Value of the Ultrasound in the Study of Ovarian Reserve for Prediction of Oocyte Recovery

**Abstract**

**Objectives** To identify which methods used in the assessment of the ovarian reserve are exclusive or complementary to identify the best response to follicle development.

**Methods** Retrospective cohort study, involving patients undergoing assisted reproduction treatment at the Instituto de Medicina Reprodutiva e Fetal, from April 2009 to July 2014. Age, biochemical tests, and ultrasound were assessed. The data were analyzed to predict the follicular development and the relation between them, using, for statistical analysis, Statistical Package for Social Sciences software.

**Results** Out of the 293 couples included, 50.2% presented infertility by ovarian factor. Considering the age as the main variable, a significant negative correlation with the volume of both ovaries was observed (right ovary, \( r = 0.21 \); left ovary, \( r = -0.22 \); both \( p < 0.0001 \)), and with the antral follicle count (right ovary, \( r = -0.38 \); left ovary, \( r = -0.47 \); both \( p < 0.0001 \)). Considering the antral follicle count as the main variable, a significant positive correlation with the total recruited oocytes was observed. When we correlated the antral follicle count with the recruited follicles larger than 18 mm, we observed that, with a cutoff of 12 antral follicles, there is a positive predictive value of 99%, and an area under the ROC curve of 0.76.

**Conclusion** We concluded from our study that age and antral follicle count are effective predictors of ovarian response in cycles of assisted reproduction. The ovarian volume, as well as the anti-Müllerian hormone dosage, seem to be adequate markers of the ovarian reserve.

**Keywords**
- in vitro fertilization
- ultrasound
- ovarian function tests
- ovarian reserve

**Resumo**

**Objetivo** Identificar quais métodos utilizados na avaliação da reserva ovariana são excludentes ou complementares na identificação da melhor resposta ao desenvolvimento folicular.

**Métodos** Estudo retrospectivo de coorte, que envolveu pacientes em tratamento de reprodução assistida no Instituto de Medicina Reprodutiva de abril de 2009 a julho de 2014. Idade, testes bioquímicos e ecografia foram avaliados. Os dados foram analisados para prever o desenvolvimento folicular e a relação entre eles, utilizando, para análise estatística, Statistical Package for Social Sciences software.

**Resultados** De 293 casais incluídos, 50,2% apresentaram infertilidade por fator ovariano. Considerando a idade como a variável principal, observou-se uma correlação negativa significativa com o volume de ambos os ovários (ovário direito, \( r = 0,21 \); ovário esquerdo, \( r = -0,22 \); ambos \( p < 0,0001 \)), e com o contagem de folículos antrais (ovário direito, \( r = -0,38 \); ovário esquerdo, \( r = -0,47 \); ambos \( p < 0,0001 \)). Considerando o contágio de folículos antrais como a variável principal, observou-se uma correlação positiva significativa com o total de oócitos recrutados. Ao correlacionar o contágio de folículos antrais com os folículos recrutados maiores que 18 mm, observou-se que, com um corte de 12 folículos antrais, há um valor preditivo positivo de 99%, e um área sob a curva ROC de 0,76.

**Conclusão** Concluímos de nosso estudo que a idade e o contágio de folículos antrais são previsores eficazes de resposta ovariana em ciclos de reprodução assistida. O volume ovariano, assim como a dosagem do hormônio anti-Mülleriano, parecem ser marcais adequados da reserva ovariana.
2014. Foram avaliadas idade, exames bioquímicos e ecografia. Os dados foram analisados na predição do desenvolvimento follicular e nas suas relações entre si, utilizando para análise estatística o programa Statistical Package for Social Sciences.

**Results** Dos 293 casais incluídos, 50,2% apresentavam infertilidade por fator ovariano. Considerando a idade como principal variável, foi observada uma correlação significativa e negativa com volume de ambos ovários (ovário direito, r = 0,21; ovário esquerdo, r = −0,22; ambos p < 0,0001), e com contagem de foliculos antrais (ovário direito, r = −0,38; ovário esquerdo, r = −0,47; ambos p < 0,0001). Considerando a contagem de foliculos antrais como a variável principal, foi observada uma correlação significativa e positiva com o total de óocitos recrutados. Quando correlacionamos a contagem de foliculos antrais com os foliculos recrutados maiores do que 18 mm, observamos que, com um ponto de corte de 12 foliculos antrais, tem-se um valor preditivo positivo de 99%, e uma área da curva ROC de 0,76.

