Testing and interpreting measures of ovarian reserve: a committee opinion

Practice Committee of the American Society for Reproductive Medicine

American Society for Reproductive Medicine, Birmingham, Alabama

Ovarian reserve is defined as the number of oocytes remaining in the ovary, or oocyte quantity (oocyte number). Markers of ovarian reserve include hormone levels and sonographically measured features of the ovaries. These markers can be useful as predictors of oocyte yield following controlled ovarian stimulation and oocyte retrieval. However, they are poor predictors of reproductive potential independently from age. This document replaces the document of the same name last published in 2012 (Fertil Steril 2012;98:1407–15).

El resumen está disponible en Español al final del artículo.

Discuss: You can discuss this article with its authors and other readers at https://www.fertstertdialog.com/posts/31403

The process of reproductive aging has traditionally centered on the principle that human oocytes peak in number during fetal life, undergo ovulation or atresia thereafter, and do not regenerate. Reproductive and ovarian senescence occurs with depletion of the number of oocytes, or ovarian reserve. Women in later stages of reproductive aging (perimenopause and menopause), when menses become irregular, have lower ovarian reserve than women in earlier reproductive stages, when menstrual cycles are normally regular (1). Ovarian reserve can be indirectly measured with the use of hormone levels or ultrasound imaging of the ovaries. This document reviews the evidence relating to the clinical utility and predictive value of ovarian reserve testing as predictors of reproductive potential.

WHAT IS OVARIAN RESERVE?

Ovarian reserve is defined as the number of oocytes remaining in the ovary, or oocyte quantity (oocyte number). Ovarian reserve, or oocyte quantity (oocyte number), is different from oocyte quality, which relates to the potential of a fertilized oocyte to result in a live-born infant. Female infants are born with ~500,000 to 1 million oocytes, follicular atresia and ovulation result in a slow depletion of oocyte number over time, and menopause subsequently ensues. Ovarian reserve correlates inversely with age, but there is considerable variation in ovarian reserve among women of the same chronologic age (2).

WHAT ARE MEASURES OF OVARIAN RESERVE?

Ovarian reserve tests include both biochemical tests and ultrasound imaging of the ovaries. Biochemical tests of ovarian reserve can be divided further into early-follicular-phase measurements of FSH, E2, or inhibin B; measurement of cycle-day-independent antimullerian hormone (AMH); and provocative tests, such as the clomiphene citrate challenge test (CCCT). Biochemical measures of ovarian reserve are intended to directly or indirectly measure the oocyte or follicular pool.

Inhibin B and AMH are glycoprotein hormones produced by small ovarian follicles and are therefore direct measures of the follicular pool. Whereas AMH is primarily secreted by primary, preantral, and early antral follicles, inhibin B is secreted primarily by preantral follicles. As the number of ovarian follicles declines with age, both AMH and early-follicular-phase inhibin B concentrations decline. Decreased inhibin B secretion lowers the level of central negative feedback, resulting in increased pituitary FSH secretion and in higher late-luteal and early-follicular FSH concentrations (an indirect measure). In turn, the earlier increase in FSH levels stimulates an earlier onset of new follicular growth and an increase in E2 concentrations, ultimately decreasing the length of the follicular phase and the overall cycle. These hormone tests are therefore used as markers of ovarian reserve.

Basal Serum FSH and E2

Basal serum FSH concentrations are elevated on day 2, 3, or 4 of the menstrual cycle in women with diminished ovarian reserve (DOR). Elevated basal serum FSH is a specific, but not sensitive, test for DOR (3). However, FSH levels have significant inter- and intra-cycle variability that limits the reliability of a single measurement. High
intercycle variability suggests more advanced DOR (4, 5). The overall correlation among different FSH assays is high, but absolute values can differ from one another (6). Basal E2 alone should not be used to screen for DOR. The test has value only as an aid to the correct interpretation of a normal basal serum FSH value, as follows. An early rise in serum E2 concentrations is a classic characteristic of reproductive aging and can lower an otherwise elevated basal FSH level into the normal range, thereby causing a misinterpretation of the test. When the basal FSH concentration is normal but the E2 level is elevated (>60–80 pg/mL), this may indicate ovarian dysfunction attributable to DOR (7).

