CLINICAL PERSPECTIVES
IN REPRODUCTIVE ENDOCRINOLOGY

OVARIAN STIMULATION MONITORING: A ROUNDTABLE REVIEW

MODERATOR

Leo Bonaventura, MD
Indianapolis Fertility Center

PARTICIPANTS

Richard E. Blackwell, MD, PhD
Division of Reproductive Biology and Endocrinology
University of Alabama at Birmingham

Robert Casper, MD
Reproductive Biology Unit
Toronto General Hospital
Toronto, Canada

William D. Schlaff, MD
Reproductive Endocrinology
University of Colorado Health Sciences Center
Denver, Colorado

Richard Scott, MD
Department of Obstetrics and Gynecology
Uniformed Services University of Health Sciences
Bethesda, Maryland

James M. Wheeler, MD, MPH
Department of Obstetrics/Gynecology
and Reproductive Endocrinology
Baylor College of Medicine
Houston, Texas

INTRODUCTION

There are three primary goals of monitoring gonadotropin-stimulated cycles: to assess the adequacy of response to ovarian stimulation regimens, to determine the optimal timing of hCG administration, and to reduce the risk of ovarian hyperstimulation syndrome (OHSS) and multiple births. When selecting the appropriate techniques, the clinician must give consideration to medical as well as financial factors.

A distinguished panel of physicians assembled in Boston in July 1993 to evaluate methods of monitoring gonadotropin-induced ovarian stimulation as well as selected techniques of assessing adequacy of response. The panel reached a consensus on many issues. These recommendations can be used by the clinician to help clarify the monitoring expectations in gonadotropin-stimulated cycles. Although these guidelines represent a consensus of the panel, many individual factors must be considered by the clinician on how to best monitor patients. Therefore, these decisions must be made on a case by case basis.

ESTRADIOL MONITORING

Richard Blackwell, MD, PhD

Thirty years ago, clinical investigators found that monitoring urinary estradiol measurements could be used to predict the incidence of complications associated with controlled ovarian hyperstimulation (COH). Today, serum estradiol levels have replaced the use of urinary measurements. Many specialists believe that estradiol monitoring should be used in conjunction with ultrasound follicular monitoring to help them make the most informed decision about treatment.

Because estradiol levels and ultrasound monitoring of follicular response are complementary, both play a useful role in reducing the potential risk of multiple gestation and ovarian hyperstimulation syndrome.

Additionally, clinical indications exist where a baseline estradiol level would also be appropriate, such as when a patient presents with a residual cystic structure on ultrasound.

In a recent review of the literature that evaluated the role of estradiol monitoring versus ultrasound monitoring, Schoemaker et al[1] reported that neither method used alone proved superior, but definite advantages exist when both methods are used together, albeit for different purposes. One purpose is to predict the potential for hyperstimulation syndrome, and the other is to indicate the time for hCG administration.

While Schoemaker's conclusions may be true, the panel agreed that estradiol and ultrasound monitoring have synergistic roles. For example, with regard to safety, a high estradiol level can help predict which patient might hyperstimulate, but it is not as accurate in predicting multiple births. Ultrasound scanning provides the information about follicle size and number that alerts the clinician to an increased risk of multiple gestation.

The panel also agreed that physicians should be aware of the following issues when interpreting raw estradiol laboratory values.

Due to the wide variability among the currently employed RIAs, serum estradiol values should be interpreted with caution.

Typically, RIA kits from one company include serum standards referenced against another company's serum standards but do not reference a national standard. Furthermore, these standards may change at any time without notice to the user. Such variability contributes to discordant findings within an institution and between institutions.

As shown in Table 1, Hershlag et al[1] found significantly different values for estradiol in the same serum samples at different laboratories even when the same RIA kit was used.

FOR SAME SAMPLES AT FIVE

Institution	Kit	E2 (pg/mL)
A	Serono (Diagnostics)	87.3 ± 25.9
B	Pentax	78.1 ± 30.4
C	Diagnostic Products	54.6 ± 34.1
D	Diagnostic Products	42.0 ± 25.8
E	Pentax	45.8 ± 32.7

At the University of Alabama, an internal estradiol assay is used, so the physicians are familiar with its idiosyncrasies. Knowing that it overestimates estradiol secondary to cross-reactivity, they interpret the results accordingly. However, many hospitals and reference labs do not have dependable internal standards. They must rely on the materials provided in commercial kits, which may yield suboptimal results.