**Conclusions** Concluímos com nosso trabalho que a idade e a contagem de foliculos antrais são eficientes preditores da resposta ovariana em ciclos de reprodução assistida. O volume ovariano, assim como a dosagem do hormônio anti-mulleriano, parecem ser marcadores adequados de reserva ovariana.

**Introduction**

Research on infertility has evolved with constant studies and technological advances due to the increase of infertile couples who seek assisted reproduction (AR) services.\(^1,2\) The recruitment and development of multiple ovarian follicles are key to treatment.\(^1,3,4\)

The correct assessment of the ovarian reserve is a central issue in the management of patients with infertility.\(^1,3,5\) The goal is to predict the chances of response to the induction and select the “optimal” dose for the ovarian hyperstimulation.\(^3,6\)

Among the aspects of clinical history, advancing age is considered a determining factor of fertility.\(^7,8\) It causes a decrease in the ovarian reserve and an impaired oocyte quality.

The ovarian volume, thanks to its good accuracy and the cost-benefit ratio, seems to be a useful tool in monitoring patients undergoing AR.\(^9\)

Studies have demonstrated an apparent correlation between the number of antral follicles and the functional status of the ovaries.\(^3,6,10–13\) When compared, the antral follicle count (AFC) is higher than the volume of the ovaries in the evaluation of poor response to in vitro fertilization (IVF).\(^7,9\)

The basal serum follicle stimulating hormone (FSH) concentration, together with the concentrations of estradiol,\(^1,14\) represents a predictive value for ovarian response.\(^15,16\) Elevated FSH levels are strongly associated with poor ovarian response, low levels of estradiol and low rate gestation.\(^1,17\) Therefore, the basal serum concentration of FSH\(^6\) has been used as a marker for ovarian insufficiency.\(^9\)

The dosage of anti-Müllerian hormone (AMH) has been used as an early and sensitive marker of the ovarian reserve.\(^1,8,18\) It reflects the amount of remaining primordial follicles, so this hormone is strongly associated with the AFC.\(^1,8,11,19\) The decline in AMH levels can be detected earlier than other hormonal changes and the AFC.\(^1,20,21\)

A poor ovarian response may be associated with low pregnancy rates, and cycles are frequently cancelled without achieving the oocyte retrieval.\(^7,22\)

Despite the importance of the ovarian reserve measurement, the best way to correctly evaluate the follicular status remains controversial.\(^1\) A good ovarian reserve test should be predictive of conception, and should indicate the probable duration of ovarian activity. Furthermore, it should point out the ideal dose of ovarian stimulation and the chance to achieve gestation successfully.\(^23\)

**Purpose**

The main objective of this work is to identify which methods used in the assessment of the ovarian reserve are exclusive or complementary to identify the best response to follicle development.

**Methods**

After approval by the Ethics and Research Committee of the Faculdade de Medicina de São José do Rio Preto (FAMERP) under number 18617, this study was conducted at the Instituto de Medicina Reprodutiva e Fetal (Institute of Reproductive and Fetal Medicine – IMR), São José do Rio Preto, São Paulo, entity convened with FAMERP. It is a retrospective cohort study, involving patients under AR treatment for conjugal infertility.

From April 2009 to July 2014, 379 patients were submitted to ovulation induction for high complexity AR at the IMR and selected for this work. After signing a free and informed consent, and without prejudice to the proposed treatment, 293 patients were included in the study.
The included patients were evaluated, regardless of their chronological age, through biochemical tests (FSH, estradiol or AMH) and ultrasound (ovarian volume and AFC) for analysis of the ovarian reserve.

Patients with suspicion of pregnancy; patients who would have a procedure with egg reception; those submitted to oophorectomy and/or oophoroplasty, unilateral or bilateral; those who underwent ultrasound in other services; the ones in which one ovary was not visualized at baseline ultrasound; those with residual ovarian cyst (greater than 16 mm average diameter) in the moment of the first ultrasound; and the ones presenting FSH dosages greater than 15 UI/mL were excluded.

Clinical data were collected through interviews held by the author of the work or members of her team, including collection of medical record data.

Blood samples were drawn for FSH and estradiol dosage on the same day of the first ultrasound, always before the beginning of the stimulus, on the third day of the menstrual cycle at most. Anti-Müllerian hormone measurements were performed before the beginning of the treatment, at any stage of the menstrual cycle. The ultrasound operators had access to the laboratory tests results.

