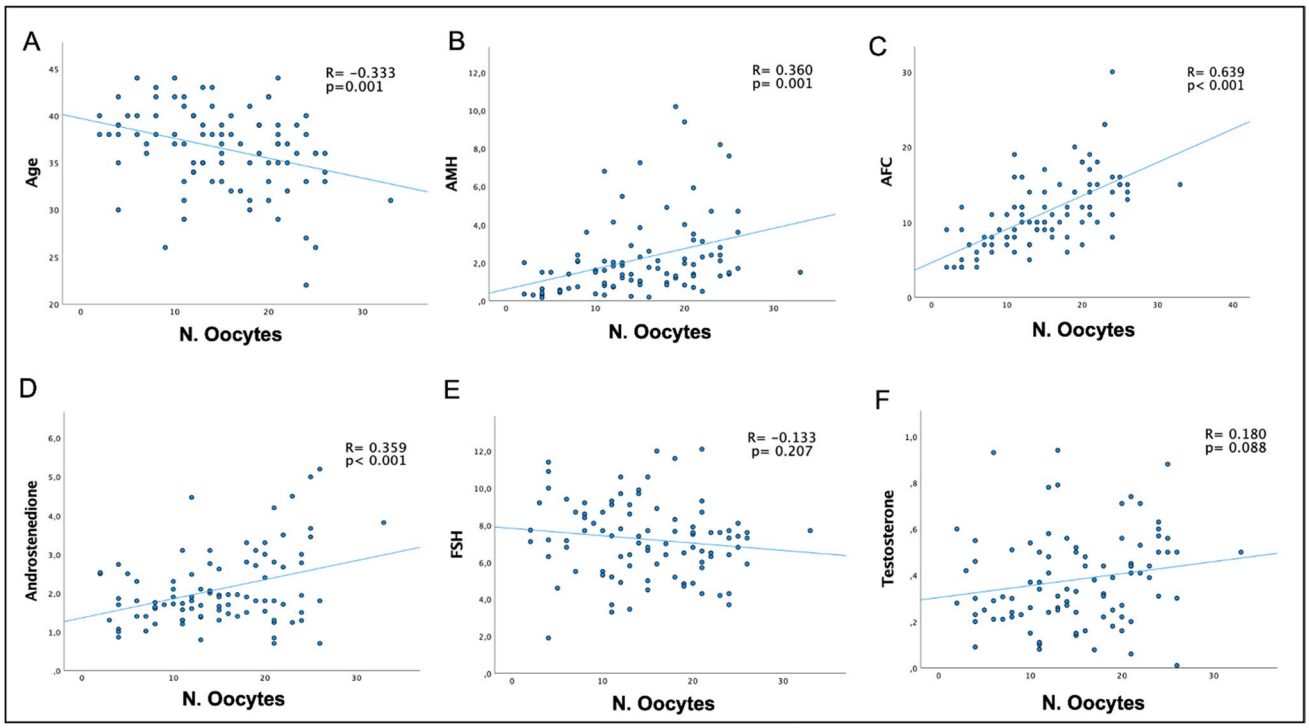


Fig. 1 Correlation between N° of oocytes and age (**A**), AMH (**B**), AFC (**C**), Δ4-androstenedione (**D**), FSH (**E**), and total testosterone (**F**). Scatter plots and fitted regression line are shown in each figure. Pearson correlation was performed for statistical analysis. Exact p-values or correlation coefficients are reported in the figure