Improved quality control in commercial kits and within labs is critical to effective estradiol level monitoring. The University of Alabama laboratory assays about 50,000 various hormone samples annually and randomly inserts into its program known standards for each assay to challenge results and confirm that they are in the right range. Furthermore, when an assay is performed, three tubes at multiple dilutions are utilized, yielding up to nine points per assay. In contrast, commercial labs often run a single or duplicate sample per assay, which results in too much variability to ensure accurate results.

TIMING AND INTERPRETATION OF ESTRADIOL LEVELS VARY FROM ONE INSTITUTION TO ANOTHER.

At the University of Alabama, estradiol levels are monitored at day 8 or 9 and then again at day 11 or 12 of a cycle. Other physicians think one estradiol level is enough and point out that an estradiol level of 800 pg/mL in a patient with seven follicles is certain to double the next day. Still others think it is possible to limit estradiol monitoring to one level taken on day 11, along with a single scan taken the same day, and probably project ovulation with a reasonable degree of accuracy.

Some centers may have poorer overall results using gonadotropins than other centers because the clinicians have less experience or treat patients very conservatively. For example, in an attempt to avoid hyperstimulation, some programs may withhold hCG when estradiol levels are above 2500 pg/mL, which could result in a large number of canceled cycles. Although everyone is concerned about ovarian hyperstimulation syndrome (OHSS), the fact remains that OHSS in the average practice of the reproductive endocrinologist is very rare.

In general, the decision to withhold hCG should be based on the follicle, the patient, and the circumstances and not solely on a specific estradiol level. Estradiol levels should be perceived as a range, not as an absolute number.

For example, hCG administration may be appropriate for a 42-year-old patient with estradiol levels of 4500 pg/mL but inappropriate for a younger patient with levels of 2000 pg/mL and six mature follicles.

In higher risk situations, such as the possibility of multiple gestations or ovarian hyperstimulation, hCG could be withheld after an in-depth discussion with the patient about the risks involved in taking hCG.

INTERPATIENT VARIABILITY IN ESTRADIOL LEVELS IS ALSO AN ISSUE

There is also significant interpatient variability in serum estradiol levels. Some patients respond more slowly than others regardless of the protocol. For example, Ian Craft's group in England reports on patients with estradiol levels of 100 on days 6, 8, or 10, but by day 16 they reach levels of 1100. This type of response is more sudden in comparison to patients with hypothalamic disorder, who are typically slow responders. Eventually they demonstrate a normal estradiol response, but it may be out of phase by a week or more.

SUMMARY

For optimal monitoring of ovulation induction, estradiol monitoring should be used in conjunction with ultrasound monitoring. Physicians should be aware that there are wide variations in the results of estradiol assays, and opinions vary about when and how often to measure estradiol levels. In addition, conservative approaches to monitoring (ie, using levels that are too low to avoid the possibility of hyperstimulation) may result in a number of canceled cycles. Physicians should also be aware that patients with low estradiol levels on days 8, 9, or 10 may eventually demonstrate a normal estradiol response.

ULTRASOUND MONITORING

Richard Scott, MD

Ultrasound monitoring during ovulation induction has been, and continues to be, an extremely valuable addition to the practice of reproductive endocrinology. Although this technology has been available for more than 20 years, it was not until the advent of high resolution transvaginal sonography in the late 1980s that its potential in ovulation induction monitoring was fully realized.

Sonography plays an important role in the assessment of the ovary and uterus during ovulation induction. In the ovary, it provides physicians with important information about the size and number of follicles, timing of hCG administration, and ovarian patterns associated with hyperstimulation syndrome.

Uterine sonography helps define and assess important endometrial patterns that have been determined to have a high predictive value in assessing implantation and pregnancy outcome.

Additional benefits of ultrasound include the fact that scanning is relatively simple to perform and well tolerated by patients, the data are visually represented, and the findings are highly reproducible.

While ultrasound scanning plays an important role in ovulation monitoring, care must be taken to assure reliable, quality scans and measurements.

WHEN TO INITIATE ULTRASOUND MONITORING

A review of the literature about ultrasound monitoring indicates that there are really no data that define the optimal time to obtain a baseline scan or whether a baseline scan is even necessary or indicated. Formal recommendations for when to initiate scanning have not been developed, although many reproductive endocrinologists do baseline scans between days 6 and 8.

