
A Practical Approach to Recent Advances in Ovarian Reserve Testing

5

Benjamin Leader and Valerie L. Baker

“It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity...”
Charles Dickens, *A Tale of Two Cities*

5.1 Introduction

5.1.1 Progress

For pre and perimenopausal women, current clinically available ORTs provide important new benefits (Table 5.1) which primarily has been driven for the last decade by advancements in clinical research, much of which incorporates the use of antral follicle count (AFC) and serum biomarkers such as anti-mullerian hormone (AMH). In the field of assisted reproductive technologies (ART), progress with ORT clinical research has led to improved clinical practice with respect to prediction of ovulatory response [1–6] and optimization of oocyte retrieval using ORTs that can help more efficiently dose medications [7] and minimize side effects such as ovarian hyperstimulation syndrome (OHSS)

[8, 9]. Although not ready for routine, general use, under appropriate guidance by a fertility specialist, women can now obtain noninvasive, widely accessible ORTs that provide clinically useful general information regarding their egg supply and likelihood of menopause being earlier relative to the population average [10–13]. The recent advances with ORTs have far reaching implications for improvements in medical care by earlier detection of primary ovarian insufficiency (POI) and polycystic ovary syndrome (PCOS), counseling regarding use of fertility preservation, assessment of ovarian injury via surgery or medications such as chemotherapy, and monitoring of ovarian-related cancers [14]. These advances can, therefore, directly improve the quality of life for many women and their partners through better medical management as well as more informed decision making across a wide spectrum of medical topics.

B. Leader, M.D., Ph.D. (✉)
Clinical Research Division, ReproSource Inc., 300 Trade Center, Suite 6540, Woburn, MA 01801, USA
e-mail: leader@reprosource.com

V.L. Baker, M.D.
Division of Reproductive Endocrinology and Infertility,
Department of Obstetrics and Gynecology, Stanford University School of Medicine, Palo Alto, CA, USA

5.1.2 Challenges

Although correctly interpreted ORT results currently have great potential to benefit patients, there remains a significant risk that the results may be misinterpreted either by the clinician or

Table 5.1 Overview of clinical applications of ORTs

ORT clinical application	Clinical outcomes assessed	Current clinical uses	Limitations	Research applications
Predict response to COS	Follicular response, # of oocyte retrieved, cycle cancellation, excessive response/OHSS	Used to modify COS medication/protocol to reduce incidence of hyper-response, counsel poor responders	No consensus, Site-specific cut points, protocols	Optimize number of oocytes retrieved and medication dosing, improve cost effectiveness, and reduce adverse reactions
Oocyte quality, ART success	Rates of fertilization, blastocyst formation, implantation, live birth	Counsel patients about likelihood of success, which is program specific	Variable approaches, strong disagreement as to clinical value of ORT use	Algorithms to establish individualized probability estimates and improve success
Natural fertility	Live birth after attempts at natural conception	At best, inform women of possible increased risk of infertility	Thresholds not available for routine clinical use	Identifying who is at risk for being infertile, predict current fertility, and fertility window
PCOS risk/diagnosis	PCOS diagnosis via Rotterdam criteria	Further PCOS eval for high risk, used by some in diagnosis if APC unavailable	No agreed upon thresholds	PCOS-specific treatment protocols
Primary Ovarian insufficiency	Menstrual irregularities or absence, infertility, poor response to COs	If abnormal proactive fertility assessments, planning of future reproductive attempts	Thresholds not available for routine screening	Establish screening protocol with associated medical care algorithms
Menopause Prediction	Last menstrual period, menopausal staging criteria	Qualitative information only regarding possibly increased risk of menopause earlier than average	Specific time estimates not ready for clinical use	Establish accurate and specific predication of time to perimenopause and last menstrual period

the patient. Furthermore, the medical community remains far from reaching consensus regarding the use of ORTs [15]. The causes for this concerning lack of consensus can be grouped into two major areas: (1) lack of standardization and (2) lack of tests that can assess egg quality. In the generation and application of ORT results, major challenges exist in standardizing testing materials and methods and in widespread definitional differences used in research and clinical care with respect to the phenotypes of patients tested, medical indications, clinical outcomes managed, and selected diagnostic cut points. Overcoming these challenges is further hampered by the testing technology itself which currently only demonstrates strong prediction of oocyte quantity not oocyte quality, both of which are needed for full assessment of ovarian reserve and chances of pregnancy. Underlying the difficulty in connecting ORT results to oocyte quality is that oocyte quality ultimately is proven by the success of an oocyte to develop into a healthy baby which is a process that requires many other factors in addition to oocyte quality.