Clomiphene Citrate Challenge Test

The CCCT involves measurements of serum FSH before (cycle day 3) and after (cycle day 10) treatment with clomiphene citrate (100 mg daily, cycle days 5–9). Whereas rising inhibin B and E2 levels derived from a growing cohort of ovarian follicles will suppress FSH in women with responsive ovaries, the smaller follicular cohorts that can be recruited in women with DOR will generate less inhibin B and E2, resulting in decreased negative feedback inhibition of FSH secretion and higher stimulated FSH concentrations. An elevated FSH concentration after clomiphene stimulation therefore suggests DOR. Provocative testing with the use of the CCCT may increase sensitivity for DOR (8). However, compared with basal FSH and ultrasonographically determined antral follicle count (AFC), the clomiphene-stimulated day-10 FSH level is not superior to nondonyclic tests for predicting poor ovarian response or pregnancy after in vitro fertilization (IVF) (8–10) or unassisted conception in women with infertility (11). For this reason, this test should be abandoned.

Antimüllerian Hormone

Serum concentrations of AMH, produced by granulosa cells of early follicles, are gonadotropin independent and, therefore, remain relatively consistent within and between menstrual cycles in both normal, young, ovulating women and women with infertility (12–15). AMH levels may be decreased in women currently using hormonal contraceptives and, therefore, should be interpreted with caution in those patients (16). AMH is a more sensitive measure of ovarian reserve than FSH and tends to decline before FSH rises (17). For this reason, AMH has largely replaced basal FSH and E2 level testing as a biomarker of ovarian reserve. Basal FSH and E2 levels may provide additional information in women with very low AMH levels.

Antral Follicle Count and Ovarian Volume

Ultrasonographic measures of ovarian reserve include AFC and ovarian volume. Ovarian volume declines with age and is therefore a potential indicator of ovarian reserve; however, this is uncommonly used for clinical prediction given the high inter- and intracycle variability and general lack of sensitivity (4). AFC is the sum of the number of antral follicles in both ovaries as observed with transvaginal ultrasonography during the early-follicular phase. Most studies have defined antral follicles as those measuring 2–10 mm in mean diameter in the greatest two-dimensional plane across the ovary. AFC has low intercycle variability and high interobserver reliability in experienced centers (18–22).

SELECTING A MARKER OF OVARIAN RESERVE

AMH levels appear to be a more sensitive marker of ovarian reserve compared with early-follicular-phase hormone levels. AMH tends to decline before FSH rises (17); thus, high FSH is a specific marker for DOR but fails to detect a more subtle decline in ovarian reserve. AMH testing is simpler to administer. AFC and AMH have been shown in multiple studies to be equivalent (23). When performed in an experienced center, AFC is a reasonable alternative to AMH.

DO OVARIAN RESERVE TESTS PREDICT REPRODUCTIVE POTENTIAL AMONG WOMEN WITH UNPROVEN FERTILITY?

Ovarian reserve declines with age, as do fertility rates. This has led to the presumption that ovarian reserve should predict reproductive potential; however, this has not been borne out in studies so far. A number of studies have been conducted to assess the value of markers of ovarian reserve as markers of current reproductive potential. In prospective cohort studies, markers of ovarian reserve were poor predictors of reproductive potential as measured by fecundability (the probability of conceiving in a given menstrual cycle), cumulative probability of pregnancy, or incidence of infertility.

The EAGER (Effects of Aspirin in Gestation and Reproduction) trial enrolled women aged 18–40 years with a history of one or more pregnancy losses (N = 1,202). Women with low AMH levels (<1 ng/mL) had similar cumulative pregnancy rates as women with normal values (1.0–3.5 ng/mL) (24). The Time to Conceive study, including a prospective cohort of 750 women aged 30–44 years without known history of or risk factors for infertility, found that women with low AMH levels (<0.7 ng/mL) or high FSH values (>10 IU/L) had similar cumulative pregnancy rates after 6 and 12 cycles of attempting pregnancy compared with women with normal levels (25). Similarly, markers of ovarian reserve do not predict the probability of conceiving after unmedicated donor-sperm insemination cycles (26).

DO OVARIAN RESERVE TESTS PREDICT REPRODUCTIVE POTENTIAL AMONG WOMEN WITH INFERTILITY?

Based on the limited data available, it appears that results of ovarian reserve tests are no more useful than age alone in predicting unassisted pregnancy in women with infertility, nor do they offer clinically meaningful improvements over already established pregnancy prediction models, such as the Hunault model. The Hunault model incorporates female age, duration of subfertility, previous pregnancy, semen analysis results, and referral status to direct infertile patients into treatment or expectant management depending on their chances for spontaneous conception over the ensuing 12 months (27). Two large prospective cohort studies found...
that while certain ovarian reserve tests (FSH and AFC) did predict unassisted pregnancy, they failed to substantially change patient management compared with the Hunault model alone (11, 28).

