COMPARATIVE STUDY OF OVARIAN RESERVE TESTS AND OVARIAN RESPONSE TO
CONTROLLED OVARIAN STIMULATION IN IVF
Karunakaran Sandeep1, Srinivas Sangisapu2

1Professor, Department of Obstetrics and Gynaecology, Institute of Naval Medicine at INHS, Asvini, Mumbai, Maharashtra.
2Associate Professor, Department of Obstetrics and Gynaecology, Institute of Naval Medicine at INHS, Asvini, Mumbai, Maharashtra.

ABSTRACT

BACKGROUND
Controlled ovarian stimulation (COS) and optimum retrieval of fertilizable oocytes is crucial for IVF success. Ovarian response to COS is related to ovarian reserve (OR) - size of primordial follicular pool capable of maturing in the presence of gonadotrophins. There is a temporal fall in follicular pool from embryonic life onwards. The quest for ideal ovarian reserve test (ORT) resulted in several markers to measure ovarian response singly or in combination increasing the cost. The present study is a comparative evaluation of these tests with a view to ascertain the single most suitable test for ovarian response.

The aim of the study is to compare and evaluate various ORTs with ovarian response and oocyte retrieval.

The Objective of the study is to measure the hormonal and sonographic ovarian reserve markers and to evaluate which measure amongst them is superior in assessing the ovarian response and oocyte retrieval.

MATERIALS AND METHODS
Single centre retrospective cohort study of 183 women planned for first fresh cycle IVF with COS by GnRH agonist long protocol. Antral follicle count (AFC) and basal hormonal assay for FSH, LH, Oestradiol (E2), Inhibin-B and AMH were compared and evaluated for ovarian response and oocyte retrieval.

RESULTS
Serum AMH levels and Total AFC are significantly (p<0.001) lower in poor responders which is further confirmed by multivariate regression analysis (<0.05). It also shows that both AMH and AFC are significant (p<0.001) predictors of number of oocytes retrieved.

CONCLUSION
Serum AMH and Antral Follicular Count (AFC) are equally significant predictors of ovarian response and number of oocytes retrieved.

KEYWORDS
Ovarian Reserve Test (ORT), AFC, AMH, Predictors of Ovarian Response, Oocyte Retrieval.

HOW TO CITE THIS ARTICLE: Sandeep K, Sangisapu S. Comparative study of ovarian reserve tests and ovarian response to controlled ovarian stimulation in IVF. J. Evid. Based Med. Healthc. 2019; 6(1), 1-8. DOI: 10.18410/jebmh/2019/1
disadvantages. These markers and their predictive role is reviewed at length by Jirge PR. An ideal ORM should be simple, accurate, easy to perform, reliable, reproducible with least inter observer, inter & intra assay variations, cycle independent and cost effective with good sensitivity, specificity and predictive capability not only in terms of ovarian response to COS but also IVF outcome across all categories of infertile women. However there appeared to be no such ideal marker that is technically superior satisfying all the criteria. Since no single test was singled out as the marker of OR, investigators have tried combination of tests with a view to improving the predictive potential of ORM. The current study is designed for comparative evaluation of these markers/tests using both basal ultrasonographic (Antral Follicular Count) & endocrine parameters (basal FSH, basal LH, basal oestradiol- E2, inhibin B and AMH) of ovarian reserve corroborate with ovarian response and oocyte yield in women undergoing COS. The terms ovarian reserve marker (ORM) and ovarian reserve tests (ORT) are used interchangeably in this article.

Aims and Objectives
1. To compare and evaluate various ORT with ovarian response and oocyte retrieval.
2. To measure the hormonal and sonographic ovarian reserve markers and
3. To evaluate as to which measure amongst them is superior in assessing the ovarian response and oocyte retrieval.

MATERIALS AND METHODS
A single centre retrospective cohort study was conducted on 200 women undergoing first fresh IVF & Embryo Transfer (ET) and COS by long protocol with GnRH agonists followed by ovulation trigger with human chorionic gonadotropin (hCG) for a period of one year. Inclusion and exclusion criteria are listed in Box-1.