5.2 Definitions

5.2.1 Ovarian Reserve

A woman's reproductive potential, as determined by her oocyte quantity and quality, is often defined as her ovarian reserve. Although multiple factors contribute to a woman's ability to have a baby, an assessment of ovarian reserve allows approximation of a woman's fertility potential as it relates to the contribution from her oocytes. However, currently, no test can definitively determine how many oocytes a woman has and/or which oocytes are capable of conceiving an embryo that can become a healthy baby. Therefore, ovarian reserve functionally is defined in the literature by those clinical outcomes that can be measured. The advent of the ART field has provided an artificial circumstance that allows measurement of a wide variety of clinical outcome parameters not available for measure in natural reproduction. In fact, until recently [16, 17], available tests of ovarian reserve were

not validated to any natural fertility parameter but were mainly calibrated to surrogates of only ovarian quantity obtained from ART treatment outcomes, such as oocytes retrieved through controlled ovarian stimulation (COS). Although future studies may prove otherwise, when compared to age alone, ORTs have not consistently demonstrated a substantially superior ability to predict chance of spontaneous conception or live birth rate with fertility treatment.

5.2.2 Oocyte Quantity

Although recently a question has been raised as to whether human oocytes may be regenerated later in life [18], most data support the concept that oocyte supply is set at birth and is depleted over time [19]. Ironically, as the number of oocytes is not actually measurable directly without removing and dissecting the ovaries, clinical measurements of oocyte quantity are defined qualitatively. Sonographic assessment of the number of growing follicles appears to correlate with the total number of oocytes as quantified histologically [20]. Another quantifiable, clinically available measure of oocyte quantity is the number of oocytes retrieved through COS. In order for response to COS to provide a reasonable assessment of ovarian reserve, gonadotropins must be administered at doses chosen to achieve an oocyte number that maximizes live birth rate without undue risk of OHSS. The number of oocytes retrieved during an IVF attempt functions as a surrogate to approximate the number of remaining oocytes in woman [4, 6, 21–24].

5.2.3 Oocyte Quality

Oocyte quality generally refers to the ability of an oocyte to perform its primary function: to produce a healthy baby in conjunction with the genetic material supplied by a sperm. However, the creation of a healthy baby involves a multitude of factors such that the oocyte plays the classic scientific “necessary but not sufficient” role. It is currently difficult to independently measure and accurately quantify the non-oocyte contributions that are required to have a healthy baby such

as sperm or endometrial quality. Furthermore, when fewer live births occur than embryos transferred to the recipient woman, it has historically not been possible to definitively link an individual oocyte to the individual baby born. Recently, however, the increased use of elective single embryo transfer and “genetic fingerprinting” of each embryo prior to transfer, allows individual assessments of oocytes or embryos to be linked to their outcome [25–27]. Currently, oocyte quality for a woman is not assessed at the individual oocyte level but generally is inferred from calculated rates or averages from clinical endpoints such as fertilization rate, blastocyst formation rate, morphologic assessment of embryo quality, implantation rate, and live birth rate.

5.2.4 Cumulative Live Birth Rate/Total Reproductive Potential

As a concept, the number of oocytes available that are capable of producing a healthy baby can be further larger concepts such as cumulative live birth rate or total reproductive potential. There is a growing sentiment that the live birth rate per cycle has perhaps been overemphasized as a measure of fertility treatment success, and instead perhaps more focus should be placed on the cumulative chance of live birth rate over a course of treatment which may include multiple cycles of intrauterine insemination or multiple fresh and frozen cycles of IVF [28]. The term “total reproductive potential” has been introduced and is defined as the chance of live birth from one ovarian stimulation and oocyte retrieval, including the pregnancies from all fresh and all frozen embryo transfers associated with this ovarian stimulation [29]. Nearly all publications to date which have examined the prognostic value of ORTs for IVF cycles have focused on the live birth rate from one fresh transfer, not the cumulative live birth rate over multiple fresh and frozen cycles, nor the total reproductive potential from a single cycle. It would be valuable that future studies examining the prognostic value of ORTs also assess cumulative live birth rate and/or total reproductive potential as outcomes of interest.

Table 5.2 Qualitative overview of ORT correlation strength to various clinical outcomes

ORT	Modality	Response to COS	Live birth rate	Natural fertility	PCOS	POI	Menopause	Comment
AFC	Imaging	++	+	+	++	++	+	Widely available in ART centers. Highly user dependent, although automated systems may reduce user variability. Obesity may limit use
AMH	Biomarker	++	+	+	++	++	++	Measure across menstrual cycle. Can vary significantly within individuals, although less than other ORTs. Multiple charging diagnostic platforms in past and in future with no reference materials
FSH	Biomarker	+	+	+	+	++	++	High cut points identify poor response or success in a small percentage of patients but with poor sensitivity
Inhibin B	Biomarker	+/-	-	-	-	+	+	Confirmatory of other markers. Recent change in available diagnostic platforms associated with poor performance in many studies
CCCT/ EFFORT (rise in FSH/Inh B)	Patient response	++	+	-	-	+	-	Can increase predictive performance of FSH and Inhibin B. Requires two measurements and medication dose with modest additive information to single marker such as AMH or AFC. Infrequently used
Outcome of Prior ART cycle	Patient response	++	++	-	-	++	+	Not possible to use in initial assessment. Stronger data in older patients
Combinations	Multivariate	+	+	-	-	-	++	Heterogeneity of study design and ORTs prevent meaningful comparisons. Published studies show at best modest improvements currently but with continuing improvement