**DO MARKERS OF OVARIAN RESERVE IN INFERTILE WOMEN PREDICT TREATMENT SUCCESS AFTER OVARIAN STIMULATION AND INTRAUTERINE INSEMINATION?**

Women with unexplained infertility are commonly treated with ovarian stimulation (OS) using oral agents or gonadotropins combined with intrauterine insemination (IUI). The goal of OS is to achieve multifollicular development to increase the odds of conception. Most studies examining the value of markers of ovarian reserve to predict success following treatment with OS/IUI have been retrospective analyses of clinical practice. Findings from these studies have been inconclusive, with some showing that DOR as detected by these markers is associated with a reduction in the probability of pregnancy after OS/IUI and others showing no association (29–31). Recently, a secondary analysis of the AMIGOS (Assessment of Multiple Intrauterine Gestations from Ovarian Stimulation) trial, a randomized clinical trial that enrolled 900 women with unexplained infertility and treated them with OS/IUI, found that AMH did not predict pregnancy rates after OS/IUI (32). In summary, the best evidence to date suggests that markers of ovarian reserve do not predict the likelihood of success following OS/IUI; however, further research is needed.

**DO OVARIAN RESERVE TESTS PREDICT OOCYTE YIELD AFTER CONTROLLED OVARIAN HYPERSTIMULATION FOR IVF?**

The ability of AMH and AFC to predict oocyte yield as well as poor and excessive response to gonadotropin stimulation in assisted reproductive technologies has been well demonstrated in numerous studies (33–35). Often, AMH and AFC are examined as linear variables in dose-response models. While both continue to show excellent promise to predict ovarian response in IVF, data do not support the use of a low value to deter or refuse treatment. An analysis of national data assessing more than 5,000 fresh autologous cycles in which the patient’s AMH level was extremely low (<0.16 ng/mL) found that 54% of cycles were canceled, 50.7% of cycles resulted in three or fewer oocytes, and 25% had no embryo to transfer; the live birth rate was 9.5% per cycle start (20.5% per transfer) (36). Therefore, extremely low AMH values may be used to appropriately counsel women regarding suboptimal response and yield, but should not be used to refuse treatment (37).

**DO OVARIAN RESERVE TESTS PREDICT PREGNANCY AND LIVE BIRTH AFTER IVF?**

Whereas the ability of ovarian reserve tests, particularly AMH and AFC, to predict quantitative outcomes such as oocyte yield and ovarian responsiveness has been well shown, the same is not true for qualitative outcomes. Studies examining AMH as a prognostic measure for oocyte quality or clinical pregnancy and live birth rates are conflicting. Meta-analyses and studies with larger data sets, however, reveal similar results: AMH as an independent variable is only weakly predictive of pregnancy at best (38–41). Much of these data are in fresh transfers, yet one study investigating qualitative outcomes in frozen-thawed autologous transfers came to the same conclusion (41). In multiple logistic regression models, AMH is clearly an important part of the puzzle, but there are many more variables, such as female age, sperm quality, embryo development, stimulation protocols, and procedural and laboratory techniques, that play a significant role.

**COMBINED OVARIAN RESERVE TESTS**

Attempts to improve the predictive value of ovarian reserve tests by using multiple markers instead of a single baseline marker have not been found to be beneficial (41–45). High-risk scoring systems (33, 46) have been explored. Other studies have used multivariable regression models to predict either poor response to ovarian stimulation or the number of follicles/oocytes retrieved (47–50). However, complicated equations are cumbersome to apply clinically and do not provide clear cutoff points for each ovarian reserve test included. A prospective analysis of a combination of AMH, inhibin B, and three-dimensional assessment of AFC and ovarian volume concluded that the prediction was no better than that derived from each test individually (51). Another study found that including AFC and AMH in predictive models for poor response improved the model fit, but that similar accuracy was reached using AMH or AFC alone (52). A further report observed no difference in live birth when using an AFC versus an AFC-plus-AMH algorithm to determine gonadotropin dosing in IVF (43).

Despite a lack of evidence that more ovarian reserve testing improves outcomes, it is common for patients to have multiple screening tests performed. In addition to adding to overall expense, discordant results may complicate management and confuse the patient. Although generally positively correlated, AMH and AFC may be discordant up to 30% of the time according to at least one study (53). Additional research is needed to guide management in the event of discordant results.

