OVARIAN RESERVE TESTING

Amanda Deadmond, MD, Positive Steps Fertility, Shreveport, LA
Christian A. Koch, MD, PhD, Professor, Division of Endocrinology, Diabetes, Metabolism, Department of Medicine, The University of Tennessee Health Science Center, Memphis, TN, Fox Chase Cancer Center, Philadelphia, PA. Christian.koch65@gmail.com, Christian.koch@fccc.edu
J. Preston Parry, MD, MPH, Adjunct Professor, Department of Obstetrics and Gynecology, Louisiana State University Health-Shreveport, Shreveport, LA, Positive Steps Fertility, Madison, MS. drprestonparry@gmail.com

Updated December 12, 2022

ABSTRACT

The ovaries affect far more than reproductive health. Estrogen affects cardiovascular, skeletal, mental health, and numerous other aspects of wellness. Additionally, ovarian dysfunction can reflect disequilibrium relating to multiple conditions. Efficient and effective ovarian testing can give women valuable answers about their fertility, time to menopause, and other conditions and symptoms they may face. Though no test is perfect, antral follicle count (AFC) and anti- Müllerian hormone (AMH) provide more sensitive and specific results that allow for the continuum of ovarian function, and have advantages over classic tests such as follicle stimulating hormone (FSH), estradiol, the clomiphene citrate challenge test (CCCT), and others. This chapter explores these and additional ovarian assays, their underlying mechanisms, and limitations that may favor one test over another depending on circumstances. Particular emphasis is given to evaluating perimenopausal status, procreation, and etiologies for amenorrhea.

INTRODUCTION

Ovarian endocrinology is dynamic. Years of quiescence are followed by oscillating secretion until near burnout, but some function remains even after menopause. “Ovarian reserve testing” assesses where the ovaries are within this spectrum. These measures seem to most clearly relate to oocyte quantity, as multiple other factors (especially age) meaningfully affect oocyte quality and fecundability. However, quantity and quality are not completely independent, as abnormal ovarian reserve testing has been linked to increased blastocyst aneuploidy (1).

This chapter will characterize the main biochemical and sonographic approaches used in both classic and modern testing. Moreover, an assay, like any tool, has value relative to the task to which it is applied. Accordingly, this chapter will also discuss application of ovarian reserve tests to several common areas: assessing perimenopausal status, evaluating ovarian reserve for fertility, and addressing primary and secondary amenorrhea. Use of these markers in assessing the male is covered elsewhere

(https://www.endotext.org/chapter/laboratory-assessment-of-testicular-function).

Because consensus can be difficult, the following summaries reflect trends, though different perspectives exist and the literature continues to evolve. Existing research on ovarian reserve testing is often confusing because of heterogeneity among tested populations (the general population, infertility patients of all ages, infertility patients more than 35 years old, etc., see also data from the Society for Assisted Reproductive Technology (SART), Figure 1, (2)). Additionally, one must always keep in mind that as with all screening tests, no single result is definitive, since findings must be interpreted in context and should be repeated or supplemented as appropriate.

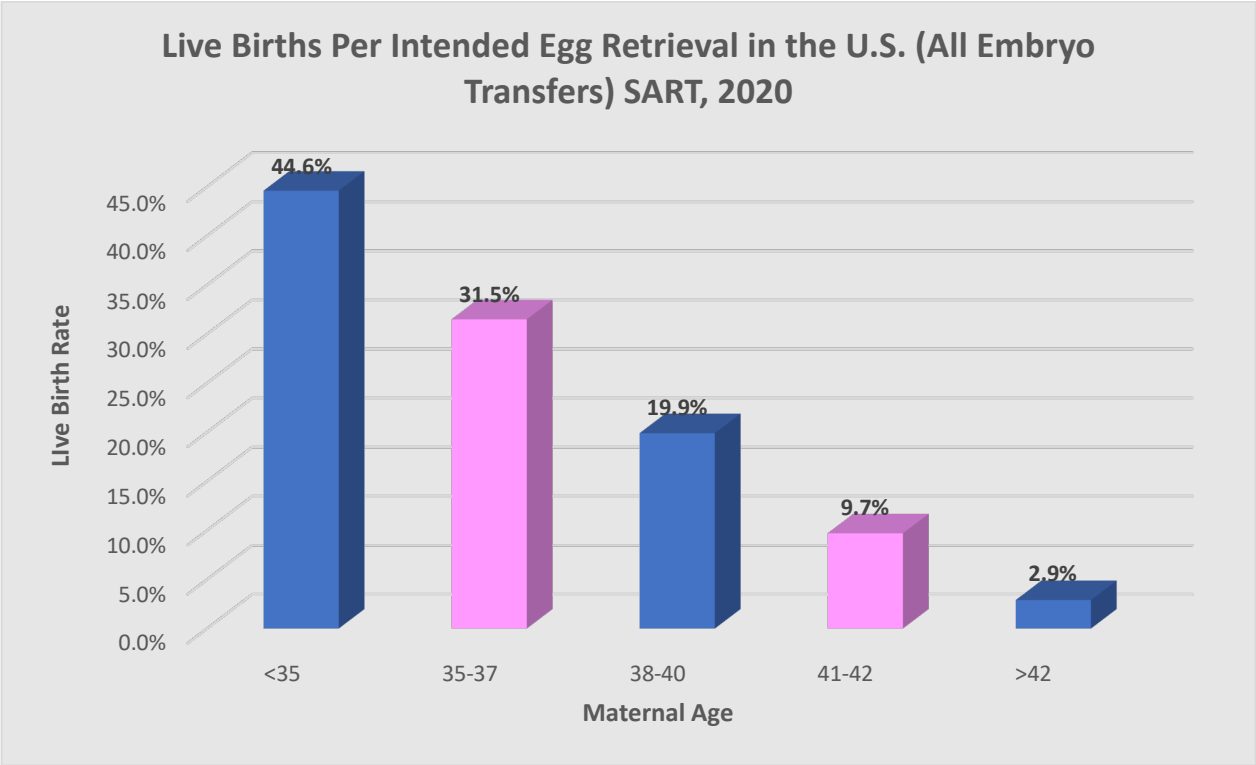

Figure 1. The relative effect of age on fecundity through in vitro fertilization (IVF) in 2020 according to the Society for Assisted Reproductive Technology (SART) (2).

MARKERS OF OVARIAN RESERVE

Follicle Stimulating Hormone (FSH)

MECHANISM

FSH was the first hormone directly linked to ovarian aging (3). It is secreted by the anterior pituitary and promotes the progression of antral follicles into dominant follicles. Feedback from estrogen, inhibin, and activin influence hypothalamic GnRH pulsatility, which determines pituitary FSH expression. Elevated

FSH levels can be seen with dwindling reserve, where a greater FSH stimulus is required to drive folliculogenesis, but elevated levels also can be found in normal ovarian reserve if measured at the time of the LH surge. Low FSH levels are seen prior to puberty or with hypogonadotropic hypogonadism. In addition to medical conditions that shift pituitary FSH expression, exogenous hormones and their modulators (clomiphene, letrozole, etc.), cimetidine, phenothiazines, and other medications can also shift levels.

TESTING

Many non-FSH substrates can induce an FSH-like effect. Without describing in detail the spectrum of FSH assays that bypass this challenge, for which an excellent review is available (4), in the clinical setting FSH is typically measured by immunoassay. The sample is usually acquired by phlebotomy (24-hour urine collections are rarely used) on menstrual cycle day three for ovulatory patients, with day one being the first full day of flow.