The sonographic examination was performed using the following ultrasound apparatus: Medison, model ACCUVIX XP (Samsung, Seoul Korea); and GE, model Voluson E8 (GE Healthcare, Austria), with an endovacitary transducer of 5–12 MHz, adequately prepared for the implementation of endovaginal examination, performed always by the same operators (the author of this paper and three other IMR clinical doctors), and lasting ten to twenty minutes. The ultrasound was performed between the first and third days of the menstrual cycle.

During the 2D ultrasound examination, the volume of the ovaries was calculated by the measurement of their three largest diameters. The antral follicles were counted, and they are represented at the ultrasound as small round anechoic images with diameters between two and ten millimeters. The register was obtained by measuring the largest diameter of each follicle. The follicle number count was performed by scanning the ovaries transvaginally.

The new serial ultrasound was performed between six and eight days after the start of the ovarian stimulation. From there, the ultrasound control was performed with individual frequency for each case. Controls were performed every 2 days until reaching follicles with a 16 mm average diameter, and daily until obtaining at least 3 follicles of 18 mm or more for the final maturation with the human chorionic gonadotropin hormone.

Later, the sonographic parameters were analyzed (ovarian volume, AFC and follicular development data), as well as the age and laboratory parameters (serum levels of FSH, estradiol and AMH), in the prediction of follicular development, and also the in the relationships among them. Data referring to egg collection, fertilization and pregnancy rates were not studied to avoid potential sources of bias.

The collected data were tabulated and statistically analyzed using the Statistical Package for Social Sciences (SPSS) software for Windows (SPSS, version 20, Chicago, US).

The association between variables was assessed using Spearman’s correlation coefficient. We used the receiver operating characteristic (ROC) curve to determine the cutoff point, the sensitivity, the specificity and the predictive values of the tests. The comparison between the subgroup of number of antral follicles and the ovarian volume was analyzed by Kruskal-Wallis test. A p value of < 0.005 was considered statistically significant.

This Project did not require any financial support to be developed.

Results

Of the 293 couples included in this study, 39.6% had infertility by male factor; 50.2% by ovarian factor; 13.7% by endometriosis; 12.3% by tubal factor; 2.7% by uterine factor; 7.8% without factors defined and classified as infertility or sterility without apparent cause; and 25.6% had more than 1 factor of infertility.

The average age was 34.7 years, ranging from 23 to 47 years; the FSH baseline had a mean of 4.3 IU/mL, ranging between 0.2 and 14.6 IU/mL; the estradiol, a mean of 18.4 ng/mL; and the AMH, a mean of 1.4 ng/mL ranging between 0.3 and 3.6 ng/mL.

The ovarian volume measurement and the AFC were calculated in 293 patients, yielding an average of 4.5 cc in volume and 7.7 in the average number of antral follicles for the right ovary and mean of, respectively, 4.3 cc and 7.3 for the left.

Considering the age as the main variable, a significant negative correlation was observed with the volume of both ovaries (right ovary, $r = -0.21$; left ovary, $r = -0.22$; both $p < 0.0001$), the AFC (right ovary, $r = -0.38$; left, $r = -0.47$; both $p < 0.0001$) and the total of recruited follicles ($r = -0.47$, $p < 0.0001$).

Considering the volume of both ovaries as the main variable, a significant positive correlation with the AFC of the respective ovary and the recruited follicles was observed, as shown in Table 1.

Considering the AFC as the main variable, a significant positive correlation was observed with the total of recruited oocytes (right ovary, $r = 0.73$; left ovary, $r = 0.72$; total antral follicles, $r = 0.77$; all $p < 0.0001$).

Considering the FSH dosage as the main variable, a significant positive correlation with the estradiol dosage ($r = 0.20$; $p = 0.004$) was observed. This means that in patients of reproductive age, the greater the dose of FSH, the higher the dosage of estradiol. Correlating the FSH with age ($p = 0.49$; $r = 0.45$), a positive correlation without statistical significance was observed. Correlating the FSH with the AFC (both ovaries, $r = 0.07$; $p = 0.24$), a positive correlation without statistical significance was obtained, probably because patients with FSH higher than 15 UI/mL were excluded from the study. In relation to the recruited follicles ($r = -0.27$;
A negative correlation was obtained, though not statistically significant. The AMH versus age showed a negative correlation (\( r = -0.33; p = 0.21 \)); correlating the AMH with the volume of the ovaries, positive values were obtained (right, \( r = 0.49; p = 0.06 \) and left, \( r = 0.28; p = 0.30 \)); and correlating the AMH with the AFC, positive values were obtained as well (right ovary, \( r = 0.48; p = 0.06 \); and left ovary, \( r = 0.37; p = 0.17 \)).

