A review of luteinizing hormone and its role in ovarian reserve testing

Vinayak Smith¹*, Tiki Osianlis², Beverley Vollenhoven²³⁴

¹Eastern Health, Department of Obstetrics and Gynaecology, Box Hill Hospital, Nelson Road, Box Hill, Victoria 3128, Australia
²Monash IVF, 252 Clayton Road, Victoria 3168, Australia
³Monash Health, Women’s and Children’s Program, Monash Medical Centre, Clayton Road, Victoria 3168, Australia
⁴Department of Obstetrics and Gynaecology, Monash University, Clayton, Victoria 3168, Australia

Received: 3 December 2013
Accepted: 15 December 2013

*Correspondence: Dr. Vinayak Smith, E-mail: smith.vinayak@gmail.com

© 2014 Smith V et al. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Background: Ovarian physiology illustrates the synergistic interaction between luteinizing hormone and follicle stimulating hormone in the process of folliculogenesis. While follicle stimulating hormone has been well established as a marker of ovarian reserve, the role of luteinizing hormone has remained somewhat controversial and it seems to have become the ‘forgotten gonadotropin’. The following review aims to investigate the available evidence surrounding luteinizing hormone as an ovarian reserve test and examine the issues which need to be addressed in order to establish it as an ovarian reserve test. It then further attempts to propose a model to direct effective research to ascertain if it does have a role to play in ovarian reserve testing.

Findings: The evidence is equivocal in the use of luteinizing hormone as an independent predictor of ovarian reserve. However, there is much stronger evidence to suggest that the follicle stimulating hormone/luteinizing hormone ratio is a useful marker of ovarian reserve—in particular when it is ≥2 and approaching 3. The evidence base for this ratio at present however is sparse. In addition, ovarian reserve tests are fraught with issues over reliability, accuracy, definition and the implications of testing itself. In order to overcome these issues, more quality research needs to be carried out to test this relationship between luteinizing hormone and follicle stimulating hormone.

Keywords: Luteinizing hormone, Follicle stimulating hormone, Ovarian reserve, Gonadotropins, Assisted reproductive technology, Ovarian physiology

PURPOSE OF REVIEW

Ovarian reserve tests (ORTs) have become part of the antecedent workup for individuals undergoing assisted reproductive technology (ART). In 1988, the foundations for ORTs were first laid by studies indicating a potential role for basal follicle stimulating hormone (FSH) in predicting pregnancy.¹ Since then basal FSH measurement has gone on to become the most commonly used marker of ovarian reserve.² Given the synergistic interplay of both FSH and luteinizing hormone (LH) in ovarian physiology which ARTs attempt to mimic, it is of interest that LH has not been explored in the same capacity. This review aims to illustrate the physiological role played by LH in folliculogenesis and critically review the available literature to determine if LH does indeed have a role in ovarian reserve testing.

OVARIAN RESERVE TESTING

The term ovarian reserve aims to correlate reproductive potential with the number and quality of remaining oocytes in women of reproductive age.³ Ovarian reserve testing aims to quantify this relationship by measuring either oocyte quality, quantity or the ability for an individual to achieve pregnancy. These tests can be conducted either through biochemical means or through ultrasonographic measures.
Ovarian reserve tests (ORTs) have gained popularity in ART for their use as screening tests for diminished ovarian reserve (DOR) so as to help identify individuals who would be unlikely to achieve pregnancy using ART. These women were expected to have a poor response to ovarian stimulation or reduced fertility (ability to achieve pregnancy) in comparison to their equivalent aged peers. The tests were expected to help exclude these couples from ART thereby reducing healthcare costs, futile medical treatment, risks of surgical procedures and negative psychological impacts. However, to be able to successfully achieve this, an ORT needs to have a high specificity and predictive value as well as reproducibility.

Basal FSH concentration, measured between days 2-4 of the menstrual cycle, is well recognized and a commonly used marker of ovarian reserve. On the basis of ovarian physiology, it would seem reasonable to suggest that LH should have a role in determining ovarian fecundity as well as reserve. LH however, is yet to be established as a marker of ovarian reserve possibly due to the fact that only a few studies have evaluated its predictive role either independently or in conjunction with FSH.