Since ovulation induction is such a highly individualized process, it may be impossible to define a specific protocol for monitoring and, as such, the process remains more of an individual's preference.

Although there is no clear agreement at this time about baseline scanning, a few general recommendations and suggestions can be made:

1. Baseline scans are more important in contiguous (consecutive) cycles than noncontiguous (30 days or longer between treatment) cycles because, in general, they provide additional information about persistent follicular or corpus luteal cysts.
2. If one is going to monitor with ultrasound only, a baseline scan is critical to determine the appearance of the uterus and ovaries. This baseline should be compared with meticulously performed follow-up scans.
3. If estradiol levels are going to be used with ultrasound monitoring, it is recommended that the baseline scan be done on the same day as the patient's first estradiol level. This combination baseline assessment, in addition to maximizing patient convenience, also provides a secondary quantitative reference point (estradiol level) from which to correlate the ultrasound findings.

WHEN TO ADMINISTER hCG

There are no prospective studies that specifically evaluate the optimal diameter of the lead follicle at the time of hCG administration. Therefore, most authors have used their own personal protocols, which were developed from a large volume of practical experience. This relatively empiric approach has been a practical way to approach hCG administration because everyone knows what results occur when hCG is administered at a particular follicular size. The range of practical experiences, as well as differences in equipment and measurement techniques, helps explain why some clinicians give hCG when follicles are 14 mm in diameter, while others wait until follicles are 18 mm in diameter.

RELIABILITY OF MEASUREMENTS

Although ultrasound scanning is being widely applied, there are a number of considerations that need to be understood about the reliability of the measurements.

A primary and fundamental consideration is this: There is no substitute for good training and technology. Because in many cases the same patient may be scanned by multiple operators, care must be taken to ensure that uniform measurement protocols are being applied.

In general, all scans should be performed utilizing a transvaginal approach and conducted with a high-frequency (between 5 MHz and 7.5 MHz) endovaginal transducer.

To maintain consistency and reproducibility, a standard measurement technique needs to be defined and followed. For example, should the follicle be measured in two dimensions (such as 8 mm by 6 mm) or in the largest dimension only (such as 8 mm)? This is less important if the follicle is round but very important if the follicle is oval.

The need for measurement protocol consistency is highlighted when one looks to the literature for the causes of unreliable measurement and determines observer variation to be critical.

Forman et al[3] conducted a study to estimate variation in repeated measurements made by the same observer and compared this with the variation in measurements made by several experienced observers. The investigators concluded that intraobserver variation was 2 mm and interobserver variation was 3 mm, an enormous range.

In light of these data, it would be ideal to have just one person perform all your follicular scanning. However, that may not be practical. Regarding satellite centers, it may also be worthwhile to have sonographers from outlying areas train for a week or two at the main institution to increase confidence in their ability.

PREDICTING OVARIAN HYPERSTIMULATION SYNDROME (OHSS)

OHSS is the most serious complication of gonadotropin treatment. The risk for OHSS has traditionally been associated with the peak estradiol level during the stimulation cycle. However, OHSS may occur in the presence of normal or even low estradiol levels. Furthermore, the presence of high estradiol levels does not always prognosticate the development of OHSS. These facts demonstrate that estradiol monitoring alone may not be sufficient to predict OHSS.

Sonographic monitoring, on the other hand, has been found to be a very useful adjunct in this area by allowing us to predict the occurrence of OHSS based on the overall number of follicles, the number of intermediate follicles, and their distribution within the ovary itself (Fig 1).

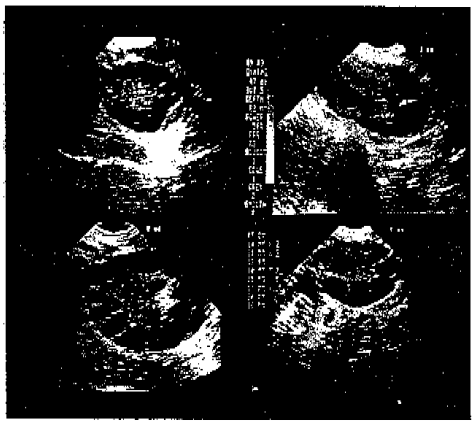


Fig 1. — Four sonographic visualizations of the "Necklace sign," which is correlated with the risk of OHSS.