++ multiple published studies from multiple sources supporting strong correlation; + some evidence establishing a correlation with possible contradictory results in other studies; +/- recent studies not supportive of association; - insufficient evidence to support association

5.3 Modalities of Ovarian Reserve Testing

Broadly speaking, in conjunction with a proper history and physical exam, there are at least three common modalities of ORTs: imaging, biomarker testing, and ovarian response itself (Table 5.2).

5.3.1 Imaging

Ultrasonography is the imaging modality of choice for testing ovarian reserve generally via a transvaginal ultrasound probe which can provide ovarian volume measurements or antral follicle counts (AFC). AFC is the more commonly used metric in the literature and identifies follicles

generally from 2 to 10 mm in diameter [30]. Although AFC is most frequently obtained manually counting follicle diameter, there are efforts to automate the processing of the images to provide count and volume measurements with the thought there would be less user-dependent variability [31, 32]. Ovarian volume has also been considered as a potential ORT, but studies demonstrate it not to be as predictive of ovarian response as AFC [5, 30, 33]

5.3.2 Biomarkers

Biomarker testing primary involves biochemical evaluation of the hypothalamic pituitary ovarian (HPO). A frequently used biomarker historically and currently is follicle stimulating hormone (FSH) which is secreted by the pituitary and is well known to begin to rise early in the menstrual cycle to stimulate follicles to mature and become candidates for ovulation [34]. Excess FSH secretion and follicle stimulation is prevented through subsequent FSH suppression by rising levels of estradiol from oocytes, as well as by the glycoprotein hormone, inhibin B, which is produced by granulosa cells of pre-antral and antral follicles [35]. FSH secretion may vary widely from cycle to cycle (perhaps warranting the nickname “Fluctuating Severely Hormone”), with the prognostic value of the test being most accurate with the highest values [36, 37]. This fluctuation creates the problem that FSH may often be falsely reassuring regarding the status of ovarian reserve [38]. Antimüllerian hormone (AMH) is also a glycoprotein secreted by granulosa cells like inhibin B but from early stage follicles and acts to inhibit FSH effects on the follicle [39]. AMH is different from FSH and inhibin b in that levels during the menstrual cycle remain fairly constant when averaged across a population [40–44]. However, it should be emphasized that within individuals, there can be significant changes in AMH levels within a cycle [45]. While AMH variability is clinically significant (perhaps also deserving a nickname, “Also Meandering Hormone”) it shows less variability than most other ORTs when remeasured. Lastly, AMH has been shown at a population level to decline gradually in an almost linear

fashion [46–49], while FSH is known to remain relatively constant or rise slowly until a rapid rise is observed in the perimenopausal stage [19]. An important area of research is to determine within individuals what patterns of AMH decline exist which underlie the gradual age-dependent decline in average AMH values observed at a population level. It also is possible that at some point in the future, genetic markers such as FMR1 will also be tested more routinely to help predict whether a woman is at risk for development of premature depletion of oocyte supply [50, 51].

5.3.3 Ovarian Response

Incorporation of the patient response to the diagnostic process can be assessed with a mixture of medication and multiple biomarker measurements, referred to as dynamic or provocative testing. In addition, the actual outcome of an ART cycle itself has been reported to predict future response in certain patient populations. Commonly cited dynamic tests include the clomiphene citrate challenge test (CCCT) which measures serum FSH just prior and after 5 days of clomiphene treatment beginning on cycle day 5; the exogenous FSH ovarian reserve test (EFORT) which measures serum FSH and/or inhibin B just prior to administration FSH on cycle day 3, then measured again 24 h later [1, 5, 52–54]. Attempting to incorporate patient response into the diagnostic assessment is expensive and logistically difficult which likely has decreased the prevalence of the use of this modality. However, ultimately, the number of high quality oocytes retrieved in COS may be considered one of the major clinical outcomes of interest and closest surrogate for quantitative aspects of ovarian reserve. Thus, the patient’s response to COS itself serves as a helpful modality to assess ovarian reserve [55, 56].