**LUTEINIZING HORMONE (LH) IN OVARIAN PHYSIOLOGY**

Our current understanding of ovarian physiology is based upon the complex and synergistic interactions between both LH and FSH across the menstrual cycle.

The premise that both gonadotrophins concurrently exert their effects to stimulate follicular development, maturation and selection was first raised in the ‘two cells, two gonadotrophins hypothesis’. The hypothesis leaned towards a model of co-dependence between LH and FSH rather than on the autonomous action of either gonadotrophin alone during follicular maturation and ovarian steroidogenesis. In addition, it also suggested the involvement of the theca and granulosa cell layers in this process.

The clinical basis of this can be best demonstrated in individuals with Kallmann’s syndrome who have undergone ovulation induction. Due to hypogonadotrophic hypogonadism, they are unable to endogenously secrete gonadotrophins but have normal ovarian physiology. Ovulation induction, undertaken with both LH and FSH, results in normal follicle growth and a high luteal phase progesterone concentration following an artificial LH surge using human chorionic gonadotrophin (HCG). In contrast, treatment with FSH alone allows for similar ultrasonographic follicle growth but results in low concentrations of estradiol in the follicular phase and no increase in progesterone after HCG injection.

In the early follicular phase, FSH plays a pivotal role in the recruitment of the preantral follicle cohort to create a pool of synchronously developing follicles. Interestingly at this stage, the theca cells in these follicles only express LH receptors whilst the granulosa cells progressively acquire FSH receptors. This factor outlines the importance of LH, even in the early follicular phase, as it is required for the production of precursor androgens in the theca cell layer which are subsequently aromatized to oestriadiol in the adjacent granulosa cell layer under the influence of FSH.

By the mid follicular phase, LH receptors start becoming expressed in the granulosa cell layer as well. This however occurs in tandem with a rise in serum oestriadiol which exerts a negative feedback on the pituitary and thereby causes a fall in FSH concentration. This switch from relative FSH to LH dependence is what drives the continued growth of the dominant follicle, which displays the most efficient androgen aromatizing activity and also brings about atresia of the remaining follicle cohort which fail to appropriately adapt.

The late follicular phase is characterized by relative LH dependence of the dominant follicle. There is a rapid synthesis of oestriadiol subsequently from the dominant follicle. The rising oestriadiol then plays an important role in promoting follicle and oocyte maturity, readying the endometrium for implantation and priming the hypothalamic-pituitary-ovarian axis for the subsequent LH surge from which subsequent ovulation results.

In its entirety, the physiology of the ovaries is suggestive of LH having a much larger role in follicle maturation and selection than it is currently given recognition for. This can firstly be inferred from examination of the gonadotrophins across the menstrual cycle. FSH levels are generally elevated in between the luteal and follicular transition period, where they play a role in follicular recruitment. However, they continuously decline from that point until midcycle where the levels are triggered again by increasing oestriadiol and progesterone concentrations. LH in contrast progressively increases across the follicular phase primarily due to LH pulse frequency augmentation. On this premise, if follicle maturation were just solely dependent on FSH levels, we should expect developmental arrest in mid to late follicular phase in line with a decline in its levels which evidently is not the case. Secondly, another suggestive feature is the change in LH receptor status of the granulosa cell across the follicular phase especially taking into account its impact on dominant follicle selection. The dominant follicle expresses LH receptors more efficiently than the smaller ovarian follicles which allows it to keep growing compared to the other smaller follicles. This creates a scenario where FSH functions as a ‘nurturer’ - promoting initial follicle recruitment and maturation whilst LH displays ‘hormonal Darwinism’ in assessing the follicles ability to achieve ovulation through successful receptor expression.

The suggestive nature of the above mentioned physiological mechanisms does prompt the clinical
question of whether LH does indeed have a bigger role to play in assessing ovarian function?