In making the correlation of the ovarian volume with the recruited follicles larger than 18 mm, when using a 2.6 cm\(^3\) cutoff point in the amount of at least one ovary, we observed: a sensitivity of 81%; a positive predictive value of 97%; and an area under the ROC curve of 0.76 (\( \text{Fig. 1} \)).

However, when we correlated the AFC with the recruited follicles larger than 18 mm, we observed that, with a cutoff of 12 antral follicles, a positive predictive value of 99% and an area under the ROC curve of 0.76 were obtained (\( \text{Fig. 2} \)).

Of the 293 people included in this study, 21 patients had their cycles cancelled due to ovulation induction failure. In this group, age (\( p = 0.03 \)) and FSH levels (\( p = 0.02 \)) higher than in the group that completed the stimulus were observed; ovarian volume and an AFC (\( p = 0.0001 \)) lower than in the group that completed the stimulus, were also observed.

### Table 1: Spearman Correlation between the ovarian volume and the AFC, and the number of recruited follicles

<table>
<thead>
<tr>
<th>Variable</th>
<th>( r )</th>
<th>AFCR</th>
<th>AFCL</th>
<th>Recruited follicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right ovary volume</td>
<td>&lt; 0.0001</td>
<td>0.593</td>
<td>0.489</td>
<td></td>
</tr>
<tr>
<td>Left ovary volume</td>
<td>&lt; 0.001</td>
<td>0.560</td>
<td>0.438</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: AFC, antral follicle count; AFCL, antral follicle count in the left ovary; AFCR, antral follicle count in the right ovary; \( p \), percentile.
Discussion

The application of assessment tests of ovarian reserves in patients who will undergo IVF has been widely studied, mainly due to the high cost and complexity of the treatment. But there is still no consensus regarding the best predictor of ovarian response, despite the fact that several studies have been published, a fact that may explain the abundance of proposed tests.

The objective of establishing the best method or combination of these tests is to reduce the number of tests that a particular patient should be submitted for the evaluation of the ovarian reserve, reducing the stress and financial burden, without compromising the necessary information to perform an IVF treatment, however.\textsuperscript{14}

Considering the age as the main variable, we observed that there is an obvious and strong correlation: the greater the age, the lower the volume of the ovaries, the number of antral follicles, and the number of recruited follicles. The inverse association of a woman’s age with low a ovarian reserve is largely reported in other published studies.\textsuperscript{9,12,24} Choi et al\textsuperscript{25} report that before the first IVF in a patient, 60% of the predicted prognostic is for the age, and 40% for other clinical factors.

Aboulghar et al\textsuperscript{8} describe that age remains as the primary determinant of success in AR cycles; however, at any age, women with high levels of AMH have a higher success rate of oocyte recovery after follicular puncture than those with lower levels of AMH in the same age range.

Surekha et al\textsuperscript{26} confirm the data from our study when comparing the lack of a significant association of the FSH with the ovarian reserve, unlike other specific markers, such as AMH and AFC.

In another retrospective study, Chuang et al\textsuperscript{27} concluded that the basal FSH is a good marker for the remaining follicular pool. Magalhães et al\textsuperscript{28} described that it is difficult to find associations between the FSH basal level and the follicular count when patients with FSH levels greater than 15 UI/mL are excluded from the study, as it was done in our work. On the other hand, Abdalla and Thum\textsuperscript{29} suggest that a high rate of FSH (\textgreater{} 10 IU/mL) should not be a criterion for the exclusion of patients for treatment with IVF, since the test represents a quantitative and not qualitative aspect of the ovarian reserve, that is, despite having a low follicular pool, one patient does not necessarily present a poor oocyte quality, especially the ones younger than 38 years of age.\textsuperscript{9}

Luna et al\textsuperscript{17} describe that the cycle cancellation rates were significantly higher in patients with elevated...
FSH ($\geq 13.03$ IU/mL) on day 3, compared with patients with normal levels of FSH in all age groups, which was similar to the rates found in our study.