Inclusion Criteria
First fresh IVF cycles, COS by long protocol with GnRH agonists followed by hCG trigger. Unexplained infertility due to absent or damaged fallopian tubes, infertility due to uterine factors, endometriosis, Absence of medical illness in either or both partners. Prepared fresh semen samples of male partners, with recovery concentration of sperms >10 million/mL.

Exclusion Criteria
Frozen embryo transfers (FET) subsequent attempts of COS, PCOS, OHSS, Non-responders to COS, Cryopreserved samples. Couple who underwent ICSI, Ovum donation Presence of medical illness in either or both partners. Prepared fresh semen samples of male partners, with recovery concentration of sperms <10 million/mL.

As a standard protocol thorough evaluation of infertile couples is done prior to recruitment for IVF- detailed history, physical examination, complete blood counts, urinalysis, hormone profile, blood sugar, blood group, coagulation profile, HIV, HBV, HCV, VDRL, Hormone profile, Mantoux test, ESR, chest X-ray, Hysterosalpingogram, ultrasonography of pelvis, ovulation studies, Diagnostic Hysterosalpingoscopy, and Seminal analysis (at least twice) as per WHO Manual 2010. All female partners received pre conceptional folic acid supplementation 5 mg twice a day and tab Aspirin 75 once daily. Women included in the study underwent 2D trans-vaginal ultra-sonography (TVS) on Day 2 of the cycle to rule out the presence of ovarian cyst and for AFC ((sum total of antral follicles measuring 2 to 6 mm of right and left ovaries). Blood samples collected on same day to assess basal levels of FSH, LH, oestradiol (E2), Inhibin-B and AMH. GnRH down regulation started on Day 1. COS was achieved by recombinant FSH. Depending upon ovarian response by TVS, gonadotrophin dose was titrated afterwards at least 5 days of stimulation. Women with at least >2 follicles of 15 mm in diameters were considered to have an adequate ovarian response. Those with ≤ 2 follicles <12 mm after day 12 of stimulation were considered as inadequate ovarian response. When more than 2 leading follicles attain 18 mm size, recombinant hCG (250 mg) and urinary hCG (5000 IU) was administered for final maturation of oocytes and ovum pick up (OPU) was done as per standardised protocol after 36 hours of hCG trigger.

Hormonal Assays
Blood samples were collected from each volunteer into two plain vacutainers. Samples were centrifuged immediately, and serum separated. One vacutainer was used for the FSH, LH, and E2 assays, which were performed within 2–3 hours after obtaining or within 24 hours in which case the serum was stored at 2°C until assayed. The inhibit and AMH were assayed batch wise. Hence serum samples collected for the same were frozen at -20°C and stored until sufficient samples were available. The FSH, LH, and E2 levels were measured by micro particle enzyme immunoassay method. The respective analytical sensitivity, the intra-assay and inter assay coefficients of variation are 0.37 IU/L, <5%, & <15% (FSH), 0.5 IU/L, <7%, & <8% (LH) and 8 p mol/L, 2.9%–11%, &4.8%–15.2% (E2).

Serum AMH levels and inhibit-B levels were assessed by their respective Gen II enzyme linked immunosorbent assay (ELISA) kit (Beckman Coulter). The respective analytical sensitivity and the intra-assay and inter assay coefficients of variation for AMH and inhibit-B are 0.08 ng/mL, <5%, and <8% (AMH), and 2.6 pg/mL, <6%, and <8% (AMH).

All semen samples were prepared by double density gradient swim up method.

Fertilisation was performed by micro droplet method with oil overlay within 2 hours of removing sperm from final swim up and embryos were cultured in sequential culture media (Vitrolife).