5.3.4 Multivariate Approaches

As more ORTs become available and more patient subphenotypes are defined, the clinician is faced with an increasing number of variables. This presents the challenges of answering which tests

are most predictive of the outcome of interest, are several tests better than one test, and how should the tests be weighted? The reality is each clinician uses a multivariate approach when making daily decisions, often referred to as the “art of medicine.” The clinician must intuitively weight dozens of variables contained in the past medical history, age, and physical exam with the ORT results but without clear data about how many of these inputs change accuracy. Attempts are now being made to potentially improve the performance of ORTs by combining them mathematically in algorithms to allow optimized weighting and produce clinically usable information [31, 57–61]. The same issues that prevent consensus with single ORT use are magnified with use of index scores and multivariate approaches—which makes it even more difficult to compare studies. Currently, the gains shown by published studies are modest at best for use of ORTs and age at predicting COS response and success of ART treatments and have conflicting conclusions. Meta analyses that seek to combine data from multiple centers and laboratories can be problematic given the heterogeneity of the testing methods, patients, and treatment protocols and it is not surprising that they obtain results that show poor associations [62–64]. Yet, if multivariate models are used to synthesize consistent ORT methodology, patient populations, and treatments, it is quite possible that the information obtained from combining biomarkers, imaging techniques, and genetic variants, will be more informative and easier to apply clinically.

5.4 Current Clinical Applications of Ovarian Reserve Tests

Descriptions regarding the current clinical applications of ORTs (Table 5.2) are provided below but there are certain caveats that apply almost uniformly to these applications:

- First, the wide variety of definitions used for patient populations, exposures, ORT selection, and methodology, prevents any actual cut points from being generally recommended without first defining the aforementioned variables precisely.

- Secondly, ORT values exist on a continuum and can fluctuate within individuals due to inherent biological variability, such that single measurements can be misleading with frequencies that depend upon the ORT and patient population. Thus, cut points for ORTs, which are useful to compare assays or establish clinical algorithms, should be used cautiously and the reliance on one ORT modality should be avoided for definitive management decisions.

The consequence of these caveats is that practical approaches may require more effort by the clinician when initially establishing a clinical strategy to navigating the use of ORTs including (a) gaining an understand from where cut points and value ranges were derived for a chosen ORT source and (b) if that relates appropriately to the clinical outcomes and patient population being managed.

5.4.1 ORTs for Predicting Response To Controlled Ovarian Stimulation

Although additional applications of ORTs are developing and in clinical practice, identifying low and high responders to COS may be the most well-established use. The term “low” rather than “poor” and “high” rather than “good” is selected here to emphasize and focus on the quantitative aspect of response to COS separately from oocyte quality and ART cycle success.

5.4.2 Low Responders

The literature can be confusing as most of the ORTs have studies demonstrating cut points which can yield sensitivities and/or specificities above 80–90 %. There now have been a number of studies that have compared the performance characteristics of most ORTs together, including basal FSH, inhibin B, estradiol, AMH, and AFC. AMH and AFC perform fairly consistently with greater overall correlation to low response than age or other single ORTs, which, given the heterogeneity of study designs, attests to their strong correlation to response to COS [1, 3, 6, 21–23, 31, 54, 65]. Although some studies have tried to

determine which performs better, AFC or AMH, results have shown fairly similar performances when both ORTs are performed well although some may believe AFC to be slightly better than AMH when in the hands of experienced clinicians [66, 67]. It should be noted, however, that none of the ORTs have demonstrated, through multiple publications from several groups, sufficient sensitivity or specificity to predict with certainty the outcome in ART, even for oocyte quantity.

Some studies have shown basal serum FSH to have clinically helpful specificity for poor response [36, 68], with one study showing of up to 100 % specificity but only when a high cut point for normal is utilized and with sensitivity too low to be used alone as an ORT [68]. Basal Inhibin B, initially showed promise as an ORT in studies using the inhibin B system from Serotec, LTD [69, 70]. However, subsequent studies [1, 6] failed to reproduce similar accuracy for inhibin B, commensurate, interestingly, with the lack of availability of the Serotec platform. Dynamic or provocative tests such as CCCT and EFORT (using both FSH and inhibin B) [2, 71–73] have consistently shown clinically useful sensitivity and/or specificity often superior to other single ORTs. However, the requirement for two measurements and medication has likely led to minimal use, especially when evidence exists that a single measurement of a single ORT may have sufficiently similar accuracy [54].

5.4.3 High Responders

Certain ORTs consistently demonstrate the significant ability to predict, independently of age, which women will likely be high responders to COS which has important benefits to reduce complications of excessive response (e.g., OHSS and cycle cancelation) and also to reduce consumption of gonadotropins. There are now numerous studies demonstrating clinical utility of ORTs with respect to a wide variety of definitions of excessive response including high estradiol levels, withdrawal of stimulation (“coasting”), cycle cancelation, high number of oocytes retrieved, and more severe conditions associated with OHSS

such as accumulation of ascites and hospitalization [8, 54, 62, 74–76]. For example, in a study of 110 patients with excessive response defined as greater than 20 oocytes retrieved, investigators could demonstrate that an AFC cut point could select 11 % of patients and identify hyper-response with 50 % and 96 % sensitivity and specificity, respectively [54]. Using moderate and severe OHSS as a clinical outcome, in a study of 262 patients, an AMH cut point which identified 25 % of the patients also performed with 91 % sensitivity and 81 % specificity, respectively for OHSS [8]. However, despite the variation in the definition of excessive response outcome and also variation with cut point selection, AFC and AMH showed across multiple studies clinical helpful performance characteristics and frequently performed better than most other ORTs for both sensitivity and specificity. As both AMH and AFC measurements exist along a continuum, for practical implementation, one must choose the definition of excessive response and identify internal thresholds for management changes.