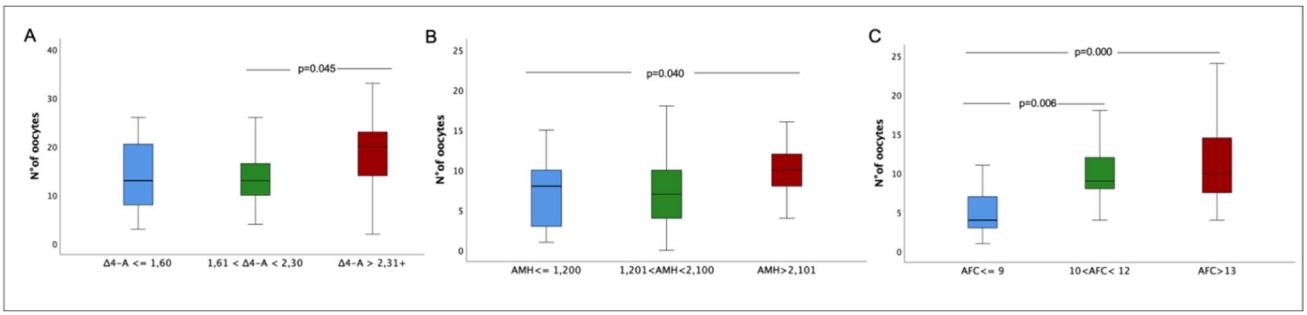


Fig. 2 The assessment of the number of retrieved oocytes was performed in 3 groups of Δ4-androstenedione (**A**), AMH (**B**), and AFC (**C**). The visual binning was used to create the 3 equal width categories of analysis and to evaluate the trend of our outcome and the statistical significance between the 3 groups identified. Exact *p*-values between groups are indicated

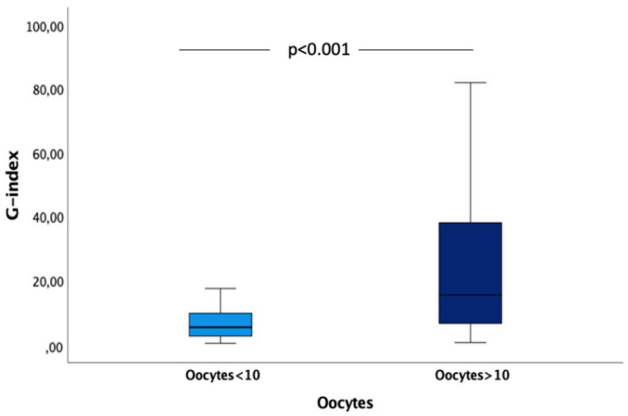


Fig. 3 The value of G-index was assessed in 2 study groups: the first one composed by patients defined as less responders (oocytes < 10) and those considered as more responders (oocytes > 10). The difference resulted statistically significant, with $p < 0.001$, confirming the predictive value of G-index

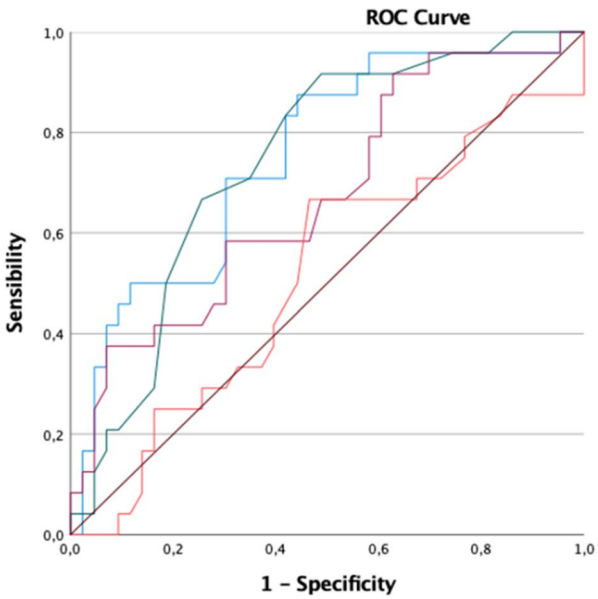
Discussion

The main purpose of the study was to investigate the possible predictive value of baseline androgen levels for ovarian stimulation outcomes in IVF cycles and to compare it with the known markers of ovarian response.

Unlike DHEA-S, T, SHBG, and FAI, our data suggest Δ4-A levels as a marker of response to ovarian stimulation. Comparing the reliability of this biomarker to that of already established predictors in daily use, it performs similarly to AMH and AFC and clearly superior to FSH. Its significance logically deserves to be better understood and validated.