**WHAT IS POOR OVARIAN RESPONSE?**

The tests discussed above reflect ovarian reserve which tends to correlate with oocyte yield, whereas ovarian response is defined as the actual oocyte yield after ovarian stimulation. Poor ovarian response is identified by a reduction in follicular response to maximal stimulation during the IVF procedure, resulting in a reduced number of retrieved oocytes. To standardize the definition of poor ovarian response, a European Society of Human Reproduction and Embryology Working Group convened in Bologna and proposed that two of the following criteria be present to define whether a given low response to stimulation is truly representative of poor ovarian response:
1) Advanced maternal age (≥ 40 years) or any other risk factor for poor ovarian response.
2) A previous poor ovarian response.
3) An abnormal ovarian reserve test.

Two episodes of poor ovarian response after maximal stimulation are sufficient to define a patient as a poor responder in the absence of advanced maternal age or an abnormal ovarian reserve test (54). In the setting of recently demonstrated repeated good or poor ovarian response, additional ovarian reserve testing is unnecessary.

WHICH PATIENTS SHOULD UNDERGO OVARIAN RESERVE TESTING?

When caring for a couple with infertility, clinicians use factors such as age and diagnoses to counsel individual patients and tailor the treatment plan. The goal of ovarian reserve testing is to add prognostic information to the counseling and planning process to help couples choose among treatment options, because ovarian reserve testing is testing only quantity, not quality, of the remaining oocyte pool. However, it is important to emphasize that ovarian reserve tests are not infallible and should not be the sole criteria used to deny patient access to assisted reproductive technologies or other treatments (55). Evidence of DOR does not necessarily equate with inability to conceive.

Markers of ovarian reserve should not be used as a fertility test for women who are not infertile or who have untested fertility. Consequently, markers of ovarian reserve should not be used to promote planned oocyte cryopreservation. Decisions regarding oocyte cryopreservation should be based on a woman’s reproductive plans and age. Age is a much stronger predictor of reproductive success than is ovarian reserve. All health care providers are encouraged to speak with their patients about their reproductive plans at all stages of reproductive life. Opening this dialogue and educating patients facilitates their decision making.

SUMMARY

- Dynamic tests such as the CCCT do not improve test accuracy for predicting poor ovarian response, pregnancy after IVF, or unassisted conception over basal markers and, therefore, should be abandoned.
- Currently, AMH and AFC are the most sensitive and reliable markers of ovarian reserve.
- Combined ovarian reserve test models do not consistently improve predictive ability over that of single ovarian reserve tests.
- Markers of ovarian reserve do not predict current reproductive potential among women with unproven fertility.
- Results of ovarian reserve tests are not useful in predicting the likelihood of unassisted pregnancy in women with infertility, nor do they offer clinically meaningful improvements over already established pregnancy prediction models.
- Markers of ovarian reserve do not appear to predict pregnancy after OS/IUI for unexplained infertility.

- The ability of AMH and AFC to predict oocyte yield as well as poor and excessive ovarian responsiveness in IVF has been well demonstrated.
- Extremely low AMH values should not be used to refuse treatment in IVF.
- AMH and AFC have only a weak association with qualitative outcomes such as oocyte quality, clinical pregnancy rates, and live birth rates.
- Poor ovarian response to maximal stimulation during IVF procedures reflects DOR, and further ovarian reserve testing is unnecessary.

CONCLUSION

AMH and AFC are currently the simplest, most sensitive, and specific measures of ovarian reserve. Markers of ovarian reserve have been shown to be good predictors of oocyte yield, but poor independent predictors of reproductive potential. Therefore, they should not be used as a fertility test or to deny access to infertility treatment.

Acknowledgments: This report was developed under the direction of the Practice Committee of the American Society for Reproductive Medicine as a service to its members and other practicing clinicians. Although this document reflects appropriate management of a problem encountered in the practice of reproductive medicine, it is not intended to be the only approved standard of practice or to dictate an exclusive course of treatment. Other plans of management may be appropriate, taking into account the needs of the individual patient, available resources, and institutional or clinical practice limitations. The Practice Committee and the Board of Directors of the ASRM have approved this report.