Assessment of fertilisation was done on Day 1 i.e.; 18 to 20 hours post insemination at two Pro Nuclei (PN) stage. Denudation of was carried out on Day 1 at the time of assessment of fertilization.
Ethical committee clearance was taken for conducting the study. All tests were carried out with written and informed consent of patients. Staff at ART Centre are blinded to measurements of ORM except AFC. Staff at endocrine laboratory measuring the endocrine markers were blinded to the outcome. Serum hormone levels and outcome were analysed retrospectively by third group of investigators.

**Statistical Analysis**

SPSS software version 20. ROC analysis, univariate and multivariate logistic regression analysis were done. P value of < 0.05 was taken as significant.

### Inclusion and Exclusion Criteria

<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) First fresh IVF cycles.</td>
<td>a) PCOS</td>
</tr>
<tr>
<td>b) COS by Long protocol with GnRH agonists followed by hCG trigger.</td>
<td>b) OHSS</td>
</tr>
<tr>
<td>c) Unexplained infertility</td>
<td>c) Non-responders to COS</td>
</tr>
<tr>
<td>d) Infertility due to absent or damaged fallopian tubes.</td>
<td>d) Couple who underwent ICSI, ovum donation</td>
</tr>
<tr>
<td>e) Infertility due to uterine factors.</td>
<td>e) Presence of medical illness in either or both partners</td>
</tr>
<tr>
<td>f) Endometriosis.</td>
<td>f) Prepared fresh semen samples of male partners, with recovery concentration of sperms &lt;10 million/ml</td>
</tr>
<tr>
<td>g) Prepared fresh semen samples of male partners, with recovery concentration of sperms &gt;10 million/ml</td>
<td></td>
</tr>
</tbody>
</table>

### RESULTS

183 women were analysed out of 200 (Flow chart) with independent variables i.e., age, BMI, Day 2 FSH, E2, Inhibin-B, AMH. AFC vis-a-vis the primary outcome variables (normal responder / poor responders and number of oocytes retrieved). The mean age, standard deviation (SD) and age range of the participants was 33.5 ±3.5 years (24–40 years). Their mean BMI, SD, BMI range was 24.3 ±3.4 kg/m² (20–35 kg/m²) (Table 1).

### Baseline Characteristics

Table 2 shows cross tabulation of base line values of independents variables (age, BMI, basal FSH, LH, E2, inhibin-B, and AFC) with primary outcome variable (normal responders and poor responders). 20 participants (10.9%) were poor responders. Both Age and BMI was similar in both groups. Poor responders had significantly lower AMH levels (P<.001) and significantly higher FSH levels (P<.01) than the normal responders. The poor responders also had a lower total AFC (P<.001). The other endocrine markers (basal LH, E2, and Inhibin-B) were similar in both groups. The number of oocytes retrieved was significantly lower (P<.001) in the poor responders compared with the rest of the group (1.8 ±1.4 vs. 11.7±5.2). A preliminary correlation analysis for “number of oocytes retrieved” has shown that Total AFC (r=0.8833), AMH (r=0.8614), Inhibin-B(r=0.8050) and Oestradiol (E2) (r=0.6267) are the top most best predictors in that order. Hence, Day 2 FSH and LH were excluded from the study. ROC analysis was done on the remaining four parameters. (graph 1 ROC analysis of AFC, graph 2 ROC analysis of AMH, graph 3 ROC analysis of inhibin, graph 4 ROC analysis of oestradiol (E2, graph 5 Comparative ROC analysis of all ORM). Based on ROC analysis an optimum cut-off was determined for predicting “adequate response” determined as per criteria >2 oocytes retrieved.