Groups of follicles, 6 mm to 8 mm in size, arranged in an annular pattern around the periphery of the ovary at the onset of stimulation have been shown to correlate fairly well with the patient's subsequent risk of OHSS. This is the "Necklace sign."

In addition, Blankstein et al[4] discovered that patients with hyperstimulation syndrome usually had a larger number of relatively smaller follicles at the time they were given hCG or on the day following hCG administration. Mild OHSS occurred in the presence of eight to nine fol-

licles (68.7% of which were 9 mm to 15 mm in size. Most of the patients who experienced moderate to severe OHSS had follicles smaller than 16 mm.

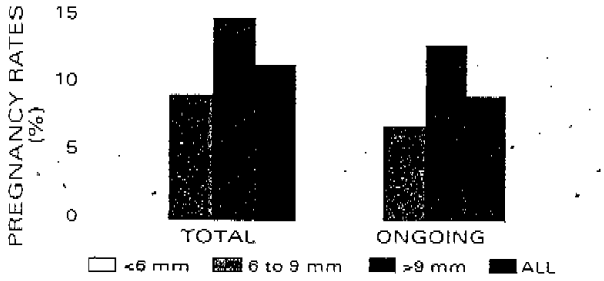
These data and the panel's experience suggest that looking at the number of mature follicles is not enough to get a clear picture of hyperstimulation syndrome. The challenge, of course, is to determine if less aggressive stimulation can lower the risk of hyperstimulation in these patients.

SONOGRAPHIC MONITORING OF THE ENDOMETRIUM

In 1989, Gonen et al[5] demonstrated a significant correlation between endometrial thickness and pregnancy rates in IVF patients. Specifically, endometrial thickness was significantly greater in women who conceived than in women who did not conceive.

These results were confirmed in patients undergoing ovulation induction with exogenous gonadotropins in a recent study by Dickey et al.[6] As shown in Table 2, no pregnancies occurred when endometrial thickness was

TABLE 2 — PER CYCLE FECUNDITY AND CONTINUING PREGNANCY RATES ACCORDING TO ENDOMETRIAL THICKNESS



less than 6 mm on the day of hCG administration. Per cycle continuing pregnancy rates were 12.6% when the endometrium was 9 mm thick or more, but only 6.9% when the endometrium was 6 mm to 9 mm thick.

There was also a correlation between endometrial patterns on the day of hCG administration and per cycle fecundity, with the total number of pregnancies being highest in patients with a triple-line pattern.

It's important to note here that an entirely homogeneous hyperechogenic endometrium (type A), an endometrium with an intermediate pattern [illegible] ized by the same reflectivity of ultrasound [illegible] myometrium (type B), and a multilayered [illegible] are [illegible] echogenic outer and middle [illegible] lines at the inner hypoechogenic region [illegible] now described in only two ways: as hyperechoic [illegible] patterns (formerly type A) and triple-line patterns (formerly types B and C).

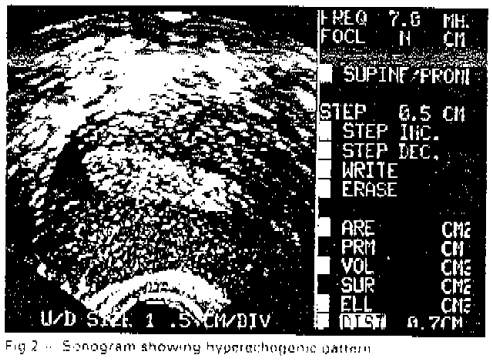


Fig 2 — Sonogram showing hyperechogenic pattern.

Fig 3 — Sonogram showing triple-line pattern.

The panel agreed that pregnancy is most [illegible] occur with a triple-line pattern endometrium [illegible] more than 9 mm thick. Pregnancy is unlikely [illegible] if the endometrium is hyperechogenic and less [illegible] 6 mm thick. In addition, the rate of pregnanc[illegible] able, but diminished, when endometrial thick[illegible] ranges from 6 mm to 9 mm.