The potential role of Δ4-A as a biomarker in ART cycles is not new. In a prospective case–control study with a study group of 46 women with PCOS, the response of Δ4-A to low-dose rFSH was more strongly associated with the number of selected follicles than with serum E2. Therefore,

Fig. 4 Receiver operating characteristic (ROC) curve was designed to compare the predictive abilities of different biomarkers, calculating the global accuracy or area under the curve (AUC) of our parameters of interest. As showed in the table, G-index showed the highest AUC value, proving to have the highest sensitivity and specificity in predicting ovarian response



Variable	AUC
G-Index	0.753
AMH	0.672
AFC	0.514
Androstenedione	0.745

Table 3 Multiple linear regression analysis of predictive determinants for response to COS

	Coefficients				
	Unstandardized coefficients		Standardized coefficients		
	β	Std. error	*β*	*t*	*p*-value
	0.213	0.070	-	3.049	0.003
G-index	0.009	0.003	0.388	3.351	**0.001***

Excluded variables	*p*-value
BMI (kg/m^2)	0.082
FSH (mUI/mL)	0.648
LH (mUI/mL)	0.054
17β-Estradiol (pg/mL)	0.051
AMH (ng/mL)	0.984
Δ4-Androstenedione (ng/mL)	0.624
DHEA-S (μg/dL)	0.402
Total testosterone (ng/mL)	0.619
SHBG (nmol/L)	0.181
FAI (%)	0.378
AFC (n°)	0.302

BMI, body mass index; *FSH*, follicle-stimulating hormone; *LH*, lutenizing hormone; *AMH*, anti-müllerian hormone; *DHEA-S*, dehydroepiandrosterone sulfate; *SHBG*, sex hormone binding globulin; *FAI*, free androgen index; *AFC*, antral follicle count

the authors proposed the stimulated Δ4-A level as an early marker of the ovarian response [33].

Moreover, Menet and coworkers demonstrated that the first injections of rFSH in ART cycles are likely to stimulate the Δ4 pathway involving two main enzymes (17α-hydroxylase and 17–20 lyase). Since Δ4-A and 17-OH progesterone are the two major steroids in the Δ4 pathway, it has been suggested that they were of interest to ensure proper management of ovarian induction and that they could influence follicular growth. Therefore, the level of circulating Δ4-A can be considered a function of the ovarian response. From this point of view, T seems to be of lesser interest as it derives from the peripheral conversion of Δ5 androgens, which become predominant only during the luteal phase [34].

The action of androgens on the female ovary is not mediated only by estrogen receptors, through the conversion of androgens into estrogens and 3ß-diol, but also and more directly through AR [35]. Inhibition of androgen activity, via antiandrogen antibodies or an androgen receptor antagonist such as bicalutamide, significantly suppresses follicular growth in in vitro cultures of pre-antral mouse follicles. Treatment with dihydrotestosterone (DHT) is able to restore follicular growth [36]. These effects are not related to the role of androgens as estrogen precursors, since the addition of estrogens or an aromatase inhibitor (fadrozole) does not affect follicular growth [37]. The role of androgens on ovarian function was also elegantly highlighted by Gleicher through the case of a patient with primary adrenal insufficiency (Addison's disease) who, despite being treated with glucocorticoids, developed secondary ovarian insufficiency due to adrenal androgen deficiency [38]. Therefore, consistent evidence supports the idea that androgens are essential factors in the development of the early follicular phase. Based on these considerations and with the aim of increasing the follicular pool, some researchers have integrated the treatment of DOR women in ART cycles with androgens,