5.4.4 Oocyte Quality, Live Birth Rate in ART

With respect to ART treatments, the studies performed to date have not demonstrated with sufficient consistency or robust predictive power a clinically helpful relationship between ORT results and oocyte quality or pregnancy success that is widely applicable with specific cut points [19]. That said there have been studies which demonstrate remarkable results in specific circumstances that could dramatically help guide care. For example, in a study of serum basal FSH measurements in over 8,000 cycles from one center with a single FSH measurement source, FSH thresholds could make clinically helpful, age group specific, robust predictions of chances of live delivery per ART cycle start along a continuum of values [68]. Values above certain thresholds demonstrated 100 % specificity for failed cycles although those thresholds only identified about 1 in 30 women tested above 40 years of age and 1 in 324 women tested under age 35. However, other differently structured studies arrive at dif-

ferent conclusions such as FSH being valuable predicting live birth only in certain age groups [77] versus no ability to predict live birth better than age alone as concluded by a recent meta-analysis which used 28 databases to aggregate data from 5,705 IVF patients and multiple FSH diagnostic platforms [64]. A number of studies indicate that AMH or AFC levels do not predict treatment success [21, 78]. This conflicts with the findings other published findings [60, 76, 79] including a recent study externally validated an AMH-based live birth prediction model to, independent of age, predict live birth in 822 patients with statistical significance, although the confidence intervals were wider than some may view as clinically helpful [59].

The lack of consensus and conflicting medical literature is not surprising given the multifactorial nature of embryo development into a healthy baby. However, the heterogeneity of study designs, an inability to control for confounding variables, and insufficiently robust biological association of ORTs to live birth rate, presents serious hurdles to overcome in the quest for consensus. Thus, applications of ORTs in predicting live birth currently must remain a user-defined, site-specific approach. Future studies that examine the prognostic value of AMH, AFC, or other tests on cumulative live birth rate or total reproductive potential as described above are needed. It is quite plausible that any measure that predicts oocyte number of retrieval may be a better predictor of the success of fresh and frozen embryo transfer combined, than it would be of fresh cycles only because more embryos are likely to be frozen if a greater number of oocytes are retrieved.

5.4.5 Overall Fertility and Recurrent Pregnancy Loss

Clinical justification for ORT use in the general population to assess fertility is beginning to appear. Evidence is mounting that infertility is associated with lower ORT values as demonstrated by lower AFC in 881 infertile women without PCOS compared to 771 women without the diagnosis of infertility [16]. In another pro-

spective study of 100 general population women attempting to conceive, early follicular phase AMH was shown to predict fertility rates [17]. Thus, it appears promising that ORT results will play a future role in fertility assessment of the general population.

Data on miscarriage and ORTs are scant and primarily derive from patients receiving ART treatment. One retrospective study showed no association with highest serum basal FSH and fetal aneuploidy [80] in 177 spontaneous miscarriages associated with 70 euploid and 107 aneuploid offspring. No association with AFC, FSH, and CCCT was demonstrated prospectively comparing values in 77 women with pregnancy loss versus 233 with ongoing pregnancy [52]. However, AFC was shown to be predictive of only first trimester loss in 67 patients with miscarriage compared to 247 controls with ongoing pregnancy, although the overall association was weak with an ROC curve AUC of 0.588 [81]. Recently, in a study of women undergoing aneuploidy screening of embryos followed by IVF of 279 women, those with reassuring FSH and AMH values generated lower rate of all aneuploid blastocysts compared to 93 women with concerning FSH and/or AMH (35 % vs. 14%, $P < 0.001$) [82]. It was further noted that when both FSH and AMH were concerning, the highest percentage of aneuploid blastocysts was observed (77 %) compared to only one being concerning (58.5 %, 58.8 %) and both reassuring (51.7 %). Thus, it appears that ORTs may be useful in predicting increased risk of miscarriage.

5.4.6 PCOS, POI, and Menopause

As research has advanced and ORTs such as AFC and AMH have become more widely used, helpful clinical information for patients can be applied to help identify, diagnose, and manage other diseases and processes not strictly related to attempts to have a child.

AMH is now also being proposed by some as an alternative criterion to diagnose women with PCOS or to identify women at high risk for PCOS [83]. One recent study, which included by 66 women without

PCOS or polycystic ovaries and 62 confirmed PCOS by hyperandrogenism and oligomenorrhea, identified an optimized AMH cut point demonstrating 92 % sensitivity and 97 % specificity for PCOS [84]. However, the use of AMH in this context remains controversial and has not been adopted in official criteria for PCOS diagnosis.