THE EVIDENCE FOR LH AS A MARKER OF OVARIAN RESERVE

In 1998, Noci et al. produced the earliest evidence that basal day 3 LH concentrations <3mIU/ml were predictive of a poor ovarian response.16 These findings of reduced ovarian response with low basal LH were confirmed by Shrim et al. and Fleming et al. in both of their retrospective analyses.17,18 Jurema et al. however, in their retrospective analysis of 230 in vitro fertilization (IVF) cycles called this into question by finding no statistically significant differences in basal LH levels between normal and poor responders.19 Based on the powering of the study, there were some questions raised as to the significance of their findings. Their assertions however were firmly backed by Bjercke et al. in their retrospective analysis of 2625 IVF cycles in 1652 infertile women where cycle day 1 LH levels post a long term GnRH downregulation protocol failed to discriminate between conception versus non conception cycles and ongoing pregnancies versus early pregnancy loss.

Table 1: Summary of available evidence for basal LH as an ORT.

<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>n</th>
<th>Study design</th>
<th>Findings/outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>Noci et al.</td>
<td>30 women</td>
<td>Case: 30 women Control: 45 women</td>
<td>Case (LH &lt;3) vs. control (LH ≥3) Basal LH &lt;3 correlated with: • Reduced estradiol peak (s.s.) • Lower number of follicles &gt;15mm during HCG administration (s.s.) • Lower number of follicles &gt;10 mm (NS)</td>
</tr>
<tr>
<td>1998</td>
<td>Fleming et al.</td>
<td>20 women</td>
<td>Case: 20 women Control: 41 women</td>
<td>Case (LH ≤0.5) vs. Control (LH &gt;0.5) LH ≤0.5 correlated with • Lowered estradiol levels (s.s.) • Lower oocyte yields (NS) • Lower fertilization rate (NS)</td>
</tr>
<tr>
<td>2003</td>
<td>Jurema et al.</td>
<td>183 cycles</td>
<td>Case: 183 cycles Control: 47 cycles</td>
<td>Case (good cycle prognosis) vs. control (poor cycle prognosis) • No difference in baseline LH levels (3.76 ± 0.15 vs. 3.85 ± 0.16 IU/L) between normal or poor responders (NS) • No difference in baseline LH levels (3.36 ± 0.29 vs. 3.85 ± 0.16 IU/L) in differentiating between cycles resulting in pregnancy (NS)</td>
</tr>
<tr>
<td>2006</td>
<td>Shrim et al.</td>
<td>41 women</td>
<td>Case: 41 women Control: 596 women</td>
<td>Case (FSH/LH &gt;3) vs. control (FSH/LH ≤3) Basal LH 1.6 IU/ml vs. 6.3 IU/ml (s.s.) Lower basal LH correlated with: • Lower estradiol peak (s.s.) • Reduced yield of oocytes and embryos (s.s.) • Lower fertilization rate (NS)</td>
</tr>
<tr>
<td>2005</td>
<td>Bjercke et al.</td>
<td>2625 IVF + ICSI cycles</td>
<td>Low LH values stratified No threshold of LH values were able to: • Discriminate between conception and non-conception cycles (s.s.) • Discriminate between early pregnancy loss and ongoing pregnancy (s.s.)</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>Weghofer et al.</td>
<td>632 women</td>
<td>Basal LH unable to discriminate between • Number of oocytes retrieved (NS)</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>Orvieto et al.</td>
<td>268 IVF cycles</td>
<td>Normal prognosis cycles analysed between LH ≤3 and LH ≥3 IU/L. LH unable to discriminate between between: • Number of oocytes retrieved (NS) • Pregnancy rate (NS) • Number of embryos transferred (NS)</td>
<td></td>
</tr>
</tbody>
</table>

*s.s. - statistically significant; NS - not statistically significant; Retro - Retrospective
Interestingly, LH concentrations even in the polycystic ovarian syndrome (PCOS) subgroup analysis, with known endogenous elevated LH levels, showed no significance in outcomes. Orvieto et al. also found LH levels <3mIU/ml to lack association with IVF outcome in women with a good prognosis prior to commencing IVF.^

While the contrasting data upon the significance of basal LH sought to confuse, a growing body of evidence started highlighting the fact that the usefulness of LH was probably best drawn upon when used in conjunction with FSH.