The results are tabulated in Table 3. Total AFC is the best predictor for the number of oocytes retrieved. This is better than AMH as a standalone test. In predicting the outcome of COS. Univariate and multivariate linear regression analysis for the prediction of the number of oocytes retrieved are shown in Table 4. Several of the parameters were predictive on univariate analysis, But, AFC and AMH were the only significant predictors on multivariate analysis. Multivariate regression analysis of basal markers of ovarian reserve for the prediction of poor response are shown in Table 5. The results were similar for prediction of poor ovarian response. i.e., AFC and AMH were the only significant predictors on multivariate logistic regression analysis. Fertilisation success as assessed by formation of PN stage of embryo was the secondary outcome studied. Only 6 out 183 had fertilisation failure. We have not included further follow up of these cases for secondary outcome since multiple confounders, modifiers and moderators other than due to ovarian reserve might actually come in to play resulting in erroneous observations. The unsuccessful fertilisation number (6) is so small that any further analysis of non-pregnancy / non-conception would give skewed results. Similarly, there is only one case of OHSS post OPU. Hence, we did not analyse ORT for ovarian hyper response.
Table 1. Base Line Characteristics - Age and BMI

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Normal Responders (n = 163)</th>
<th>Poor Responder (n = 20)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Years)</td>
<td>33.3±3.6 (24-40)</td>
<td>35.7±1.9 (33-39)</td>
<td>&lt;0.5</td>
</tr>
<tr>
<td>Body Mass Index (kg/m²)</td>
<td>24.4±3.4 (20-35)</td>
<td>24.0±2.9 (20-30)</td>
<td>.60</td>
</tr>
<tr>
<td>Basal FSH Level (IU/L)</td>
<td>7±1.8 (2.95-11.96)</td>
<td>8.3±1.5 (4.8-10)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Basal LH Level (IU/L)</td>
<td>5.5±3.0 (1.3-28.3)</td>
<td>5.3±1.9 (2.0-9.0)</td>
<td>.71</td>
</tr>
<tr>
<td>FSH: LH Ratio</td>
<td>1.5±0.8 (0.2-5.5)</td>
<td>1.8±0.8 (0.9-4.2)</td>
<td>.22</td>
</tr>
<tr>
<td>Basal E2 Level (p mol/L)</td>
<td>160.7±56.0 (42-373)</td>
<td>182.2±77.6 (76-357)</td>
<td>.31</td>
</tr>
<tr>
<td>Inhibin-B (pg/mL)</td>
<td>51.6±28.7 (7-164)</td>
<td>58.7±62.6 (7-264.9)</td>
<td>.45</td>
</tr>
<tr>
<td>Anti-Mullerian Hormone (ng/mL)</td>
<td>1.48±0.75 (0.19-4.31)</td>
<td>0.58±0.28 (0.12-0.99)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Total Antral Follicle Count (AFC)</td>
<td>15.7±4.3 (5-22)</td>
<td>8.6±1.9 (5-12)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Number of Oocytes retrieved</td>
<td>11.7±5.2</td>
<td>1.8±1.4</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 2. Comparison of Baseline Clinical, Endocrine and Ultrasound Characteristics between Normal and Poor Responder Groups