Perhaps the most exciting developments relate to early detection of POI and long-term prediction of the menopausal transition and menopause. As AMH and AFC levels, at a population level, demonstrate a gradual almost linear decline [46–49], these ORTs have applications in both early detection of POI prior to symptoms and long-term prediction of menopause onset. Earlier identification of women at risk of POI may help them avoid the most severe consequences of this disease such as missing the opportunity to have children with their own eggs as well as other complications associated with early menopause such as bone loss and increased cardiovascular events [14].

Although AMH, AFC, FSH, and Inhibin B have all been published as being able to add significantly more predictive power to prediction of menopause than age alone, AMH and AFC appear to show the better performance characteristics [11–13, 85]. Furthermore, it may be rates of change are more predictive than single measurements [85, 86]. As increasing amounts of individualized longitudinal data are becoming available, the confidence intervals around age of the predicted last menstrual period are becoming narrower [87]. Subphenotypes may be further defined that can increase predictive information, such as genetic interactions with ovarian reserve. For example, one study of 240 women indicated that FMR1 repeat length was associated with a 54 % difference in AMH level [50]. Another recent study identified several genetic markers in 450 women that were associated with ovarian follicle number and menopause [88]. At this juncture, the published literature on menopause prediction appears sufficiently consistent such that, if a woman has an AMH or AFC value very low for her age using a well-calibrated testing source, it would be questionable not to alert her at least about the increased possibility of earlier than average menopause. This knowledge can allow a woman to proactively address her desired plan

for future childbearing. In addition, a woman with ORT results substantially low for her age can proactively address the risk of long-term medical issues such as osteoporosis, cardiovascular disease, and certain forms of cancer which are more prevalent in women with early menopause [14].

With the availability of clinically validated egg preservation technologies, there is now the ability to dramatically increase the length of time a woman has to have a child with her own eggs [89]. This significant advancement has clear immediate application to preserve eggs, for example, prior to receiving ovarian toxic treatments such as chemotherapy [90]. However, the combination of egg preservation and the developing predictive power of ORTs, presents society with the double-edged sword of providing a safety net for possible future ovarian reserve-related infertility, but the risk of encouraging women to delay natural attempts at conception.

5.4.7 Exogenous Hormone Use

Influence on AMH levels by exogenous hormones has been clearly demonstrated [91]. While some publications suggest that oral contraceptive pills (OCPs) do not affect AMH or AFC levels [92], it now is becoming clear that OCPs such as monophasic estrogens can lower AMH and AFC levels [93, 94]. In one study of 25 women on OCPs for more than 3 months significant improvement in AMH and AFC parameters were observed after the second menstrual cycle without OCPs [95]. This was confirmed in a complementary study with 44 women off OCPs for at least 3 months who showed an average reduction of approximately 50 % in AMH by week 9 of OCP use [96]. This indicates that if a woman has a concerning AMH or AFC while on an estrogen OCP, it may be helpful to retest after stopping the OCP use for two cycles if the retesting would change management. However, if the AMH level is reassuring while on estrogen OCPs, the above recently published studies indicate it will likely remain reassuring off OCPs. While there may be logical ways to extract clinically helpful information in certain scenarios with patients taking

OCPs, careful attention should be paid to the use of exogenous hormones when interpreting ORTs.

5.5 Current Challenges

5.5.1 Biology

One of the biggest barriers for current ORTs in achieving the desired narrowness of confidence intervals for predicting clinical outcomes is the inherent biological flux associated with biomarkers of the HPO axis. If ORT results can fluctuate in clinically significant amounts with some frequency, there is an intrinsic limit to the accuracy of the test regardless of study design, sample size, and uniformity of patient population. It has long been recognized that FSH levels fluctuate dramatically between from cycle to cycle [36, 37]. The recently more popular ORTs, AFC and AMH, receive much focus in part because the average value in the population does not show the same dramatic dependence on the stage of the menstrual cycle as FSH, inhibin B, LH, or estradiol [44, 97]. While this has important logistical benefits by not requiring measurement at a particular time of the menstrual cycle, especially in those women who do not regularly menstruate or have had a hysterectomy, this does not address the larger issue of values being clinically significantly different in the same individual when retested even within the same menstrual cycle. For example, Sowers et al. measured AMH every day of the menstrual cycle, demonstrating a consistent AMH average throughout the menstrual cycle in five groups of five women with similar AMH values [44]. However, closer examination of the data points showed two of five women with similar average AMH values having daily values of approximately 0.6 and 0.75 ng/ml for half the cycle and nearly 2 ng/ml for the other half of the menstrual cycle. This finding was recently observed again in a population of 44 women retested within a menstrual cycle [45].