Mukherjee et al. appear to be one of the first to demonstrate the value of the FSH/ LH ratio. In their retrospective analysis of 74 patients undergoing IVF, the group with a FSH/LH ratio ≥3.6 showed a poor response to controlled ovarian hyperstimulation (sensitivity 85.7%, specificity 95%), lower peak oestradiol levels and fewer follicles less than 15mm. When compared to the group with an FSH/LH ratio <3.6 they also demonstrated a higher cancellation rate and lower pregnancy rate. Liu et al. added further evidence to these findings in their retrospective analysis of 297 women who were less than 40 year of age and were undergoing their first cycle of IVF with an interest in the FSH/ LH ratio (FSH/LH ratio <2 vs. FSH/LH ratio ≥2). The primary outcomes were pregnancy and cancellation rates. There was a statistically significant higher cancellation rate in the FSH/LH ratio ≥2 group (25.5% vs. 18.6%). The FSH/LH ratio also demonstrated a strong negative correlation with Day 3 LH levels (r=-0.699 P<0.005) and a strong correlation with clinical pregnancy (r=0.10 P=.08). They however, failed to demonstrate lower rates of clinical pregnancies in the group. Brodin et al. in their prospective study also explored the impact of basal gonadotrophin levels (Day 2-4) and used live pregnancy and delivery rates as a primary outcome. The study encompassed 745 women who underwent 1328 IVF/ ICSI treatments. Their study demonstrated diminished ovarian reserve (DOR) and poor IVF treatment outcomes in women with low basal LH levels. The FSH/LH ratio however, demonstrated a much stronger statistical correlation to those endpoints then the basal LH levels. Low FSH levels (<6.7 U/I) combined with high LH levels (>4.9 U/I) were associated with the highest success rates (39%) and the opposite (high FSH / low LH) were correlated with the lowest pregnancy rate (22%). The study however, found it more useful to examine absolute values of FSH and LH individually in contrast to setting an absolute value for an FSH/ LH ratio as they found the statistical errors to be increased by doing so. Weghofer et al. in their retrospective analysis of 632 women failed to demonstrate any statistically significant correlation between basal LH levels (0.05 P <0.19) on ovarian reserve - defined as the number of oocytes retrieved. There was however, a demonstrable significant correlation between the FSH/ LH ratios (-0.11 P <0.01) and ovarian reserve in their analysis. They also demonstrated a low baseline LH with borderline FSH levels (10-15mIU/ml) further enhanced the negative predictive values of borderline FSH levels. Shrim et al. evaluated 1434 IVF cycles retrospectively and demonstrated impaired cycle outcomes in women with LH <12 mIU/ml where their FSH/LH ratio was ≥3. In order to potentially exclude patients with a potential peri menopausal state, a secondary analysis was conducted where patients in the borderline normal to elevated range (LH >8mIU/ml) were excluded. There was still a statistically significant reduction in cycle outcomes. Orvieto et al. payed special attention to patients undergoing both agonist (n = 131) and GnRH antagonist (n=137) protocols , and once again both subgroups achieved lower pregnancy rates in contrast to their controls (11.1 vest 27.7 and 8.3 vs. 31.9) with FSH/LH ratios ≥2 or ≥3. It is of note that the FSH/ LH ratio still maintained its discriminative ability in this population who were selected on the premise that they had a favourable prognosis for ART (women <35 years of age with day 3 FSH <15U/L undergoing up to their 3rd IVF cycle). Barosso et al. showed similar findings in a subgroup analysis of 28 patients with an FSH/ LH ratio >3.^

**DISCUSSION**

The evidence for the role of LH in ovarian reserve assessment is equivocal when used independently. The divided opinions on its use as an independent predictor are not without physiological basis. It may be the case that low early follicular LH levels probably show reduced activity of one or more of the known ovarian autocrine and paracrine regulators. Also, they are possibly indicative of an impaired balance between the gonads and pituitary itself. However, we must bear in mind that measured levels of LH do not necessarily correlate with LH’s bioactivity- which understandably has not been explored by any of the above mentioned studies but will be a parameter of interest in judging actual hormonal activity.

The incongruent data probably adds strength to the existence of an ‘LH ceiling’ which is unique to each individual - a hypothesis which proposes that ovarian follicles require a threshold level of LH stimulation for early follicular development but then reaches a ceiling value beyond which normal development ceases. This could potentially make LH difficult to independently interpret in the earlier follicular period and potentially explain the contrasting findings of the above mentioned studies. Also, a potentially confounding factor may be the fact that it may have a large intra and inter cycle variability similar to FSH as well which makes reproducibility an issue.