They also agreed that the endometrium s[illegible] always be evaluated in patients taking clomiph[illegible] citrate. Because 20% of women will have a thi

endometrium while taking this drug and this hyperechogenic pattern is generally inconsistent with a positive pregnancy outcome. It is recommended that an ultrasound scan be done 1 week after the first cycle on clomiphene citrate. If it shows a hyperechogenic (type A) pattern, then clomiphene citrate should be discontinued, and the patient should be considered for possible estradiol supplementation or gonadotropin therapy.

SUMMARY

Because of the highly individual nature of ovulation induction monitoring, there is no specific protocol for initiating ultrasound follicular monitoring or administering hCG. Most physicians still follow the empiric stimulation schedules they developed before the advent of high-resolution ultrasound.

The panel recommends scanning on the same day the first estradiol level is measured. To get the most reliable, reproducible measurements, the panel recommends training and development and adherence to a scanning protocol.

Ultrasound monitoring of follicles has greatly improved the ability to predict ovarian hyperstimulation syndrome while monitoring of the endometrium has increased the ability to predict pregnancy outcome.

Pregnancy is more likely to occur if a triple-line pattern is observed and the endometrium is more than 9 mm thick. It is unlikely to occur if the endometrium is less than 6 mm thick, and it occurs at a variable rate when thickness ranges from 6 mm to 9 mm. It should be noted that when clomiphene citrate is administered alone for ovulation induction, an endoscan is useful to observe and characterize the endometrium.

MONITORING LUTEINIZING HORMONES (LH) WITHIN THE STIMULATED CYCLE

William Schlaff, MD

Almost 10 years ago, investigators reported reduced pregnancy rates in IVF patients who had increased levels of LH during the follicular phase.[7] In 1993, Shoham et al[8] reviewed the published literature and concluded that a raised serum LH concentration during the follicular phase may increase the risk of infertility and early pregnancy loss.

It has been demonstrated that advanced knowledge of premature luteinization may be clinically useful in the care of patients undergoing in vitro fertilization. If this proves to be a consistent observation, physicians could cancel oocyte retrieval or advance the time of retrieval, if appropriate.[9] In this way, pregnancy rates could equal those achieved in cycles during which no premature LH surge occurred.

Knowledge of premature luteinization is probably even less relevant in women undergoing ovulation induction with gonadotropins because timing is not as critical as it is in IVF.

In addition, Hofmann et al[10] demonstrated that women who prematurely luteinize on gonadotropins alone are at high risk for premature luteinization in a subsequent ovarian stimulation cycle. In fact, the risk is equally high despite pituitary desensitization in the subsequent cycle.

These data seem to suggest that there is a difference between eliminating spontaneous LH surges and eliminating levels of LH that can cause premature luteinization.

The panel does not believe that luteinization, premature luteinization, or pregnancy rates can be predicted based on LH measurements. As a monitoring tool for stimulated cycles, LH levels seem to have no relevance and there appears to be no justification for incurring extra expense to measure them.

SUMMARY

While monitoring LH levels to predict premature luteinization in stimulated cycles has been proven to be of some value in IVF programs, it appears to be of little or no clinical value in ovulation induction.

PROGESTERONE MONITORING

James M. Wheeler, MD, MPH

Progesterone increases in the mid-luteal phase, and traditionally that is when these levels are measured. As indicated in Figure 4,[11] progesterone levels, unlike LH levels, have a fairly wide peak, which allows easy assessment.