The other major biologic barrier for ORTs to assess accurately the ability of a woman's oocytes to produce a healthy baby, is that from fertilization onward, numerous other confounding variables

are required in the process. A successful pregnancy depends upon many factors such as a sufficiently healthy sperm and a receptive endometrium. This presents a significant challenge both in the current ability to diagnostically assess these variables accurately and separately, and, statistically, in the number of patients needed to appropriately power studies that would seek to perform the extensive subset analysis required.

5.5.2 Standardization

While the biology of the human reproductive system is difficult to control, the fertility field is challenged with lack of consistency in almost every aspect of ORT study design to the point that the latest American Society for Reproductive Medicine practice guideline concluded that there is no consensus as to the definition of ovarian reserve and the evidence for the tests which measure it is at best "fair" [19]. Substantial variation can be seen in study population phenotypes, treatment regimes, clinical outcomes assessed, choice of ORT(s), and method of analysis, and use of cut points.

When it comes to performance of the ORTs, dramatic differences can exist in the reported value and clinical performance for the same sample depending upon the diagnostic platform chosen (Fig. 5.1). The best example perhaps of this is the history of inhibin B which showed clinically useful performance with the Serotec kit [69, 70] and not with the DSL kit that replaced it, leading to the likely unreparable clinical distrust of this biomarker [1, 6]. One misconception is that automation and FDA clearance resolve issues with consensus. While FDA clearance and automation improved the assay performance and ease of measurement for serum FSH, this has not led to establishment of consensus regarding FSH testing despite over 20 years of publications regarding its use [19]. Differences in diagnostic platforms are not clear on reports provided to clinicians frequently. These differences can be substantial as in the recent College of American Pathology Surveys 2011 Y-B Ligand publication demonstrated that 434 laboratories produced an

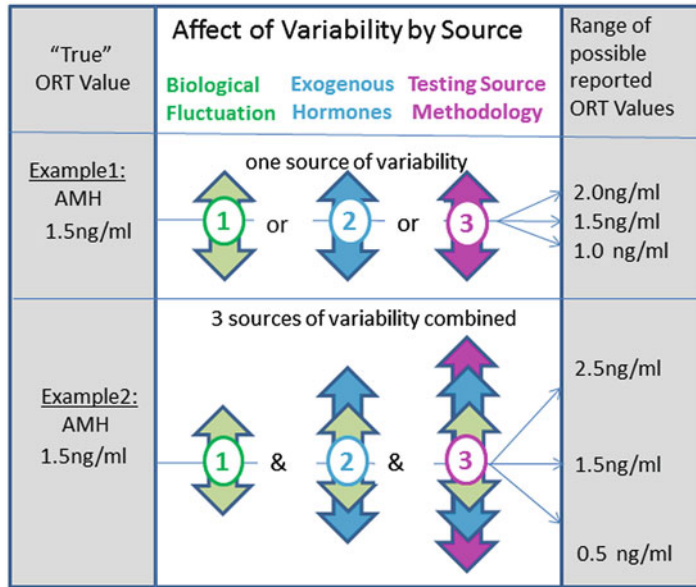


Fig. 5.1 Effect on reported ORT value by three different sources of variability. When retesting the same patient with an ORT, a minority, but significant fraction of the time one value is clinically different from a patient’s “true” or most representative value. At least three factors can affect this. (1) Biological flux of ORTs can be substantial. (2) Exposures to medications such as oral contraceptive are now known to

affect results of ORTs such as AMH and AFC. (3) Although testing methods may have minimal variability within a chosen source, the between source assay differences may be substantial. The affect of any single source of variability can be clinically significant (example 1) and even more so if multiple sources of variability are present and combine in the same direction (example 2)

acceptable mean FSH value of 34 IU/L while, with the same reference sample, 151 other laboratories produced an acceptable mean value of 19 IU/L, the difference being the FSH analyzer platform.

Unfortunately, for the two ORTs currently receiving the most attention, AFC and AMH, reference standards don’t even exist. AFC is a highly user-dependent modality and, despite attempts at international standardization, there remains some inconsistency in the size of follicle to include in the AFC with obesity further complicating interpretation or rendering it impossible [30]. The measurement of AMH has undergone three kit changes (Immunotech, DSL “GenI,” Beckman/DSL “GenII”) in the past 3 years, with a new one arriving on the market shortly along with automated platforms and blood spot tests on the way [98–100]. Although the clinical correlations observed with different AMH kits are consistent, different AMH kits often have inconsistent con-

version equations published between the others. This makes extrapolation of results from one kit to another risky to interpret a clinical report for a patient without conducting careful validation experiments. It is also very important for clinicians to be aware that values in the literature may have been performed using a different assay and thus may not be readily applicable to the results of their patients. Additionally, the use of the same AMH kit can produce dramatically different values depending on a variety of factors including the treatment of sample and the laboratory methodology. Furthermore, as previously discussed, the influence of exogenous medications such as OCPs were once considered of no consequence now are recognized as significantly affecting ORT results. However, most importantly clinical value ranges, which determine the treatment, are frequently set by the laboratory based upon CLIA requirements to establish a general mean and distribution in a general population and not upon the