There can however, be more confidence in evaluating the data with regards to the FSH/ LH ratio. Most of the studies support the correlation between it and the assessment of ovarian reserve. The favourability of the ratio is on the premise that a lower FSH and a higher LH
correlates better with a larger number of antral follicles than either marker alone. The ratio is also probably more of a real time validation of the synergistic nature of the relationship between LH and FSH in the early follicular period of the ovarian cycle. Most of the evidence points towards a FSH/ LH ratio >2 being correlated with DOR and becoming more significant whilst approaching a ratio of 3. It is of note that a number of the studies have moved to reiterate the importance of generating individualized data from centres in terms of using LH as it seems likely that the available data is not universally applicable due to factors such as the heterogeneity of individual patient populations, laboratory assays which themselves show inter laboratory variation of LH quantification and stimulation protocols. For the purposes of establishing LH as an ORT, these findings need to be viewed in line with the current opinion and controversies surrounding ORTs and their value. In order to give further context to these issues, we use two ovarian markers which look the most promising in becoming established as the ‘ideal’ markers of ovarian reserve - Antimüllerian hormone (AMH) and the antral follicle count (AFC) - as reference points.

The issues are:

Reliability

For the ORTs in general, the reliability of the test becomes a factor of paramount importance. This is the ability of a test to achieve a consistency of measure across time and demonstrate reproducibility. The ability to do so would give clinicians much more confidence in the tests and its results temporally.

AMH is produced from the granulosa cells of the preantral and antral follicles and is easily measurable in the serum. This serves as a direct representation of the ovarian follicular pool and seems to mirror the decrease in ovarian reserve through reproductive life. In addition, it has an intra- cycle and inter - cycle variability that is low which allows for it to be measured at any part of the menstrual cycle. Even with such a firm physiological backing however, recent evidence has shown issues with reliability due to inter assay variation and this has prompted caution in interpreting and applying available data on AMH to clinical practice.

AFC on the other hand, is the sum of antral follicles on both ovaries as visualized through transvaginal ultrasonography. Histologically, the number of antral follicles visualized appear to correlate with the number of primordial follicles at any age. Though arguably subjective in nature due to ultrasonography being operator dependent, it has demonstrated good inter observer and inter-cycle reliability in experienced centres. As analysed above, LH alone does not seem to be able to provide the sort of consistency of measurement that would deem it reliable as it is endogenously prone to disturbance. However, it does seem that the FSH/LH ratio shows some potential in quantifying ovarian reserve.

Accuracy

The accuracy of a test refers to its ability to predict the outcome of interest correctly. In terms of ovarian reserve testing, this would entail accurately quantifying ovarian reserve. Through this, the ORT should be able to identify women who display severe DOR and hence a poor IVF prognosis and on that premise exclude them from treatment or at the very least adjusting their expectation. On the basis of available data, current ovarian reserve tests are rather poor as screening tests of ovarian reserve either individually or when used in combination. Amongst the basal ORTs however, both AFC and AMH have demonstrated the best performances for predicting poor ovarian response. One of the means by which ORT performance is tested is through receiver operating characteristics (ROC) regression curves. Various cutpoints for sensitivity and specificity of a diagnostic test are plotted graphically and a diagnostic threshold which optimizes both the sensitivity and specificity for the test is elicited.

For AMH, the ROC have demonstrated a high degree of accuracy (AUC 0.78: 95% CI 0.72-0.84) for poor ovarian response. In addition, more than half the studies available on AMH have reported a specificity of more than 85%. The ROC for AFC seem to mirror this as well and have a high degree of accuracy (AUC 0.76 95% CI 0.70-0.82) in predicting poor ovarian response. A low AFC of 3-4 follicles across both ovaries also displays a high specificity (73-100%) to predicting poor response. This however, is marred by its consistently lower sensitivity (9-73%) which thereby limits its clinical use.