Testing on cycle days two, four, or five is not unreasonable, but if a normal result would prompt retesting, a day three measurement or a different assay is preferred. Multiple cutoffs are used, with FSH levels of >16.7, >11.4, and <10 mIU/mL reflecting high, moderately high, and normal levels based on the World Health Organization (WHO) Second International Standard (5).
Because ovarian reserve is on a continuum, any cutoff selected should relate to goals of balancing positive and negative predictive values, and this is an issue that applies to other measures of ovarian reserve as well. In amenorrheic patients, a random sample is preferred to testing after hormonally induced menses. In the setting of amenorrhea, a concurrent progesterone level (<2 ng/mL) is a reasonable control to ensure that one is in the follicular phase.

LIMITATIONS

Interpersonal and intercycle variation can be meaningful in patients at risk for moderately elevated FSH, which is why it has been called “Fluctuating Severely Hormone.” The problems with FSH’s sensitivity in part stem from it being a late marker of dwindling ovarian function, as summarized in Stages of Reproductive Aging Workshop + 10 conclusions (6). This limited predictive value is reflected in the NHANES III data, which showed 75% of women aged 40 to 44 years having normal levels at less than 10 mIU/mL, even though ovarian function is typically the rate limiting step at this age, and half of women aged 45 to 49 years had levels less than 11 mIU/mL (7). Sensitivity for FSH is often worse than specificity, with findings ranging from 11-86% and 45-100%, respectively (8). With anti-muellerian hormone (AMH) and antral follicle count (AFC) demonstrating better predictive value for ovarian response than FSH, these are more likely to be the tests of choice (9). Accordingly, relative to emerging alternatives, FSH testing increasingly is seen as less valuable than it used to be for procreative testing and more useful for evaluating perimenopausal status, hypergonadotropic and hypogonadotropic hypogonadism, and central precocious puberty.

Estradiol

MECHANISM

As with FSH, estradiol levels vacillate over the course of a menstrual cycle, peaking in both the late follicular and mid luteal phases. As ovarian reserve

declines, the follicular phase shortens because of decreasing feedback inhibition by follicles recruited during the previous cycle. (This is why the first clinical sign of decreasing ovarian reserve is shortening menstrual cycle length.) With the follicular phase starting earlier, estradiol levels start rising closer to menses (and the classic day three FSH peak actually can occur prior to menses). As a result, an elevated day three estradiol level could reflect diminishing ovarian reserve.

Elevated estradiol (>60-80 pg/mL) may also lead to an artificially normal FSH, where higher estradiol levels lead to feedback suppression of FSH. Conversely, estradiol levels <20 pg/mL on day three depending on the circumstances can be consistent with normal ovarian function, hypogonadotropic hypogonadism, or ovarian failure.

TESTING

Estradiol is also typically measured by immunoassay after phlebotomy. The sample is usually drawn at the same time as FSH levels or randomly when assessing amenorrhea. Estrone, the primary postmenopausal estrogen, and estriol, the primary pregnancy estrogen, are not typically tested when evaluating ovarian function. Also, because oral estrogens are typically metabolized into many byproducts (with varying activity), serum estradiol levels often won't reflect exogenous exposure. (However, transdermal estrogen administration can be monitored through serum levels.) Medical conditions, glucocorticoids, sex steroids, clomiphene, letrozole, GnRH agonists and antagonists, and other medications can alter estradiol levels, just as they could shift FSH levels.

LIMITATIONS

For many conditions, an estradiol level is a reasonable proxy for ovarian inactivity. However, for assessing decreasing ovarian reserve, estradiol is neither a sensitive nor specific assay (9). Accordingly, when used for measuring ovarian reserve, estradiol has its greatest value as an internal control to ensure that one is testing at the expected portion of the menstrual cycle (Figure 2).

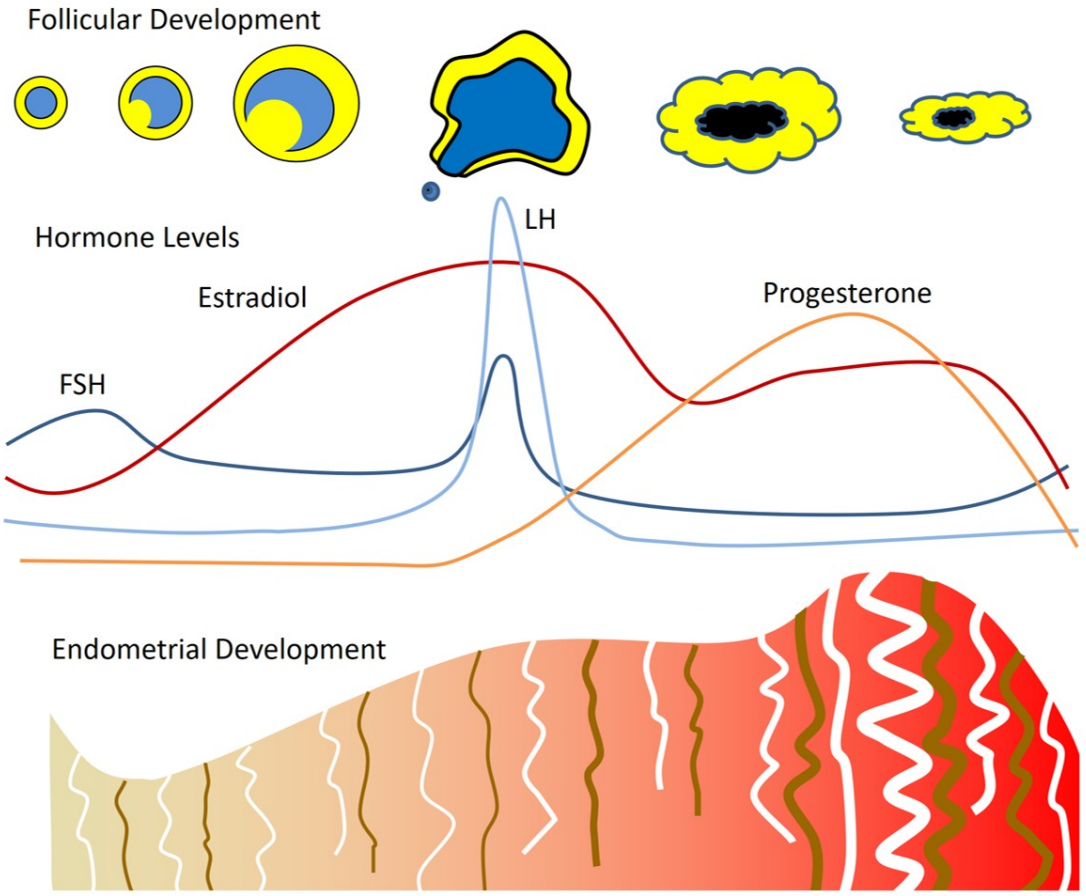

Figure 2. Ovarian, hormonal, and endometrial changes over the menstrual cycle. Adapted from Hall, et al., Hypothalamic gonadotropin-releasing hormone secretion and follicle-stimulating hormone dynamics during the luteal follicular transition (10).