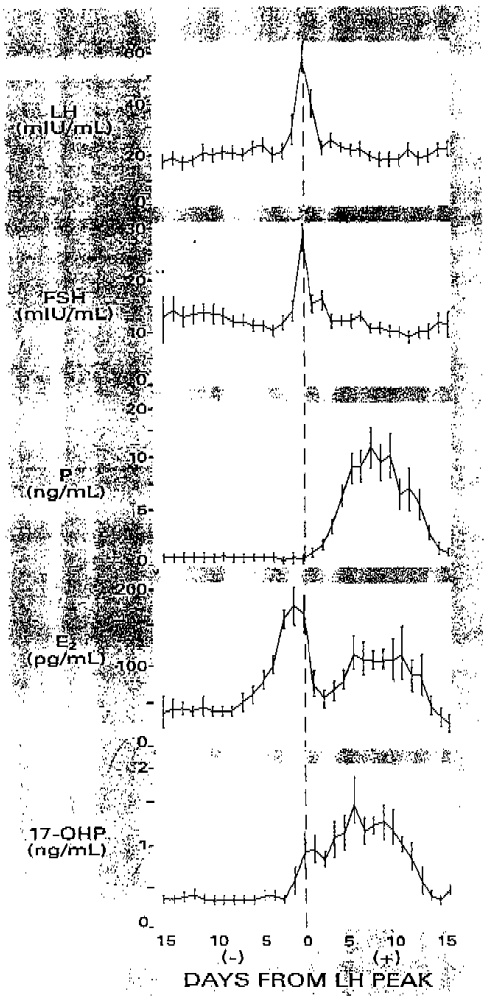


Fig 4.— Mean values of LH, FSH, progesterone (P), estradiol (E_2), and 17-hydroxyprogesterone (17-OHP) in daily serum samples of 9 women during ovulatory menstrual cycles. Data from different cycles combined with the use of the day of the mid-cycle LH peak as the reference day (day 0). The vertical bars represent one standard error of the mean.

Everyone has been taught that 3 ng/mL to 5 ng/mL of serum progesterone indicates ovulation, but it is also important to consider the quality of ovulation and the quality of corpus luteum function.

The algorithm in Figure 5 reflects one approach to progesterone assessment. Low levels (less than 5 ng/mL) are suggestive of anovulation. High levels (10 ng/mL or greater) indicate that corpus luteum function is probably adequate.

When levels are in the mid-range (5 ng/mL to 10 ng/mL), repeat measurements and a late luteal biopsy should be considered.

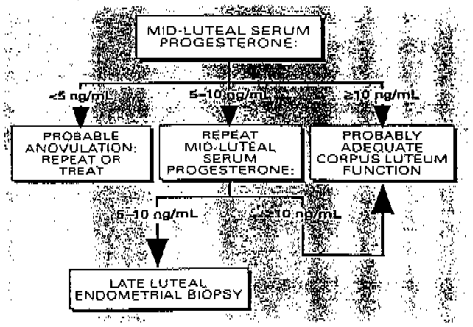


Fig 5.— Algorithm.

METHODS OF MEASUREMENT

At Baylor College of Medicine, mid-luteal progesterone is timed indirectly using the most economical and easy methods, usually basal body temperature (BBT) and an ovulation prediction kit. If the level is not satisfactory, they try to obtain a more precise timing, usually by sonographic folliculogram. Fortunately, there are good quality assays available so that progesterone is reliably assessed.

There used to be a firm belief in endometrial biopsies, but today, reliance on endometrial biopsy has really decreased because of concerns about cost, discomfort, and, most importantly, a high incidence of "out of phase" biopsies in normal, and even conception cycles.

Pillet et al[12] analyzed standard methods of assessing ovulation (ie, BBTs, endometrial biopsy, or serum progesterone) and showed that endometrial biopsy was a little more accurate than serum progesterone. Progesterone alone was not really accurate enough. In fact, it was not much more accurate than BBTs. Therefore, the authors concluded that combining BBTs and serum progesterone is probably the most cost-effective way to assess ovulation in the luteal phase.

FUTURE PROSPECTS

It is exciting to look at the work that is being done with urinary and salivary metabolites of progesterone because both of these biologic sources allow a naturally more integrated assessment of progesterone function. As a result, there may be more use of urinary or salivary collection devices in the future.

In addition, the future may bring answers to these important questions about progesterone: What does it have to do with end-organ effects? How does it correlate with effects at the cellular level? Is it possible to have excellent progesterone levels yet still have a receptor defect? Is the more mature uterus different from the younger uterus in this respect?

And have we determined the best time to measure progesterone? Could it be that later luteal phase levels correlate better with biologic function? Or is the most important progesterone assessment preovulatory? We may actually be more interested in progesterone levels in the last third of the luteal phase. So far, monitoring progesterone pre-hCG has no clear prognostic value.

In conclusion, progesterone assessment is certainly widespread and measured with reasonable accuracy. And with the decline of endometrial biopsies, there may be a need for a little more research on the role of progesterone in ovulation assessment.

SUMMARY

There were mixed opinions about progesterone monitoring among the panel members. One thought it was the standard of care today, in keeping with the need to be cost-effective, and another thought it was not clinically useful. Obviously, more research needs to be done and the decision left to the individual physician.