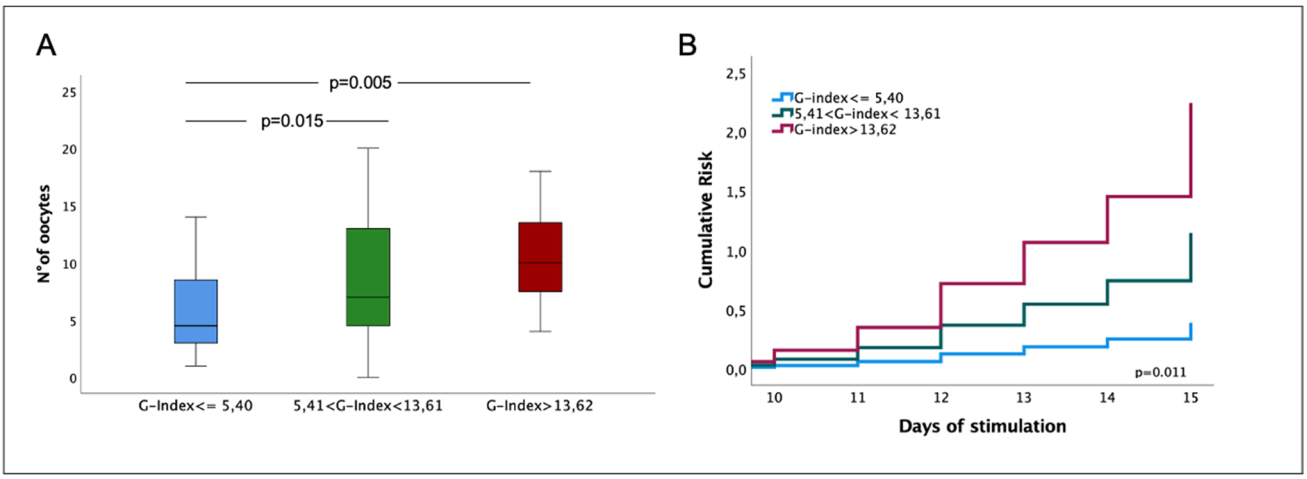


Fig. 5 **A** The number of retrieved oocytes was assessed in 3 study groups, defined accordingly to G-index value. By this visual binning, it was possible to define 2 optimal and statistically significant cut-off able to validly forecast patients' ovarian response. Exact *p*-values between groups are indicated. **B** Cumulative risk for the production of more than 10 oocytes in patients aged > 35 years, grouped for cut-off values of G-index. Statistical analysis was performed by Cox regression. Exact *p*-value is reported

reporting a significant improvement in COS output [39]. On this topic, there have been several studies with initially promising findings concerning not only the oocyte yield but also the quality of eggs and embryos and IVF pregnancy rates [40–44]. These results have not been confirmed by well-designed randomized controlled trials (RCTs) [45, 46].

Among androgens, the role of particolar interest played by Δ4-A can also be inferred from the increase in its concentration at midcycle. Since, in fact, the adrenal contribution to peripheral androgen levels is relatively constant during the spontaneous menstrual cycle, it is clear that it is the ovarian contribution of Δ4-A that makes the difference. On this basis, it has already been suggested as a reliable marker of the initial ovarian response to gonadotropins [33, 47]. What appears likely is that Δ4-A may be involved in the regulation of follicular growth, progression beyond the pre-antral stage, and prevention of follicular atresia, as well as facilitation of the follicular response to FSH [17].

Not only do our results show a statistically significant correlation between Δ4-A and the number of oocytes retrieved, proposing it as a predictor of reliability similar to AMH and AFC. The index derived from combining these markers, the G-index, accurately reflects response to stimulation and, in women > 35 years of age, has the greatest sensitivity and specificity in predicting good ovarian response compared with AMH, AFC, and Δ4-A alone. This may be because traditional markers primarily refer to the number of follicles present, not taking into account that some will not respond to stimulation. In fact, the management of ovarian stimulation is currently performed essentially on the basis of values expressing the ovarian content of pre-antral and antral follicles (AMH and AFC). The ovarian response to stimulation is a complex and dynamic process, probably dependent on multiple factors. Among these, the ovarian reserve is certainly a priority, but the ovarian ability to respond to gonadotropins does not necessarily depend exclusively on the presence of follicles.