Clomiphene Citrate Challenge Test (CCCT)

MECHANISM

The clomiphene citrate challenge test combines measurement of FSH and estradiol levels prior to clomiphene exposure and FSH levels after clomiphene exposure. Clomiphene is a selective estrogen receptor modulator (SERM) that inhibits negative feedback inhibition by estradiol on the hypothalamus. Normally, increased estrogen levels decrease GnRH pulsatility, resulting in lower FSH levels through negative feedback. By using clomiphene to block feedback inhibition by estradiol, there is an increase in FSH, which enhances follicular recruitment, and which is why clomiphene can be used for ovulation induction and superovulation.

TESTING

FSH and estradiol levels are assessed through immunoassay, as previously described. The CCCT is performed by having an FSH level drawn on the third day of the menstrual cycle, taking 100 mg of clomiphene orally cycle days five to nine, and then repeating the FSH level on cycle day number ten (11, 12). An estradiol level is also frequently drawn on the third day and sometimes on the tenth day as well.

LIMITATIONS

When assessing ovarian reserve for fertility, FSH is a limited measure of ovarian response and a poor predictor of pregnancy and estradiol is predictive of neither (9). When combining the two through the CCCT, it is difficult to assess the degree of benefit through receiver operator curves (9). If benefit is unclear, cost-effectiveness is even less so. Accordingly, other measures of ovarian reserve are increasingly used instead of the CCCT, although this assay is still more commonly used than other provocative tests, such as the exogenous FSH ovarian reserve test (EFORT) and the GnRH agonist stimulation test (GAST). The CCCT has particularly suboptimal value in anovulatory patients. The reason is that the CCCT is primarily used to help discriminate normal ovarian reserve from poor reserve in patients with potentially borderline function. However, the typical anovulatory patient tends to have robust reserve (PCOS, hypogonadotropic hypogonadism) or poor reserve (primary ovarian insufficiency), so relative to alternative assays, a test designed to elicit subtleties is typically less important in this population.

Antral Follicle Count (AFC)

MECHANISM

Follicular recruitment is in constant flux during the reproductive years, with less than 0.1% of oogonia present at birth ever making it to ovulation. Fluid surrounding numerous oocytes not selected to be the dominant follicle can be seen sonographically prior to regression. The more follicles visualized within the ovary, the greater the probable ovarian reserve, and AFC has been shown to correlate closely with the primordial follicular pool on histologic analysis. (13, 14). Though it remains for debate as to how much a dwindling follicular pool reflects oocyte quality as well as quantity, women with infertility are more likely to have lower antral follicle counts than those without infertility (15). Similarly, women with low antral follicle counts are much more likely to have cancellation for under response with IVF than those with normal counts (16). However, though low quantity in younger women may reflect fewer oocytes with which blastocysts can form, it does not clearly seem associated with higher rates of aneuploidy or miscarriage (17).

TESTING

Antral follicle count can be measured at any time during the menstrual cycle, as well as when a woman is on hormonal contraceptives or is pregnant. Classically, a woman's AFC is the total number of ovarian follicles measuring between two and nine millimeters, though many studies count follicles up to and including ten millimeters in size (Figure 3).

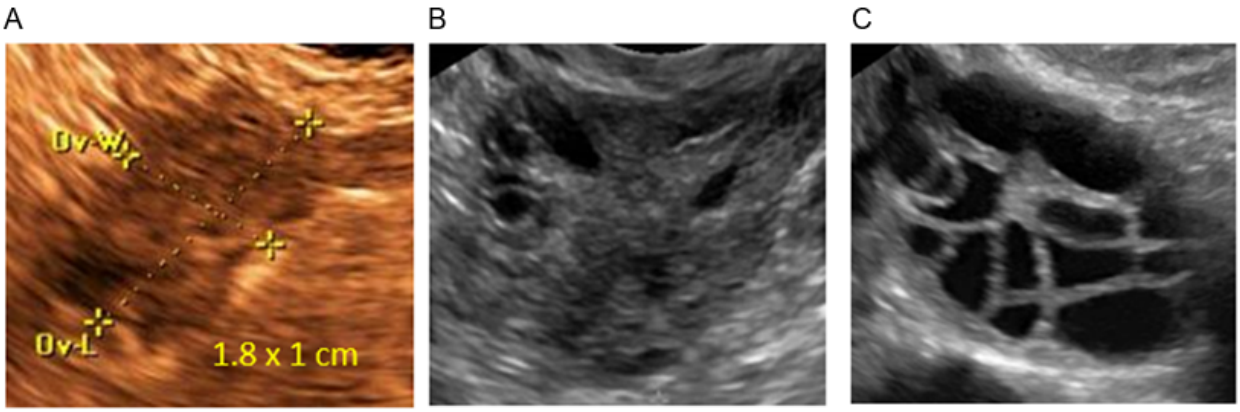

Figure 3. Ovarian sonographic imaging of women in their mid-30's. Figure 3A is from a woman with premature ovarian failure and there are no visualized antral follicles (the sonographically anechoic regions measuring approximately two to nine millimeters within the ovary). Figure 3B is from a woman with tubal factor infertility, and for whom seeing a few follicles within a single plane of the ovary would be normal. Figure 3C is from a woman with polycystic ovarian syndrome. Though her ovary is arguably more multicystic than polycystic (which would typically have follicles concentrated on the periphery of the ovary), she met the criteria for PCOS and her ovary is clearly distinct from those shown in 3A and 3B. Of note, all three ultimately conceived with their own oocytes, so it should be remembered that the absence of visualized antral follicles makes conception far less probable, but not impossible.

Multiple cutoffs are used for what constitutes normal and poor ovarian reserve. Given that antral follicle count varies among cycles, it is reasonable to view the AFC as a continuum, with four total antral follicles reflecting limited reserve, but five antral follicles not being entirely reassuring. Additionally, what constitutes normal is age dependent, where ten total antral follicles may be common for women in their 30's, but not their teens. Though many measures have been used to define polycystic ovarian morphology, the most accepted standard is that used in the Rotterdam criteria of, "12 or more follicles in each ovary measuring 2 to 9 mm in diameter, and/or increased ovarian volume (>10 ml)." This cutoff was chosen, as it was associated with 75% sensitivity and 98% specificity for distinguishing polycystic ovarian morphology (PCOM) from normal ovaries (18). Another frequently used definition comes from Adams, who considered an ovary polycystic if there were ≥ 10 follicles measuring <9 mm (19). Of note, in the development of guidelines for the WHO on PCOS, sonography was deemed preferable to AMH levels from a pragmatic standpoint. (20)

LIMITATIONS

There is debate as to how much moving outside of the early follicular phase or hormonal modulation

such as pregnancy and oral contraceptives will shift the measurement of antral follicle count. Both central and paracrine effects can occur and these are more likely to be meaningful in patients with suboptimal ovarian reserve. However, patients with reassuring ovarian reserve are unlikely to move into a non-reassuring category through these conditions if the ultrasound resolution allows for early antral follicle visualization and measurement.

Patient dependent and observer dependent limitations should also be considered. Patients with elevated BMI (particularly with increased vaginal adiposity) and/or scarring of the pelvis may be more likely to have ovaries with limited resolution for assessment, which could potentially underestimate ovarian reserve. Similarly, large cysts or endometriomas could exert a temporary paracrine effect underestimating reserve. Patients with previous ovarian surgery could also have inclusion cysts appearing similar to antral follicles, but these won't develop with stimulation or have oocytes at follicular aspiration for IVF. For observer dependent limitations, it should be noted that in some multi-center studies where anti-Müllerian hormone (AMH) is found superior to antral follicle count, one can find most institutions having AMH and AFC equally predictive, but one site has an observer where there is a meaningful difference. This has led some to conclude AMH superior to AFC, but failure to properly train observers prior to research is a limitation to study design and may not necessarily reflect true diminished value in utilizing AFC for assessing ovarian reserve. ASRM 2022 Practice Committee Guidelines note, "When performed in an experienced center, AFC is a reasonable alternative to AMH" (29).