<table>
<thead>
<tr>
<th>Serial No.</th>
<th>Independent Variable</th>
<th>Cut Off Value</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Response Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Total AFC</td>
<td>≥4</td>
<td>96.18%</td>
<td>96.15%</td>
<td>“adequate response” if &gt;2 number of oocytes retrieved</td>
</tr>
<tr>
<td>2.</td>
<td>AMH</td>
<td>≥1.6 ng/ml</td>
<td>92.99%</td>
<td>100%</td>
<td>“adequate response” if &gt;2 number of oocytes retrieved</td>
</tr>
<tr>
<td>3.</td>
<td>Inhibin</td>
<td>≥57 pg/ml</td>
<td>96.82%</td>
<td>57.69%</td>
<td>“adequate response” if &gt;2 number of oocytes retrieved</td>
</tr>
<tr>
<td>4.</td>
<td>Oestradiol (E 2)</td>
<td>≥50.3 P mol/ml</td>
<td>96.82%</td>
<td>34.62%</td>
<td>“adequate response” if &gt;2 number of oocytes retrieved</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Regression Coefficient (Mean &amp; 95% CI)</th>
<th>p-Value</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.147 (-0.340, 0.047)</td>
<td>0.132</td>
<td></td>
</tr>
<tr>
<td>Basal FSH</td>
<td>-0.061 (-0.441, 0.324)</td>
<td>0.752</td>
<td></td>
</tr>
<tr>
<td>Basal E2</td>
<td>-0.007 (-0.018, 0.002)</td>
<td>0.155</td>
<td></td>
</tr>
<tr>
<td>Anti-Mullerian Hormone</td>
<td>2.352 (1.040, 3.664)</td>
<td>&lt;0.001</td>
<td>0.470</td>
</tr>
<tr>
<td>Antral Follicle Count</td>
<td>0.480 (0.230, 0.732)</td>
<td>&lt;0.002</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Univariate and Multivariate Analysis of Baseline Markers of Ovarian Reserve for Prediction of the Number of Oocytes Retrieved at Ovum Pickup (OPU)
<table>
<thead>
<tr>
<th>Parameters</th>
<th>Odds Ratio</th>
<th>95% CI</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.172</td>
<td>0.842-1.634</td>
<td>0.35</td>
</tr>
<tr>
<td>Basal FSH</td>
<td>1.156</td>
<td>0.794-1.686</td>
<td>0.45</td>
</tr>
<tr>
<td>Anti-Mullerian Hormone</td>
<td>0.130</td>
<td>0.018-0.935</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Antral Follicle Count</td>
<td>0.650</td>
<td>0.446-0.948</td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>

Table 5. Multivariate Regression Analysis of Basal Markers of Ovarian Reserve for the Prediction of Poor Ovarian Response to Controlled Ovarian Stimulation (COS)

DISCUSSION

Our study suggests that AFC, measured by TVS (2D ultrasound), and AMH are the most significant predictors of the number of oocytes retrieved and poor ovarian response. When compared with each other, AFC and AMH appear equally predictive of poor response. AMH may be relatively more likely valid for poor response as it is the most significant predictor of the number of oocytes retrieved (table 4 p<0.001 for AMH and p< 0.002 for AFC). AMH has the potential to replace basal FSH which has been widely used for as an ORT for over two decades. Meta-analysis and systematic reviews failed to demonstrate any combination of specificity and sensitivity for basal FSH either for poor ovarian response or prediction of non-pregnancy. Though slightly different cut-off values are used, our findings are in agreement with previous studies as well as literature review examining the ability ORT to predict both...
the number of oocytes retrieved and the chance of poor ovarian response.\textsuperscript{25,26,29-32} The ability to discriminate poor responders from normal responders as a combined test, as assessed by ROC curve analysis, was not significantly different from that as individual tests (AUC for AFC and AMH). Hence there is no additional advantage of combination of tests.\textsuperscript{25} We did not analyse non-fertilisation and non-conception for reasons mentioned in results section of the article. Antral follicular count is done on day 2/3 bay taking the mean of two perpendicular diameters and adding the total count of both ovaries by conventional 2D TVS.\textsuperscript{28} Ovarian aging characterized by progressive temporal reduction of primordial follicular cohort measuring 2-6 mm,\textsuperscript{25} influences the response to stimulation.\textsuperscript{33} The AFC is positively correlated with the primordial follicular population\textsuperscript{34} and is a significant predictor of poor ovarian response with limited inter cycle variability.\textsuperscript{25} These gonadotrophin responsive follicles can be selected for further maturation to the preovulatory stage. Therefore, AFC is a direct marker of the follicular cohort that can be recruited for maturation but lacks the sensitivity and specificity to predict the non-occurrence of pregnancy.\textsuperscript{28} AFC provides an optimum sensitivity and specificity of 96.18 and 96.15 in our study. There is no advantage of 3D ultrasound over 2D for measurement of ovarian reserve.\textsuperscript{37}