This document was reviewed by ASRM members and their input was considered in the preparation of the final document. The following members of the ASRM Practice Committee participated in the development of this document: Alan Penzias, M.D.(Chair), Ricardo Azziz, M.D., M.P.H., M.B.A., Kristin Bendikson, M.D., Tommaso Falcone, M.D., Karl Hansen, M.D., Ph.D., Micah Hill, D.O., William Hurd, M.D., MPH, Sangita Jindal, Ph.D., Suleena Kalra, M.D., M.S.C.E., Jennifer Mersereau, M.D., Catherine Racowsky, Ph.D.: Robert Rebar, M.D., Richard Reindollar, M.D., Chevis N. Shannon, Dr.P.H., M.P.H., M.B.A., Anne Steiner, M.D., M.P.H., Dale Stovall, M.D., Cigdem Tanrikut, M.D., Hugh Taylor, M.D., and Belinda Yauger, M.D. The Practice Committee acknowledges the special contribution of Amber Cooper, M.D., M.S.C.L., and Violet Klenov, M.D., in the preparation of this document. All Committee members disclosed commercial and financial relationships with manufacturers or distributors of goods or services used to treat patients. Members of the Committee who were found to have conflicts of interest based on the relationships disclosed did not participate in the discussion or development of this document.

REFERENCES


21. Tal R, Seifer DB, Wantman E, Baker V, Tal O. Antimullerian 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. Fer-
43. Magnusson A, Nilsson L, Olerod G, Thurn-Kjellberg A, Bergh C. The addi-
tion of anti-müllerian hormone in an algorithm for individualized hormone
dosage did not improve the prediction of ovarian response-a randomized,
44. Jayaprakasan K, Campbell B, Hopkisson J, Johnson I, Raine-Fenning N. A
prospective, comparative analysis of anti-müllerian hormone, inhibin-B,
and three-dimensional ultrasound determinants of ovarian reserve in the
prediction of poor response to controlled ovarian stimulation. Fertil Steril
2010;93:855–64.
dose selection using markers of ovarian reserve for women undergoing
in vitro fertilisation plus intracytoplasmic sperm injection (IVF/ICSI). Cochrane
46. Nelson SM, la Marca A. The journey from the old to the new AMH assay:
how to avoid getting lost in the values. Reprod Biomed Online 2011;4:
47. McIlveen M, Skull JD, Ledger WL. Evaluation of the utility of multiple
endocrine and ultrasound measures of ovarian reserve in the prediction of
cycle cancellation in a high-risk IVF population. Hum Reprod 2007;22:
778–85.
challenge test: cycle to cycle variability of cycle day 10 follicle stimulating
49. Hendriks DJ, Mol BW, Bancsi LF, te Velde ER, Broekmans FJ. Antral follicle
count in the prediction of poor ovarian response and pregnancy after
in vitro fertilization: a meta-analysis and comparison with basal follicle-
50. Ferraretti AP, la Marca A, Fauser BC, Tarlatzis B, Nargund G, Gianaroli L.
ESHRE Working Group on Poor Ovarian Response Definition. ESHRE
consensus on the definition of ‘poor ovarian response’ to ovarian stimulation
of different ovarian reserve markers for predicting the numbers of oocytes
52. Broer SL, van Disseldorp J, Broeze KA, Dollemann M, Opmeer BC, Bossuyt P,
et al, IMPORT Study Group. Added value of ovarian reserve testing on patient
characteristics in the prediction of ovarian response and ongoing pregnancy:
53. Li HW, Lee VC, Lau EY, Yeung WS, Ho PC, Ng EH. Ovarian response and cu-
mulative live birth rate of women undergoing in-vitro fertilisation who had
discordant anti-müllerian hormone and antral follicle count measurements:
54. Jarvela IY, Sladkevicius P, Kelly S, Ojha K, Campbell S, Nargund G. Quan-
tification of ovarian power Doppler signal with three-dimensional ultraso-
nography to predict response during in vitro fertilization. Obstet Gynecol
55. Scott RT Jr, Elkind-Hirsch KE, Styne-Gross A, Miller KA, Frattarelli JL. The pre-
dictive value for in vitro fertility delivery rates is greatly impacted by the
method used to select the threshold between normal and elevated basal
Revisando e interpretando las medidas de la reserva ovárica: un comité de opinión.

La reserva ovárica se define como el número de ovocitos que permanecen en el ovario o cantidad de ovocitos (número de ovocitos). Marcadores de la reserva ovárica incluyen niveles hormonales y parámetros medidos por ecografía de los ovarios. Estos marcadores pueden ser muy útiles como predictores del rendimiento de los ovocitos siguiendo una estimulación ovárica controlada y extracción de ovocitos. Sin embargo, estos son predictores pobres del potencial reproductivo independientemente de la edad. Este documento reemplaza el último documento con el mismo nombre publicado en 2012 (Fertil Steril 2012;98:1407–15).