This is probably the reason why no ovarian response marker has been proved to be exempt from a false positive rate of at least 10–20% [9, 12]. Here, we developed a new index, the G-index, capable of contemplating not only the ovarian reserve but also a probably crucial element in order to guarantee a regular early follicular growth, namely the baseline serum level of Δ4-A. The contribution that Δ4-A presumably gives to the prediction on the outcomes of the stimulation cycle lies in the individual's ability to respond to gonadotropins. As encouraging as our results are, they certainly need to be validated in studies with larger populations. However, they pave the way for a reconsideration of the significance of the ovarian response markers in ART cycles.

Our study has some potential causes for bias and limitations, first of all, the small population under study. Furthermore, although the same stimulation regimen and starting dose of gonadotropin was used, individualization of treatment certainly influenced the results. In fact, after the first few days, the therapy was adapted to the response outputs. This was considered as an insurmountable bias since, for obvious ethical reasons, women's health and the outcome of the stimulation cycle deserved to be guaranteed. However, we believe that the value of our conclusions is not strongly affected by this limitation.

In summary, our results suggest that baseline Δ4-A levels may serve as a predictor of stimulation outcomes in IVF cycles, the reliability of which seems comparable to the best markers currently used. Furthermore, it could positively contribute to the generation of a new index, which we have suggested consisting of Δ4-A, AMH, and AFC and called G-index, which better meets clinical expectations of predicting response to gonadotropins.

Author contribution E.G. conceived the presented idea and managed the project. V.G., G.M., G.Ma, E.Gu, and J.R. contributed to acquisition of data. A.M.M. and P.S. helped supervise the project. L.G. helped analyze the data. V.G., L.G., and E.G. drafted and finalized the manuscript. All authors agreed with the submission of the manuscript and gave approval to the final version to be published.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Sunkara SK, et al. 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.
2. Verberg MF, et al. Why do couples drop-out from IVF treatment? A prospective cohort study. Hum Reprod. 2008;23(9):2050–5.
3. Baker VL, et al. Association of number of retrieved oocytes with live birth rate and birth weight: an analysis of 231,815 cycles of in vitro fertilization. Fertil Steril. 2015;103(4):931-938.e2.
4. Fauser BC, Diedrich K, Devroey P. Predictors of ovarian response: progress towards individualized treatment in ovulation induction and ovarian stimulation. Hum Reprod Update. 2008;14(1):1–14.
5. de Boer EJ, et al. A low number of retrieved oocytes at in vitro fertilization treatment is predictive of early menopause. Fertil Steril. 2002;77(5):978–85.
6. te Velde ER, Pearson PL. The variability of female reproductive ageing. Hum Reprod Update. 2002;8(2):141–54.
7. Klinkert ER, et al. A poor response in the first in vitro fertilization cycle is not necessarily related to a poor prognosis in subsequent cycles. Fertil Steril. 2004;81(5):1247–53.
8. Broer SL, et al. Prediction of an excessive response in in vitro fertilization from patient characteristics and ovarian reserve tests and comparison in subgroups: an individual patient data meta-analysis. Fertil Steril. 2013;100(2):420-9.e7.
9. Broekmans FJ, et al. A systematic review of tests predicting ovarian reserve and IVF outcome. Hum Reprod Update. 2006;12(6):685–718.
10. La Marca A, et al. Anti-Mullerian hormone (AMH) as a predictive marker in assisted reproductive technology (ART). Hum Reprod Update. 2010;16(2):113–30.