Anti-Müllerian Hormone (AMH, Müllerian Inhibiting Substance, MIS)

MECHANISM

AMH is a homodimeric glycopeptide that in reproductive aged women is predominantly granulosa cell derived. The role of systemic AMH is not clear, but at the level of the ovary, it is believed to downregulate FSH mediated folliculogenesis. AMH expression is highest in secondary, preantral, and small antral follicles up until approximately 4 mm in size, and it stops being expressed by granulosa cells when the follicle measures in the 4 to 8 mm range (Figure 4).

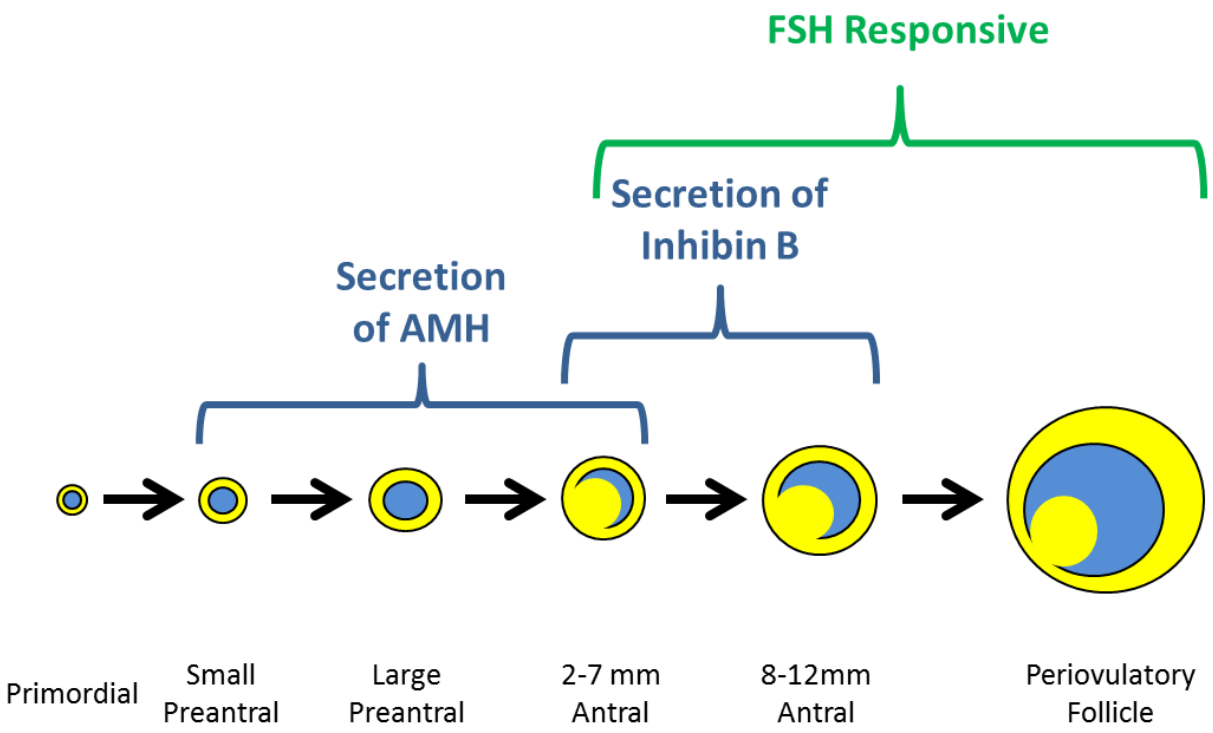

Figure 4. The interplay of follicular development and hormonal secretion and responsiveness.

AMH seems to have a role in selecting the dominant follicle in addition to generally mediating preantral follicular recruitment. AMH levels start undergoing a log-linear decline approximately fifteen years prior to menopause and drop to very low levels approximately five years before menopause (21).

The AMH level associated with diminished ovarian reserve is assay specific and depends on the desired balance of sensitivity and specificity, but is typically below 1 ng/mL. The threshold for menopause is typically lower than the lower detectable limit for many assays, being slightly below 0.1 ng/mL (22). AMH <0.5 ng/mL seems associated with fewer than three follicles available at retrieval, 0.5-1 ng/mL with reduced response, 1-3.5 ng/mL with normal response, and >3.5 ng/mL with overresponse, reflecting greater risk for ovarian hyperstimulation syndrome (23). Normal AMH values often exceed 2 ng/mL at 30, 1.5 ng/mL at 35, and 1 ng/mL at 40 as a quick reference for expected reserve at a given age.

The role of weight loss on AMH levels is open for debate, but should be substratified in women with and without PCOS. (51) Though a lack of association cannot be excluded due to limitations in sample size, there does not seem to be a clear shift in AMH with weight loss for non-PCOS patients. However, for those with elevated AMH from PCOS that has been effectively treated through diet, exercise, and/or bariatric surgery, there is improved fertility, even as AMH lowers to more normal levels. (51) It is unclear which signal transduction pathways drive these lower levels, as one would expect weight loss-associated

shifts in adiponectin, leptin, and insulin to actually increase AMH through recognized mechanisms. Lifestyle may have a broader impact beyond weight on ovarian reserve, as both AMH and AFC are statistically lower in women with lower socioeconomic status (56).

TESTING

AMH levels are measured through immunoassay on a sample obtained through phlebotomy. Values obtained have the distinct advantage of being equally valid at any point in the menstrual cycle. Because AMH is expressed primarily before FSH responsiveness occurs, it is believed that AMH remains a valid assay even when ovarian suppression occurs through smoking, oral contraceptives, GnRH agonists, and pregnancy (24). Though these factors can lead to transient ovarian suppression, they are unlikely to change levels so much as to meaningfully underestimate true reserve. The magnitude of effect through these reversible factors seems to be low, with age-specific AMH percentiles decreasing by 11% with oral contraceptives and 17% with pregnancy (25). Additionally, AMH levels drawn on day seven of the pill free interval seem to closely correlate with levels seen after oral contraceptive discontinuation (26).

A popular misconception is that just because it is valid to assess AMH throughout the
menstrual cycle and under a variety of inhibitory conditions, this should not be mistaken as meaning that AMH levels are static. Though levels of AMH tend to be steady state in perimenopausal patients, for those with higher ovarian reserve, AMH levels fluctuate significantly. This fluctuation, however, is not to the point where a person with robust ovarian reserve is likely to be categorized as having limited reserve (21).

LIMITATIONS

AMH seems to have fewer limitations than most other assays. In fact, it seems to have the advantage that it not only is useful in predicting ovarian response to gonadotropin stimulation, but may even have limited value in predicting pregnancy rates (27). However, like other ovarian reserve assays, it does not appear particularly valuable in predicting viability once pregnancy has already been established. When there is discordance between AFC and AMH levels (e.g., low AFC but normal AMH or vice versa), ovarian response is often a hybrid of the two findings (above those with diminished reserve but less than that of those with normal reserve) (28).