AMH production by granulosa cells of preantral (primary and secondary) and small antral follicles (2-6 mm) begins with primordial follicular transition to the primary follicles and continues till the antral stages.\textsuperscript{30,34,38} AMH production declines with age in tandem with the declining primordial follicular pool of the small follicles till undetectable at and after menopause.\textsuperscript{39} There is a strong correlation between levels of AMH and day 2 antral follicle count (AFC).\textsuperscript{40} AMH measurement is cycle independent\textsuperscript{41,42} without any inter cycle variability.\textsuperscript{43} Threshold values varying from 0.2 to 1.26 ng/ml were used to predict poor responders with 80–87% sensitivity and 64–93% specificity.\textsuperscript{29,44,45} wide range of serum AMH levels are described and a reliable cut-off level is yet to be defined. There is no international assay standard for AMH measurement. This may probably explain the discordance between different studies and makes comparison between laboratories difficult.\textsuperscript{26} AMH can predict a hyper-response also.\textsuperscript{46} Age-related decline of AMH levels thereby ovarian reserve can be identified by nomograms and abnormal deviations can be used for counselling couples planning to delay childbearing.\textsuperscript{47} AMH can be used as a marker to predict pregnancy.\textsuperscript{48,49} AMH can be screening test in a general sub fertile population as well.\textsuperscript{50}

Since AFC as an integral part of the IVF protocol can be performed before or after down-regulation and prior to ovarian stimulation with an equal predictive accuracy of AMH for poor OR,\textsuperscript{31} it is logical that AFC should be the first choice of tests until a assay techniques cut-off level of AMH are standardized. The disadvantage of AFC assessment is operator dependent.\textsuperscript{50,52}

In the present study neither Inhibin-B nor any of the other conventional markers proved to be predictive of ovarian response. Inhibin-B and E\textsubscript{2} are produced by the granulosa cells of early antral follicles and therefore, reflect the size of the growing follicular cohort.\textsuperscript{53-55} However, the levels of Inhibin-B and E\textsubscript{2} are regulated through pituitary FSH secretion\textsuperscript{56} and the negative feedback loops within the hypothalamic–pituitary–ovarian axis, which means that the levels of these markers are inter-related and dependent on each other and not simply the number and the size of the growing follicles levels.\textsuperscript{23,57}

The accuracy of prediction by inhibin-B of ovarian response and non-pregnancy at very low threshold levels, is very modest.\textsuperscript{58} Therefore, its routine use is not to be recommended. Basal E\textsubscript{2} according to meta-analysis does not add to the predictive value of other commonly used ORTs and hence its routine use in clinical practice is not recommended.\textsuperscript{26} Combination of ovarian markers with a view to improving the predictive value of single basal marker have not been found to be beneficial.\textsuperscript{27,29,60} Ovarian reserve tests have limited value in the prediction of non-conception and their routine use in clinical practice has been questioned.\textsuperscript{14} Pregnancy can occur even at extreme cut-off levels for an abnormal test result. Hence IVF treatment cannot be denied based on these tests, especially in participants who are seeking first cycle of IVF treatment.\textsuperscript{33} However, identification of participants who may poorly respond during IVF treatment is clinically pertinent as the couple require to be counselled and informed that there is an increased chance of cycle cancellation and a lesser chance of success, thereby allowing them to make an informed choice for IVF. ORTs allow clinicians to formulate individualised treatment protocols to optimise ovarian response.

CONCLUSION

AFC and AMH are the most significant predictors of the number of oocytes retrieved and of poor ovarian response to stimulation in IVF. Combination of these tests do not significantly improve prediction. AMH has a distinct advantage being cycle independent without any inter cycle variability but lacks assay and universally accepted cut-off value standardization. AFC on the other hand is operator dependent and hence inter observation variation is likely.

Recommendations

Since AFC is integral to IVF protocols AMH estimation is recommended as a single additional endocrine parameter as an ovarian reserve test.

Limitations of the Study

Its design as a single centre retrospective study with relatively small sample size with unknown prevalence of diminished ovarian reserve.

REFERENCES