11. Verhagen TE, et al. The accuracy of multivariate models predicting ovarian reserve and pregnancy after in vitro fertilization: a meta-analysis. Hum Reprod Update. 2008;14(2):95–100.
12. Ferraretti AP, et al. 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.
13. Kwee J, et al. Comparison of endocrine tests with respect to their predictive value on the outcome of ovarian hyperstimulation in IVF treatment: results of a prospective randomized study. Hum Reprod. 2003;18(7):1422–7.
14. Yang MY, Fortune JE. Testosterone stimulates the primary to secondary follicle transition in bovine follicles in vitro. Biol Reprod. 2006;75(6):924–32.
15. Walters KA, Allan CM, Handelsman DJ. Androgen actions and the ovary. Biol Reprod. 2008;78(3):380–9.
16. Sen A, Hammes SR. Granulosa cell-specific androgen receptors are critical regulators of ovarian development and function. Mol Endocrinol. 2010;24(7):1393–403.
17. Nielsen ME, et al. 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. 2011;17(1):63–70.
18. Burger HG. Androgen production in women. Fertil Steril. 2002;77(Suppl 4):S3-5.
19. Meldrum DR, et al. Role of decreased androgens in the ovarian response to stimulation in older women. Fertil Steril. 2013;99(1):5–11.
20. van der Stege JG, et al. Decreased androgen concentrations and diminished general and sexual well-being in women with premature ovarian failure. Menopause. 2008;15(1):23–31.
21. Gleicher N, et al. Hypoandrogenism in association with diminished functional ovarian reserve. Hum Reprod. 2013;28(4):1084–91.
22. Norman RJ. Hyperandrogenaemia and the ovary. Mol Cell Endocrinol. 2002;191(1):113–9.
23. Pigny P, et al. Elevated serum level of anti-mullerian hormone in patients with polycystic ovary syndrome: relationship to the ovarian follicle excess and to the follicular arrest. J Clin Endocrinol Metab. 2003;88(12):5957–62.
24. Sen A, et al. Endocrine autoimmune diseases and female infertility. Nat Rev Endocrinol. 2014;10(1):37–50.
25. Frattarelli JL, Peterson EH. Effect of androgen levels on in vitro fertilization cycles. Fertil Steril. 2004;81(6):1713–4.
26. Frattarelli JL, Gerber MD. Basal and cycle androgen levels correlate with in vitro fertilization stimulation parameters but do not predict pregnancy outcome. Fertil Steril. 2006;86(1):51–7.
27. Qin Y, et al. Association of basal serum testosterone levels with ovarian response and in vitro fertilization outcome. Reprod Biol Endocrinol. 2011;9:9.
28. Sun B, et al. Basal serum testosterone levels correlate with ovarian response but do not predict pregnancy outcome in non-PCOS women undergoing IVF. J Assist Reprod Genet. 2014;31(7):829–35.
29. Sullivan GM, Feinn R. Using effect size—or why the P value is not enough. J Grad Med Educ. 2012;4(3):279–82.
30. Broekmans FJ, et al. The antral follicle count: practical recommendations for better standardization. Fertil Steril. 2010;94(3):1044–51.
31. Abbara A, et al. Follicle size on day of trigger most likely to yield a mature oocyte. Front Endocrinol (Lausanne). 2018;9:193.
32. Ji J, et al. The optimum number of oocytes in IVF treatment: an analysis of 2455 cycles in China. Hum Reprod. 2013;28(10):2728–34.
33. Ozyurek ES, Yoldemir T, Artar G. Androstenedione response to recombinant human FSH is the most valid predictor of the number of selected follicles in polycystic ovarian syndrome: (a case-control study). J Ovarian Res. 2017;10(1):34.
34. Menet MC, et al. rFSH in medically assisted procreation: evidence for ovarian follicular hyperplasia and interest of mass spectrometry to measure 17-hydroxyprogesterone and Δ4-androstenedione in serum. Mol Cell Endocrinol. 2017;450:105–12.
35. Walters KA, Handelsman DJ. Role of androgens in the ovary. Mol Cell Endocrinol. 2018;465:36–47.
36. Murray AA, et al. Effect of androgens on the development of mouse follicles growing in vitro. J Reprod Fertil. 1998;113(1):27–33.
37. Wang H, et al. Effect of adrenal and ovarian androgens on type 4 follicles unresponsive to FSH in immature mice. Endocrinology. 2001;142(11):4930–6.
38. Gleicher N, et al. The importance of adrenal hypoandrogenism in infertile women with low functional ovarian reserve: a case study of associated adrenal insufficiency. Reprod Biol Endocrinol. 2016;14:23.
39. Casson PR, et al. Dehydroepiandrosterone supplementation augments ovarian stimulation in poor responders: a case series. Hum Reprod. 2000;15(10):2129–32.
40. Barad D, Gleicher N. Effect of dehydroepiandrosterone on oocyte and embryo yields, embryo grade and cell number in IVF. Hum Reprod. 2006;21(11):2845–9.
41. Gleicher N, Barad DH. Dehydroepiandrosterone (DHEA) supplementation in diminished ovarian reserve (DOR). Reprod Biol Endocrinol. 2011;9(1):67.
42. Mamas L, Mamas E. Premature ovarian failure and dehydroepiandrosterone. Fertil Steril. 2009;91(2):644–6.
43 Sunkara SK, Coomarasamy A. Androgen pretreatment in poor responders undergoing controlled ovarian stimulation and in vitro fertilization treatment. Fertil Steril. 2011;95(8):e73-4 (author reply e75).
44. Barad D, Brill H, Gleicher N. Update on the use of dehydroepiandrosterone supplementation among women with diminished ovarian function. J Assist Reprod Genet. 2007;24(12):629–34.
45. Yeung TW, et al. A randomized double-blinded placebo-controlled trial on the effect of dehydroepiandrosterone for 16 weeks on ovarian response markers in women with primary ovarian insufficiency. J Clin Endocrinol Metab. 2013;98(1):380–8.
46. Narkwichean A, et al. Efficacy of dehydroepiandrosterone (DHEA) to overcome the effect of ovarian ageing (DITTO): a proof of principle double blinded randomized placebo controlled trial. Eur J Obstet Gynecol Reprod Biol. 2017;218:39–48.
47. Ferrario M, et al. Ovarian and adrenal androgens may be useful markers to predict oocyte competence and embryo development in older women. Gynecol Endocrinol. 2015;31(2):125–30.