Inhibin B

MECHANISM

Inhibin B is similar to AMH in that it is a glycoprotein secreted by preantral follicles, with levels declining with age. Both inhibin A and B downregulate pituitary FSH secretion. However, Inhibin A levels are not used to predict ovarian reserve because they arise primarily from the dominant follicle rather than an earlier follicular cohort and therefore are less predictive. Inhibin B levels are relatively more useful, but overall remain suboptimally predictive, as they are a late finding for diminished ovarian reserve and typically start falling around four years prior to menopause (21).

TESTING

Inhibin levels are measured by immunoassay after phlebotomy. Inhibin B levels fluctuate over the menstrual cycle, with peaks in the early to mid-follicular phase, as well as during ovulation. Accordingly, inhibin B is typically measured on the third day of the menstrual cycle in ovulatory women. Outside of ovarian reserve testing, in postmenopausal women, where inhibin B levels should be consistently low, a random level is particularly good for following granulosa cell tumors (>89% have elevated inhibin B) and also can be useful for following some epithelial cell ovarian tumors.

LIMITATIONS

In addition to significant variation within the cycle, there is also meaningful variation among cycles. Because of limited sensitivity and specificity, this assay has greater value in those far more likely to have diminished reserve. Some have proposed using inhibin B in combination with other assays, but it is the opinion of the American Society for Reproductive Medicine that “combined ovarian reserve tests models do not consistently improve predictive ability over that of single ovarian reserve tests.” (29).

Ovarian Volume

MECHANISM

Follicles, stroma, and vasculature all contribute to ovarian volume. The percentage that each contributes depends on the individual, her age, underlying gynecologic conditions, and where she is during the menstrual cycle.

TESTING

Typically, ultrasound is used to measure the ovary in all 3 dimensions. These measurements are then applied in the formula for calculating the volume of an ellipse (D1 x D2 x D3 x 0.523). An ovarian volume of >10 cm^3 is considered consistent with PCOS. Although increased ovarian stromal volume distinguishes polycystic ovarian morphology from the multicystic ovary, stromal volume is not routinely measured. Alternative approaches that may improve the effectiveness of ovarian volume include the use of trapezoidal volume (30), 3D ultrasound (31), and color Doppler (32).

LIMITATIONS

Ovarian volume shifts in response to normal physiologic changes (such as the presence of a dominant follicle) and coexisting medical conditions (such as endometriomas). Exogenous hormones can decrease ovarian volume (33), even though ovarian reserve itself has not changed. For these reasons, if evaluating the ovaries by ultrasound, antral follicle count is believed to be a better proxy for ovarian reserve.

APPLICATION OF OVARIAN RESERVE TESTS

Assessing Perimenopausal Status

Classically, ovarian insufficiency and failure have been defined as present when persistent FSH levels >40 μIU/mL are found with at least two radioimmunoassays more than a month apart. No detectable antral follicles in a patient without ovarian suppression is consistent with a perimenopausal state and fewer than two antral follicles has been deemed a more sensitive cutoff (29). The reason to not require the complete absence of follicles is that minimal follicular development is not unusual in

postmenopausal women, as there can be a 14% prevalence and an 8% incidence of simple cysts in a given year (34). Similarly, though an undetectable AMH level would be consistent with menopause, in women with primary ovarian insufficiency, approximately a quarter of them will have below normal but detectable AMH levels and a sixth will have normal AMH levels (35). Though women with advancing age will have higher FSH levels, it remains unclear if women with elevated FSH earlier in their reproductive life will go through menopause earlier (36). Finally, it should be remembered that confirmation of primary ovarian insufficiency does not automatically mean completion of testing, as fragile X carrier screening and other evaluations may be appropriate.

Evaluating Ovarian Reserve for Fertility in Ovulatory Patients

For ovarian reserve testing prior to fertility therapy, there is more data on FSH than other measures. Generally, women of the same age with higher FSH levels seem to have lower fecundability (37). However, younger women with elevated FSH levels often have much better fecundability than older women with comparably elevated FSH (38) and age can be a better predictor of outcome than FSH (39). Though differences in pregnancy rates can be shown between those with high and low FSH, the assay in general has suboptimal sensitivity for both ovarian response and pregnancy rates, as reflected by receiver-operator curves (9).

A rarely addressed caveat is that though it is true that multiple studies are showing AMH and AFC to have a better balance of sensitivity and specificity than FSH, meta-analyses regarding the predictive value of FSH run the risk of being biased towards the null. The reason is that the earliest ovarian reserve testing research (using FSH) was done at a time when IVF success rates were lower. This caveat won't apply to modern studies were FSH is directly compared with AMH or AFC, but one should account for temporal bias in meta-analyses if studies from the 1990s are included. One should also note, that 20% of the time there will be AMH and FSH discordance, particularly in older women where this can be as high as 33% (55).

Anti-Müllerian hormone levels and antral follicle count seem to be emerging as the best approaches to procreative testing. A survey of 796 centers noted 51% thought AMH the best measure for ovarian reserve, while 40% selected AFC, though ultimately 80% felt age was the best predictor for pregnancy (50). After accounting for age, AFC and AMH seem highly accurate in predicting poor response with IVF, while FSH does so only moderately (40). Not only are these measures commonly used for predicting under response, but they can also be used to predict hyperstimulation (41). Ovarian reserve assessment for reproductive purposes is fraught with controversy because different practitioners prefer different balances of sensitivity and specificity. At the minimum it should be recognized that this type of testing is meant to be screening for women who are more likely to have a poor response to ovarian stimulation, and findings are not necessarily diagnostic of ovarian failure or the degree of risk for premature menopause. However, results consistent with perimenopausal findings should be confirmed and appropriate counseling given. As stated by ASRM, "Extremely low AMH levels should not be used to refuse treatment in IVF" (29). The American College of Obstetricians and Gynecologists (ACOG) draws similar conclusions (42).

Since combined tests do not consistently improve the ability to predict ovarian response, many clinicians are simply using either AMH or AFC in the context of the patient's age and reserve additional testing for atypical clinical pictures or to confirm significant ovarian insufficiency. In spite of this ASRM recommended approach, some argue that combined testing improves sensitivity in detecting suboptimal ovarian reserve. Whether or not the literature ultimately demonstrates this, a way of side-stepping this debate is by noting that combined testing is unlikely to be cost-effective. The reason is that if additional testing is unlikely to change management (especially when the vast majority of patients have normal results), it is very hard to show cost-effectiveness when doubling or tripling costs without clear benefit. Accordingly, if using combined testing, it should be selective rather than universal.

An additional note on ovarian reserve tests in procreation relates to their limitations. Though some appear better than others in predicting ovarian response to stimulation, most are limited at best in predicting pregnancy, and this predictive value is highly dependent on patient demographics within a study. This is not inherently a flaw in the assays; rather, infertility is often multifactorial, so when ovarian reserve testing is a subset of factors, this tends to bias its relevance towards the null. Studies showing an ovarian reserve test to be predictive of pregnancy in general tend to have older populations. (It has been argued as to whether this constitutes enrollment bias or limits external validity; however, it is reasonable to find a test having greater value when applied to a population at risk.)