LUTEAL PHASE MONITORING

Robert Casper, MD

To a great extent, luteal defects are a function of poor follicular development. Therefore, monitoring the luteal phase seems to have little relevance because it does not affect clinical outcome. However, if some physicians feel they need to monitor in this phase, the panel suggests that progesterone is probably a better monitoring tool than endometrial biopsy.

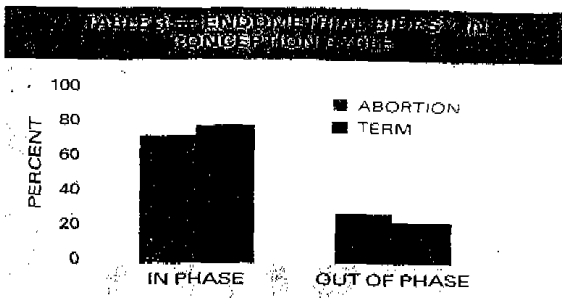
ENDOMETRIAL BIOPSIES HAVE NO CLINICAL IMPACT

A study by Balasch et al[13] showed that luteal phase deficiency diagnosed by endometrial biopsies really had no impact whatsoever on the outcome of pregnancy (Table 3).

The group did 1492 endometrial biopsies in more than 1000 infertile women and compared them with endometrial biopsies in a control group of 51 fertile women. All biopsies were done 1 to 3 days prior to the onset of menses, so they were all in the late luteal phase.

The results showed that 317 of 1036 endometrial biopsies in the first cycle were out of phase by more than 2 days—a 30.6% abnormality rate. Of the 207 women who had repeat biopsies, 95 were out of phase—a rate of 45.8%.

This is exactly the same percentage that would occur by chance. The investigators also reported that there was no effect of age on luteal phase defect diagnosis.



When the investigators did an endometrial biopsy in the cycle of conception, about 70% to 80% were in phase and about 20% to 30% were out of phase. However, as you can see in Table 3, there were no significant differences in the relative rates of spontaneous abortions and term pregnancies whether the biopsy was in phase or not.

In a study that included 12 women who received donor eggs during cycles induced by exogenous hormones,[14] an endometrium that was 6 mm thick with a triple-line pattern prior to the addition of progesterone was predictive of a favorable outcome. However, there was a poor correlation between endometrial biopsy findings and pregnancy. All but one endometrial biopsy was out of phase by 3 or more days. Thus, it appears that the endometrial biopsy is not a good test to determine the effectiveness of exogenous hormonal preparation of the endometrium. Therefore, it may be important to monitor progesterone during the luteal phase, but it is not recommended that single progesterone levels be used as a way of monitoring.

SUMMARY

Monitoring in the follicular phase is probably much more important than any kind of luteal phase monitoring. However, if physicians feel there is a need to monitor during this phase, the panel believes that progesterone monitoring is better than endometrial biopsies.

CONCLUSION

Leo Bonaventura, MD

Optimal monitoring for ovulation induction with gonadotropins may not exist. To use all of our monitoring techniques and protocols may be cost prohibitive indeed. The obvious goal is to make sure that we produce the ideal number of follicles to keep multiple births and ovarian hyperstimulation to a minimum. It was to that goal this panel directed its efforts. I will try to summarize the recommendations.

The use of estradiol alone is not satisfactory. Although it may help us predict poor responders, as well as the high responders at risk for ovarian hyperstimulation syndrome, it cannot tell us how many mature follicles are present. Estradiol works best when used with ultrasound. It is critical that the estradiol be done in the same laboratory for each patient. It is not possible to interpret estradiol reports on the same patient from different laboratories. It is also very important that ultrasound be done by vaginal technique and by the same ultrasonographer for each patient. Estradiol may be the rough guide that tells us when ultrasound should be started, but it is clearly the ultrasound that will tell how many follicles are mature, as well as the size of the endometrial lining. These are the signals to the practitioner to administer human chorionic gonadotropin to initiate ovulation. An estradiol in the area of 800 pg, a follicle size of approximately 18 mm in at least two dimensions, and an endometrial thickness of greater than 9 mm all serve to indicate a good induction and that hCG should be given. We recognize that follicle sizes and estradiol levels may vary from individual protocol to protocol, but the 18 mm follicle size, along with an estradiol level of 800 and an endometrial thickness of greater than 9 mm, would be a good guideline to follow.