Regarding the ovarian volume, we found an association of this variable with the ovarian response. Some studies have shown that the reduction is related to poor response to ovarian stimulation, but the sensitivity and specificity are lower compared with the AFC.

Considering the AFC as a main variable, we found in our work that a higher AFC relates to a greater number of recruited follicles, similar to the findings of Souza et al. Magalhães et al. describe that the predictive power of the AFC is substantially similar to the serum concentration of AMH, but with a higher sensitivity. The data from our work are also similar to Barbakadze et al., and show negative correlation between age and the AFC.

Iliodromiti et al. describe a progressive decline of AMH concentration with advancing age. Other studies have reported a strong correlation between the number of antral follicles and the serum basal level of AMH. This relationship can be explained by the fact that the AMH is produced by the theca cells of the antral follicles; therefore, the greater the number of antral follicles, the greater the amount of hormone produced, increasing its concentration in the blood.

Similarly to our study, Aflatoonian et al. when correlating the AFC with the recruited follicles greater than 18 mm, observed good sensibility and specificity (89 and 92% respectively), but used a cutoff point of 16 antral follicles, while in our study we used 12.

Most of the bibliographic data obtained confirm that the AFC and the AMH can be used as a screening method to detect probable poor responders, or responders with lower reserves, and to predict the risk of ovarian hyperstimulation syndrome (OHSS) and the cancellation of cycles, having the best predictive value of the number of oocytes collected in IVF cycles.

Broer et al. found a sensitivity of 82% and a specificity of 80% for a prediction of poor response. Kwee et al. found low response with an AFC lower than 6, the same cutoff point we used, with a sensitivity of 73% and a specificity of 95%.

To predict OHSS, Aflatoonian et al. showed that the AFC and the AMH have similar accuracy, with an area under the ROC curve of 0.961 and 0.922 respectively. Martins et al. used 20 follicles as the cutoff point, and showed that these women have an increased risk of OHSS. In a systematic review, Nastri et al. presented that the evaluation by AFC and AMH dosage, before the ovulation induction cycle, allows the prediction of the risk of OHSS, and that other baseline parameters, such as age and FSH levels, are less accurate for prediction.

Although the combination of existing tests appears to be the most effective and useful conduct for counseling infertile patients, in an attempt to provide them with some expectations about the success of the proposed treatment, our work shows stability in predicting ovarian reserve and subsequent oocyte recovery after follicular puncture through a simple ultrasound routine exam used in preconception gynecological evaluation.

Therefore, it is important to investigate the ovarian reserve through the AFC, since it allows the optimization of treatment protocols, and a reduction of the traumatic occurrence of cancelled cycles and of the fearful side effects of controlled stimulation for poli-folliculogenesis, such as OHSS.

Our AMH correlations were not statistically significant, probably due to the small number of patients who collected the serum sample for this type of marker.

Cohort studies may be subject to selection bias, as in our study, in which infertility patients submitted to AR were selected. Only women with FSH dosages lower than 15 UI/mL were included; therefore, bad potential responders were excluded and this selection may have attenuated the overall strength of the correlations.

One aspect that still requires further research and further studies is the fact that all existing tests, at the moment, are quantitative predictors of the follicular pool, but are not able to evaluate the oocyte quality or the pregnancy rates.

Conclusions

We concluded that age and the AFC are effective predictors of ovarian response in AR cycles. The ovarian volume and the AMH dosage also appear to be suitable markers of ovarian reserve.

Future Perspectives

The AFC, when performed by the three-dimensional method, appears to offer advantages over the two-dimensional method, among which we mention: the ability to produce images in three different planes; the possibility of virtual coloration of the follicles, eliminating the need to repeat the process or the failure to identify them, thus eliminating the technical issue that the AFC is operator-dependent, and greater speed of the test execution and greater comfort for the patients, since the obtained images can be analyzed later. The 3D ultrasound with sonography-based automated volume count (SonoAVC, GE Healthcare, Austria) is used to obtain more precision and better reproduction; however, early studies show no statistical difference between the two and three-dimensional evaluations.

In our service, this technique is already being used, and, in the future, we can compare our data (two-dimensional evaluation) to study the correlation of the three-dimensional evaluation with the serum levels of AMH, the randomization of ovarian hyperstimulation protocols, with consequent follicular response, and the dopplerfluorometric analysis of oocyte quality, results that could help to achieve better pregnancy outcomes.

References