Another limitation is that abnormal ovarian reserve testing does not always increase miscarriage rates (43), despite an association between abnormal testing and blastocyst aneuploidy (1). Studies show roughly a 25% increase in the probability per embryo of being aneuploid in the setting of diminished ovarian reserve (52, 53). This may underestimate the magnitude of effect, as many aneuploid embryos may not survive to biopsy and evaluation. This is why one can see the bottom quartile having higher rates of all embryos being aneuploid (19.3% vs 10.3%) and of having only one embryo to biopsy (31% vs 11%) (52). Further muddying the waters, for a particular age, though women with diminished ovarian reserve may have quantity issues without this always translating to quality (euploidy), looking at the 5%-10% of patients with lowest reserve may give a different answer than looking at the bottom quartile. The better performers among the population with the lowest reserve may bias the data towards the null. This is why when comparing the cited Jaswa (53) and Fouks (52) articles, one sees a higher rate of aneuploidy in the Jaswa study (71% DOR vs 55% controls) relative to Fouks (50% bottom quartile vs 60% middle quartiles), as the Jaswa DOR group arguably had more pronounced diminished ovarian reserve. A study by Morin focused on those with the bottom 10% for ovarian reserve and noted a decline in quantity didn't shift quality (54). However, the bottom decile had triple the rate of no usable blastocysts (17% vs 5.3%) and lower live birth rates (41.2% vs 53.1% per cycle start), where arguably successful pregnancy is the ultimate metric of quality. The Morin DOR population was meaningfully younger than the Jaswa population, so DOR associated with aging is likely to be more concerning than comparable ovarian reserve in younger women (53, 54).

To further complicate the quantity vs. quality debate, there can be heterogeneity, where some women with DOR have far lower quality than others. Finally,

follicular quantity is more valuable in identifying ovarian factors for subfertility patients but does not seem predictive of outcomes in patients who have suboptimal reserve, but have never tried to conceive (44).

Evaluating Amenorrhea in the Post-menarche, Pre-menopausal Patient

Numerous conditions can cause amenorrhea in reproductive aged women. Testing falls into two categories: diagnosing etiology and reassuring the patient that she is not in ovarian failure. For a more comprehensive discussion of how to evaluate etiology, please see the Endotext section on the Endocrinology of Female Reproduction. In general, and in addition to remembering to exclude pregnancy as a cause for amenorrhea, ASRM Practice Committee guidelines recommend FSH, TSH, and prolactin levels in addition to the usual history and physical exam (45) (An important contextual caveat is that AMH and AFC were not as well established when these guidelines came out in 2008). Though increasingly AFC is used in place of FSH, especially for evaluating hyperandrogenic women since AFC is part of the Rotterdam criteria (46), FSH still has a role in differentiating PCOS and forms of functional hypothalamic amenorrhea (47). See figure 5 for interpretation of test results.

Parameter	FHA (stress, exercise)	FHA (anorexia)	FHA + PCOS	PCOS	POF/POI	Obesity
LH	Low to normal	Low to undetectable	Normal to low	Normal to high	Normal to high	Normal to low
FSH	Low	Normal to low	Normal to low	Normal to low	High	Normal to low
E2	Low	Lowest	Normal to low	Normal to low	Normal to low	Normal to low
Testosterone	Low to normal	Low	Normal to high	High	Normal to low	Normal to high
DHEAS	Normal to low	Low	Normal	Normal to high	Normal to low	Normal to low
AMH (MIS)	Normal	Normal	Normal to high	Normal to high	Low	Normal to low

POF, Premature ovarian failure; POI, premature ovarian insufficiency; E2, estradiol; DHEAS, dehydroepiandrosterone sulfate; AMH (MIS), anti-Mullerian hormone (Mullerian-inhibiting substance).

Figure 5. Laboratory parameters in the setting of amenorrhea (43). Reproduced with permission of the author.

Regarding reassuring the patient, for gynecologists a transvaginal ultrasound to assess antral follicle count is relatively easy to perform, can often be performed promptly at the initial office visit, and can have a reassuring tangibility to patients when antral follicles are identified and their importance is explained. When sonographic evaluation of ovarian reserve is less available, a normal AMH level should be reassuring. Additionally, for patients who have been placed on oral contraceptives or other hormonal therapy without a diagnosis of etiology for amenorrhea, AFC and AMH levels may be lowered, but are still likely to remain within the normal range if the patient truly has normal reserve. Figure 6 provides information on the relative strengths of assays.

Clinical question	AFC	AMH	FSH	E2	CCCT	Inhibin B	Ovarian Volume
Evaluating ovarian reserve for fertility	+++	+++	+	-	+	-	-
Confirming postmenopausal status	+++	+++	+++	+	+	++	+
Distinguishing PCOS from hypogonadotropic hypogonadism	++	++	+++ *	+	-	++	+

Figure 6. Relative strengths of assays for determining ovarian reserve (balancing positive and negative predictive values, typical costs, and available alternatives). AFC= antral follicle count, AMH= anti-mullerian hormone, CCCT= clomiphene citrate challenge test. * When combined with LH

LOOKING TO THE FUTURE OF OVARIAN RESERVE TESTING

Advancing technologies and improving cost has made direct to consumer (DTC) fertility testing a reality. DTC testing enhances patient access care, but can do so at the expense of oversight, insight, and broader perspective. Moreover, though abnormal results may encourage a patient to seek out a physician for counseling, normal results may provide false security for fertility, which is multifactorial and more than ovarian reserve. Additionally, ovarian reserve testing in patients with untested fertility may provide limited predictive value for fecundability and ultimately fecundity (57). Ultimately, how do physicians interpret results in patients who may not warrant assessment? If DTC ovarian reserve testing is typically performed by fingerstick blood sampling, relative to a larger, better-preserved sample through traditional phlebotomy, sample accuracy at times can be suboptimal. Additionally, lack of assay standardization among labs can also hinder counseling patients on their results (58). If more information is available, the true questions are if we can use and trust the results?

When (and if) ovarian reserve testing is predictive, therapy can be focally applied early. Current testing methods are often more reliable after DOR has already occurred and better predict response to therapies than pregnancy itself. The known causative association between Fragile X premutation in women and risk of decreased ovarian reserve leads to interest in genetic causes of decreased ovarian reserve. Studies have linked target genes and epigenetic changes to patients with diagnosed DOR (59, 60). If testing Fragile X, as well as testing for blepharophimosis, ptosis, and epicanthus inversus (BPES), can anticipate DOR, the challenge is not just in making the most of ovarian reserve where possible, but there is also a role for potentially reducing transmissible morbidity.

CONCLUDING REMARKS

It has been said that a Rolex keeps time well, but makes for a lousy hammer. All ovarian reserve tests are merely tools and their value relates to the task to which they are applied. Even as we see increased use of AFC and AMH (48), we have to remember that ideal testing is "systematic, expeditious, and cost-effective" (49). In other words, when evaluating ovarian reserve, one should account for not only the symptoms and probable diagnosis, but also the turnaround time for results, and how to maximize value in testing. These latter two factors vary by site, so clinicians will have to find the right balance for their practice. Finally, one of the most important and cost-effective predictors is age (see Figure 1). In the procreative setting, after age is combined with another ovarian reserve test, the marginal benefit from further assays tends to be less (8). Accordingly, and with the exception of premature ovarian failure where independent confirmation is appropriate (due to discordance between age and the assay), until further studies justify effectiveness and cost-effectiveness, simultaneously using multiple ovarian reserve tests should be for selected patients rather than universal.

REFERENCES

1. Katz-Jaffe MG, Surrey ES, Minjarez DA, Gustofson RL, Stevens JM, Schoolcraft WB. Association of Abnormal Ovarian Reserve Parameters With a Higher Incidence of Aneuploid Blastocysts. Obstet Gynecol. 2013; 121(1): 71-77.

2. "National Summary Report." Society for Assisted Reproductive Technology. https://www.sartcorsonline.com/rptCSR_PublicMultYear.aspx?reportingYear=2020. Accessed December 4, 2022.