The use of luteinizing hormone (LH) and progesterone levels to try to predict follicle maturity has not been reproducible enough to be used as routine in ovulation induction therapy. The costs of these tests do not warrant their use, and they should only be used in specific instances. One cannot deduce from LH or progesterone whether or not that cycle will be good or bad.

The use of endometrial biopsies and progesterone in the luteal phase as a monitoring device has been used for many years. In this monograph, there is in-depth discussion and review of endometrial biopsies and progesterone in the luteal phase. There is no general recommendation for monitoring. However, if a physician feels that the luteal phase needs to be monitored, this panel recommends that the use of progesterone is a better alternative than endometrial biopsy. Endometrial biopsies are not cost-effective for the information obtained.

Finally, the panel recommends that the use of estradiol with ultrasound is the most efficacious and cost-effective means of monitoring patients undergoing ovulation induction with gonadotropins.

REFERENCES:

1. Schoemaker J, v d Meer M, v Weissenbruch MM. Re-evaluation of the role of estrogens as a marker for ovulation induction. In: Lunenfeld B, ed. *FSH Alone in Ovulation Induction.* An Official Supplement to the Proceedings of a Symposium held at the 3rd World Congress of Gynecological Endocrinology, Madonna di Campiglio, Italy, February 1992. Pearl River, NY: The Parthenon Publishing Group; 1993:23-28.

2. Hershlag A, Lesser M, Montefusco D, et al. Interinstitutional variability of follicle-stimulating hormone and estradiol levels. *Fertil Steril.* 1992;58(6):1123-1126.

3. Forman RG, Robinson J, Yudkin P, et al. What is the true follicular diameter: an assessment of the reproducibility of transvaginal ultrasound monitoring in stimulated cycles. *Fertil Steril.* 1991;56(6):989-992.

4. Blankstein J, Shalev J, Saadon T, et al. Ovarian hyperstimulation syndrome: prediction by number and size of preovulatory ovarian follicles. *Fertil Steril.* 1987;47(4):597-602.

5. Gonen Y, Casper RF, Jacobson W, et al. Endometrial thickness and growth during ovarian stimulation: a possible predictor of implantation in in vitro fertilization. *Fertil Steril.* 1989;52(3):446-450.

6. Dickey RP, Olar TT, Taylor SN, et al. Relationship of endometrial thickness and pattern to fecundity in ovulation induction cycles: effect of clomiphene citrate alone and with human menopausal gonadotropin. *Fertil . Steril.* 1993;59(4):756-760.

7. Stanger JD, Yovich JL. Reduced in vitro fertilization of human oocytes from patients with raised basal luteinizing hormone levels during the follicular phase. *Br J Obstet Gynaecol.* 1985;92:385-393.

8. Shoham Z, Jacobs HS, Insler V. Luteinizing hormone: its role, mechanism of action, and detrimental effects when hypersecreted during the follicular phase. *Fertil Steril.* 1993;59(6):1153-1161.

9. Droesch K, Muasher SJ, Kreiner D, et al. Timing of oocyte retrieval in cycles with a spontaneous luteinizing hormone surge in a large in vitro fertilization program. *Fertil Steril.* 1988;50(3):451-456.

10. Hofmann GE, Bergh PA, Guzman I, et al. Premature luteinization is not eliminated by pituitary desensitization with leuprolide acetate in women undergoing gonadotrophin stimulation who demonstrated premature luteinization in a prior gonadotrophin-only cycle. *Hum Reprod.* 1993;8:695-698.

11. Thorneycroft IH, Mishell DR, Stone SC, et al. The relation of serum 17-hydroxyprogesterone and estradiol-17 ß levels during the human menstrual cycle. *Am J Obstet Gynecol.* 1971;111(7):947-951.

12. Pillet MC, Wu TF, Adamson GD, et al. Improved prediction of postovulatory day using temperature recording, endometrial biopsy, and serum progesterone. *Fertil Steril.* 1990;53(4):614-619.

13. Balasch J, Fabregues F, Creus M, et al. The usefulness of endometrial biopsy for luteal phase evaluation in infertility. *Hum Reprod.* 1992;7(7):973-977.

14. Shapiro H, Cowell C, Casper RF, et al. The use of vaginal ultrasound for monitoring endometrial preparation in a donor oocyte program. *Fertil Steril.* 1993;59:1055-1058.