It is unfortunate that the available data on LH does not permit the comparison of the statistical measures of performance at present. We can however, draw upon the success and failures of both AFC and AMH to help delineate criteria for further studies relating to basal LH or the FSH/ LH ratio to help delineate if the test is clinically useful. The ORT ideally has to display a high degree of validity (high sensitivity and specificity). These measures of validity can then be plotted on ROC curves which can help identify test cut off points that optimize both sensitivity and the specificity of the test and hopefully make it applicable in a clinical setting.

Implications

Amidst all this controversy, it must be emphasized that ORTs are meant as screening tests and hence do not
diagnose but are suggestive of DOR. Even when suggestive, it must also be borne in mind that many women who have displayed DOR can and have achieved pregnancy and live births especially if they are younger. On this basis, it would be very difficult to completely exclude women from ART given the psychosocial implications it may have on their lives. Hence a more realisable aim of using ORTs at present would be in counselling couples regarding the potential benefit versus risks of initiating ART and in realistic expectations of outcome as well as in modifying treatment strategy.

To counteract this element of uncertainty, the evidence does seem to suggest a greater benefit in individualizing these results to patients – probably on the premise that the determinants of ART success are multifactorial in nature and not just based on gonadotrophin levels alone. For example, a good screening test requires a high positive predictive value (probability that a woman who tests positive truly has DOR) to gain widespread use. However, this becomes dependent on the prevalence of the condition in the screening group as well. If inappropriately used in a group with low risk i.e. women who are curious about their fecundity it has the potential to produce a large number of false positives. Another example is seen by the way many current prediction models fail to take female age into consideration which is puzzling given its documented correlation with IVF outcome.

**Definition**

It should be evident from the heterogeneity of outcomes described with regards to LH that there is no current consensus or universally accepted definition of ovarian reserve. This has resulted in a variety of good quality studies that have used various differing end points. This has indirectly made the task of formulating strong evidence, for instance through a meta-analysis or systematic review harder. In addition, given the small amount of data available with regards to LH, it would be expected that less value may be procured from the data.

An example of this is mirrored in the studies surrounding AFC as well. Though well characterised as an ORT, there is still no consensus on which size of follicles have to be measured on ultrasound. Also, the literature seems to have differing cut off values with regards to poor and hyper response also. One of the means that has proven successful in negating this uncertainty is through developing centre based cut off values to determine outcomes. Hence, while flagging the possible contribution that LH may add to the ORTs, it must not be forgotten that it is not without its inherent pitfalls.

**CONCLUSION**

Based on the available evidence, it does appear that LH has a yet under recognized role in assessing ovarian reserve or response, especially with regards to the FSH/LH ratio. This is especially significant with a FSH/LH ratio ≥ 2 as it approaches 3. Ovarian physiology serves as a strong validation to the above relationship and the preliminary evidence at least suggests that FSH/LH ratio can be useful in clinical practise if not guide it. The fact that it is a readily available test prior to the initiation of ART increases its appeal.

It must be reiterated however, that the available evidence for LH use alone is remarkably sparse. Its role needs to be validated through further studies and especially through higher quality data models i.e. systematic reviews. Given that the available ORTs are fraught with question regarding their usefulness and reliability, LH may add further information either on its own or may help optimize available models.

With regards to further research initiatives, the following issues have been identified in the reviewed literature and should be taken into consideration in order to formulate quality evidence to guide clinical practice:

1. A uniform definition or consensus statement of ovarian reserve - to create comparable data pools from multiple centres.
2. Statistical measures of test performance
3. Individualized centre based prediction models of data to create personalized cut off levels.
4. Considering age to be part of the prediction model.
5. Accounting for laboratory assay variability in biochemical measurements.

If examined in this manner, the preliminary evidence does hold promise for the FSH / LH ratio to be used in estimating ovarian reserve. Having said that, it must be reiterated that it is highly unlikely that LH will be the missing piece of the jigsaw in ORTs but rather offers hope in complementing or expanding the clinical information we can garner from them.

**Funding: No funding sources**

**Conflict of interest: None declared**

**Ethical approval: Not required**

**REFERENCES**


30. Broer SL, Disseldorp JV, Broeke KA, Dolleman M, Opmeer BC, Bossuyt P et al. Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing


DOI: 10.5455/2320-1770.ijrcog20140302