3. Sherman BM, West JH, Korenman SG. The menopausal transition: analysis of LH, FSH, estradiol, and progesterone concentrations during menstrual cycles of older women. J Clin Endocrinol Metab. 1976; 42: 629–36.

4. Rose MP, Gaines Das RE, Balen AH. Definition and measurement of follicle stimulating hormone. Endocr Rev. 2000 Feb;21(1):5-22.

5. Esposito MA, Coutifaris C, Barnhart KT. A moderately elevated day 3 FSH concentration has limited predictive value, especially in younger women. Hum Reprod. 2002;17:118-23.

6. Harlow SD, Gass M, Hall JE, Lobo R, Maki P, Rebar RW, Sherman S, Sluss PM, de Villiers TJ; STRAW + 10 Collaborative Group. Executive summary of the Stages of Reproductive Aging Workshop + 10: addressing the unfinished agenda of staging reproductive aging. J Clin Endocrinol Metab. 2012 Apr;97(4):1159-68.

7. Backer LC, Rubin CS, Marcus M, Kieszak SM, Schober SE. Serum Follicle-Stimulating Hormone and Luteinizing Hormone Levels in Women Aged 35-60 in the U.S. Population: The Third National Health and Nutrition Examination Survey (NHANES III,1988-1994). Menopause. 1999; 6(1): 29-35.

8. Broekmans FJ, Kwee J, Hendriks DJ, Mol BW, Lambalk CB. A systematic review of tests predicting ovarian reserve and IVF outcome. Hum Reprod Update. 2006 Nov-Dec;12(6):685-718.

9. Broer SL, van Disseldorp J, Broeze KA, Dolleman M, Opmeer BC, Bossuy P, Eijkemans MJC, Mol BJ, Broekmans FJM. Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: an individual patient data approach. Hum Reprod Update. 2013; 19(1): 26-36.

10. Hall JE, Schoenfeld DA, Martin KA, Crowley WF Jr.. Hypothalamic gonadotropin-releasing hormone secretion and follicle-stimulating hormone dynamics during the luteal-follicular transition. J Clin Endocrinol Metab. 1992; 74 (3): 600-7.

11. Scott RT, Leonardi MR, Hofmann GE, Illions EH, Neal GS, Navot D. A prospective evaluation of clomiphene citrate challenge test screening of the general infertility population. Obstet Gynecol. 1993 Oct;82(4 Pt 1):539-44.

12. Scott RT Jr, Illions EH, Kost ER, Dellinger C, Hofmann GE, Navot D. Evaluation of the significance of the estradiol response during the clomiphene citrate challenge test. Fertil Steril. 1993 Aug;60(2):242-6.

13. Gougeon A, Chainy GB. Morphometric studies of small follicles in ovaries of women at different ages. J Reprod Fertil. 1987; 81: 433–42.

14. Faddy MJ, Gosden RG, Gougeon A, Richardson SJ, Nelson JF. Accelerated disappearance of ovarian follicles in mid-life: implications for forecasting menopause. Hum Reprod. 1992; 7: 1342–46.

15. Rosen, MP, Johnstone, E, Addauan-Andersen, C, Cedars, MI. A lower antral follicle count is associated with infertility. Fertil Steril. 2011; 95(6): 1950-54.

16. Frattarelli JL, Lauria-Costab DF, Miller BT, Bergh PA, Scott RT. Basal antral follicle number and mean ovarian diameter predict cycle cancellation and ovarian responsiveness in assisted reproductive technology cycles. Fertil Steril. 2000; 74(3): 512-17.

17. Bishop LA, Richter KS, Patounakis G, Adraiani L, Moon K, Devine K. Diminished ovarian reserve as measured by means of baseline follicle-stimulating hormone and antral follicle count is not associated with pregnancy loss in younger in vitro fertilization patients. Fertil Steril. 2017; 108(6):980–987.

18. Jonard S, Robert Y, Cortet-Rudelli C, Pigny P, Decanter C, Dewailly D. Ultrasound examination of polycystic ovaries: is it worth counting the follicles? Hum Reprod. 2003; 18: 598–603.

19. Adams J, Franks S, Polson DW, Mason HD, Abdulwahid N, Tucker M, Morris DV, Price J, Jacobs HS. Multifollicular ovaries: clinical and endocrine features and response to pulsatile gonadotrophin releasing hormone. Lancet. 1985 Dec; 2(8469-70): 1375-79.

20. Balen AH, Morley LC, Misso M, Franks S, Legro RS, Wijeyaratne CN, Stener-Victorin E, Fauser BCJM, Norman RJ, Teede H. The management of anovulatory infertility in women with polycystic ovary syndrome: an analysis of the evidence to support the development of global WHO guidance. Hum Reprod Update 2016 Nov;22(6):687-708

21. Sowers MR, Eyvazzadeh AD, McConnell D, Yosef M, Jannausch ML, Zhang D, Harlow S, Randolph Jr JF. Anti-mullerian hormone and inhibin B in the definition of ovarian aging and the menopause transition. J Clin Endocrinol Metab. 2008; 93: 3478–83.

22. Dolleman M, Verschuren WMM, Eijkemans MJC, Dolle MET, Jansen EHJM, Broekmans FJM, van der Schouw YT. Reproductive and lifestyle determinants of anti-Müllerian hormone in a large population-based study. J Clin Endocrinol Metab. 2013; 98 (5): 2106-2115.

23. Toner JP, Seifer DB. Why we may abandon follicle-stimulating hormone testing: a sea change in determining ovarian reserve using antimullerian hormone. Fertil Steril. 2013 Jun; 99(7): 1825-30.

24. Ledger WL. Clinical utility of measurement of anti-mullerian hormone in reproductive endocrinology. J Clin Endocrinol Metab. 2010 Dec;95(12):5144-54.

25. Dolleman M, Faddy MH, van Disseldorp J, van der Schouw YT, Messow CM, Leader B, Peeters PHM, McConnachie A, Nelson SM, Broekmans FJM. The relationship between anti-Müllerian hormone in women receiving fertility assessments and age at menopause in subfertile women: evidence from large population studies. J Clin Endocrinol Metab. 2013; 98 (5): 1946-53.

26. van den Berg MH, van Dulmen-den Broeder E, Overbeek A, Twisk JWR, Schats R, van Leeuwen FE, Kaspers GJ, Lambalk CB. Comparison of ovarian function markers in users of hormonal contraceptives during the hormone-free interval and subsequent natural early follicular phases. Hum Reprod. 2010; 25: 1520-7.

27. Tal R, Seifer DB, Wantman E, Baker V, Tal O. Antimüllerian hormone as a predictor of live birth following assisted reproduction: an analysis of 85,062 fresh and thawed cycles from the Society for Assisted Reproductive Technology Clinic Outcome Reporting System database for 2012-2013. Fertil Steril. 2018; 109(2):258–265.

28. Li HW, Lee VC, Lau EY, Yeung WS, Ho PC, and Ng EH. Ovarian response and cumulative live birth rate of women undergoing in-vitro fertilization who had discordant anti-Müllerian hormone and antral follicle count measurements: a retrospective study. PLoS One. 2014; 9(10): e108493

29. The Practice Committee of the American Society for Reproductive Medicine. Testing and interpreting measures of ovarian reserve: a committee opinion. Fertil Steril. 2020; 114(6): 1151-7.