Authors and Affiliations

Emanuele Garzia[1,2] · Valentina Galiano[1,3] · Laura Guarnaccia[4,5] · Giovanni Marfia[2,4,5] · Giulia Murru[6] · Ellade Guermandi[1] · Jennifer Riparini[1] · Patrizia Sulpizio[1] · Anna Maria Marconi[7]

Valentina Galiano
valentina.galiano@asst-santipaolocarlo.it

Laura Guarnaccia
laura.guarnaccia@policlinico.mi.it

Giovanni Marfia
giovanni.marfia@policlinico.mi.it

Giulia Murru
giuliamurru87@gmail.com

Ellade Guermandi
ellade.guermandi@asst-santipaolocarlo.it

Jennifer Riparini
jennifer.riparini@asst-santipaolocarlo.it

Patrizia Sulpizio
patrizia.sulpizio@asst-santipaolocarlo.it

Anna Maria Marconi
annamaria.marconi@unimi.it

[1] Reproductive Medicine Unit, Department of Mother and Child, San Paolo Hospital Medical School, ASST Santi Paolo e Carlo, 20142 via di Rudinì, 8, Milan, Italy

[2] Istituto Di Medicina Aerospaziale "A. Mosso," Aeronautica Militare, Milan, Italy

[3] Gynecology and Obstetrics Unit, San Carlo Borromeo Hospital, ASST Santi Paolo e Carlo, Milan, Italy

[4] Laboratory of Experimental Neurosurgery and Cell Therapy, Neurosurgery Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

[5] Aldo Ravelli Research Center, Milan, Italy

[6] Gynecology and Obstetrics Unit, ASST Rhodense, Ospedale G. Salvini, Garbagnate Milanese, Italy

[7] Gynecology and Obstetrics Unit, Department of Mother and Child, Department of Health Sciences, San Paolo Hospital Medical School, University of Milan, Milan, Italy