30. Giacobbe M, Mendes Pinto-Neto A, Simoes Costa-Paiva LH, Martinez EZ. The usefulness of ovarian volume, antral follicle count and age as predictors of menopausal status. Climacteric 2004; 7: 255-60.

31. Merce LT, Gomez B, Engels V, Bau S, Bajo JM. Intraobserver and interobserver reproducibility of ovarian volume, antral follicle count, and vascularity indices obtained with transvaginal 3-dimensional ultrasonography, power Doppler angiography, and the virtual organ computer-aided analysis imaging program. J Ultrasound Med. 2005; 24: 1279-87.

32. Jarvela IY, Sladkevicius P, Tekay AH, Campbell S, Nargund G. Intraobserver and interobserver variability of ovarian volume, gray-scale and color flow indices obtained using transvaginal threedimensional power Doppler ultrasonography. Ultrasound Obstet Gynecol. 2003; 21: 277-82.

33. Christensen JT, Boldsen J, Westergaard JG. Ovarian volume in gynecoglogically healthy women using no contraception, or using IUD, or oral contraception. Acta Obstet Gynecol Scand. 1997; 76: 784-89.

34. Greenlee RT, Kessel B, Williams CR, Riley TL, Ragard LR, Hartge P, Buys SS, Partridge EE, Reding DJ. Prevalence, incidence and natural history of simple ovarian cysts among women over age 55 in a large cancer screening trial. Am J Obstet Gynecol. 2010 April; 202(4): 373.e1–9.

35. Méduri G, Massin N, Guibourdenche J, Bachelot A, Fiori O, Kuttenn F, Misrahi M, Touraine P. Serum anti-Müllerian hormone expression in women with premature ovarian failure. Hum Reprod. 2007 Jan;22(1):117-23.

36. Soules MR, Sherman S, Parrott E, Rebar R, Santoro N, Utian W, Woods N. Executive summary: Stages of Reproductive Aging Workshop (STRAW). Fertil Steril. 2001; 76: 874-8.

37. Caroppo E, Matteo M, Schonauer LM, Vizziello G, Pasquadibisceglie A, Vitti A, D'Amato G. Basal FSH concentration as a predictor of IVF outcome in older women undergoing stimulation with GnRH antagonist. Reprod Biomed Online. 2006; 13: 815-20.

38. Abdalla H, Thum MY. An elevated basal FSH reflects a quantitative rather than qualitative decline of the ovarian reserve. Hum Reprod. 2004; 19: 893-98.

39. Chuang CC, Chen CD, Chao KH, Chen Su, Ho HN, Yang YS. Age is a better predictor of pregnancy potential than basal follicle stimulating hormone levels in women undergoing in vitro fertilization. Fertil Steril. 2003; 79: 63-68.

40. Broer SL, van Disseldorp J, Broeze KA, Dolleman M, Opmeer BC, Bossuyt P, Eijkemans MFC, Mol BWJ, Broekmans FJM. Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: an individual patient data approach. Hum Reprod Update 2013; 19(1): 26-36.

41. Broer SL, Dolleman M, Opmeer BC, Fauser BC, Mol BW, Broekmans FJM. AMH and AFC as predictors of excessive response in controlled ovarian hyperstimulation: a meta-analysis. Hum Reprod Update. 2011; 17(1): 46-54.

42. ACOG Practice Committee. Committee Opinion No 618: Ovarian Reserve Testing. Obstet Gynecol 2015; 125(1): 268-73.

43. Haadsma ML, Groen H, Fidler V, Seinen LHM, Broekmans FJM, Heineman MJ, Hoek A. The predictive value of ovarian reserve tests for miscarriage in a population of subfertile ovulatory women. Hum Reprod. 2009; 24(3): 546-552.

44. Steiner AZ, Pritchard D, Stanczyk FZ, Kesner JS, Meadows JW, Herring AH, Baird DD. Association between biomarkers of ovarian reserve and infertility among older women of reproductive age. JAMA 2017; 318(14): 1367-1376.

45. The Practice Committee of the American Society for Reproductive Medicine. Current evaluation of amenorrhea. Fertil Steril. 2008; 90: S219-25.

46. The Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. Fertil Steril. 2004; 81(1): 19-25.

47. Santoro, N. Update in hyper- and hyogonadotropic amenorrhea. J Clin Endocrinol Metab 2011; 96(11): 3281-88.

48. Grisendi V, Mastellari E, La Marca A. Ovarian Reserve Markers to Identify Poor Responders in the Context of Poseidon Classification. Front Endocrinol (Lausanne). 2019 May 8;10:281

49. The Practice Committee of the American Society for Reproductive Medicine. Optimal evaluation of the infertile female. Fertil Steril 2006; 86(4): S264-67.

50. Tobler KJ, Shoham G, Christianson MS, Zhao Y, Leong M, Shoham Z. Use of anti-Müllerian hormone for testing ovarian reserve: a survey

of 796 infertility clinics worldwide. J Assist Reprod Genet. 2015;32(10):1441–1448.

51. Kloos J, Coyne K, Weinerman R. The relationship between anti-Müllerian hormone, body mass index and weight loss: A review of the literature. Clin Obesity. 2022; 12(6): e12559.

52. Fouks Y, Penzias A, Neuhausser W, Vaughan D, Sakkas D. A diagnosis of diminished ovarian reserve does not impact embryo aneuploidy or live birth rates compared to patients with normal ovarian reserve. Fertil Steril. 2022; 118(3): 504-12.

53. Jaswa EG, McCulloch CE, Simbulan R, Cedars MI, Rosen MP. Diminished ovarian reserve is associated with reduced euploid rates via preimplantation genetic testing for aneuploidy independently from age: evidence for concomitant reduction in oocyte quality with quantity. Fertil Steril. 2021; 115(4): 966-73.

54. Morin SJ, Patounakis G, Juneau CR, Neal SA, Scott Jr RT, Seli EJ. Diminished ovarian reserve and poor response to stimulation in patients< 38 years old: a quantitative but not qualitative reduction in performance. Human Reproduction. 2018; 33(8): 1489-98.

55. Leader B, Hegde A, Baca Q, Stone K, Lannon B, Seifer DB, Broekmans F, Baker VL. High frequency of discordance between antimüllerian hormone and follicle-stimulating hormone levels in serum from estradiol-confirmed days 2 to 4 of the menstrual cycle from 5,354 women in US fertility centers. Fertil Steril. 2012 Oct 1;98(4):1037-42.

56. Barut MU, Agacayak E, Bozkurt M, Aksu T, Gul T. There is a positive correlation between socioeconomic status and ovarian reserve in women of reproductive age. Medical science monitor: Int Med J Experimental and Clin Research. 2016; 22: 4386-92.

57. Penzias A, Azziz R, Bendikson K, et al. Testing and interpreting measures of ovarian reserve: a committee opinion. Fertil Steril. 2020; 114(6):1151-1157.

58. Kyweluk M, Feinberg EC. Direct-to-consumer (DTC) ovarian reserve testing benefits the company, not the consumer. Fertility and Sterility Dialog [Internet]. 2020 Feb 17. Available from: https://www.fertstertdialog.com/posts/59741-kyweluk-consider-this

59. Moiseeva AV, Kudryavtseva VA, Nikolenko VN, et al. Genetic determination of the ovarian reserve: a literature review. J Ovarian Rev 14, 02 (2021).

60. Olsen KW, Castillo-Fernandez J, Chan AC, et al. Identification of unique epigenetic profile in women with diminished ovarian reserve. Fertil Steril. 2021 Mar; 115(3): 732-41.