Table 5.3 Practical steps to optimize the use of ORTs

6 Steps to optimize ORT use	
Recognize lack of consensus for ORT use	Recognize consensus does not currently exist regarding ORT interpretation and that utility but depends upon your chosen ORT, clinical parameters, patient population, and clinical purpose
Identify at least two ORTs	Choose at least two ORTs from preferably two different modalities if available (e.g., AFC and AMH). Minimizes impact of variability of an individual ORT fluctuation
Establish how ORT clinical value ranges were generated	Identify a consistent source of ORTs if possible and understand how value ranges relate to clinical outcomes being managed. Avoid applying thresholds used in publications which utilize ORT sources with no link to your current source. Ideally establish internal value ranges for any source of ORT
Verify no changes in calibration on a regular basis	Unfortunately, there are frequent changes in assay materials, lab methodologies, treatment modalities that can change the interpretation of results. Verify if practical once every 6 months to a year with the laboratory director of any consistently used sources of ORTs
Be aware estrogen-based OCPs may lower AMH and AFC	Recent data suggest estrogen-based OCPs can lower AMH and AFC results. However, if a patient already has reassuring values on OCPs, it is likely they would remain reassuring off OCPs. If consider retesting on the second natural cycle off OCPs if a different ORT result would change management
Avoid definitive predictions based	ORT information should refine, not define clinical management

clinical outcomes being managed by the test. For example, a “normal range” for AMH can be 0 ng/ml to 6.9 ng/ml which spans the gamut of ovarian failure (depleted ovarian reserve) to high risk for OHSS or PCOS (high ovarian reserve).

The above challenges can unfortunately be additive and pose a significant risk of clinically miscategorizing a patient if careful steps are not taken to avoid this (Fig 5.1). Fortunately, there are practical ways to minimize the chance of misguiding care with ORT use.

5.6 Practically Optimizing the Use Of ORTs

5.6.1 It's The Approach, Not Just the Test

The pattern that consistently emerges from literature assessing ORTs is that performance and utility depend upon the user's decisions regarding patient populations, treatments, ORT selection, and methodology. Furthermore, the value of a particular ORT's PPV and NPV depends upon prevalence of the clinical outcome in the intended use population, which can vary dramati-

cally, for example, with diminished ovarian reserve in an oocyte donor screening program as compared to counseling a woman about IVF using her own oocytes. Thus, minimizing the risk of misinterpretation of ORT values requires a methodical approach, which may involve some initial effort to establish (Table 5.3). One approach is described below:

- The first recommended practical step is to recognize that consensus does not currently exist regarding ORT interpretation and utility and expend the effort necessary to establish one's desired approach.
- Second, the fluctuation of ORT results and possible sources of error makes important utilizing at least two different ORTs when evaluating a patient. Frequently, this is possible as other ORTs, such as FSH and estradiol, have other utilities in the initial assessment of a patient, and therefore to combine this with AFC and/or AMH is logistically reasonable. The use of different modalities such as imaging and serum testing has the added benefit of it being less likely to have an error, such as improper specimen handling, affect both modalities.

- Third, it would be ideal, but frequently not possible, to establish a consistent source of ORTs and obtain an understanding of how the clinical value ranges are determined. The ideal scenario is that each practitioner ultimately clinically calibrates his/her ORTs against his/her own outcome data, but this is often times not feasible. Practically speaking, as it is not possible to track down the source of every outside laboratory result, judicious use of retesting at a familiar source should be considered if retesting could significantly change clinical management.
- Although perhaps an unpleasant truth, the materials and methods change, not infrequently, for ORTs and vigilance with respect to the affect this change would have on interpretation is important. If one uses regularly one or two sources for ORT results it would not be unreasonable to perform a brief inquiry of the laboratory director once or twice per year as to if there were any changes with a chosen source of ORT that could affect value ranges.
- Fifth, as many women use OCPs and it is difficult at times to stop taking them, a practical method for tests such as AMH or AFC is to obtain the values and if reassuring consider it sufficient to use this value as recent data indicates it is likely that the ORT result remain reassuring if not more so off estrogen OCPs. If AMH and AFC are concerning while on an estrogen-based OCP, one can consider then retesting off OCPs if management decisions would change.
- Sixth, overall, one should be very cautious and avoid, if possible, counseling a patient solely based upon ORT values since the certainty of outcome for these tests is not definitive. Ultimately, it is advisable to use ORTs to influence rather than direct clinical management.

5.6.2 Clinical Example

Given the especially ambiguous nature of ORT results and lack of consensus, a short case sce-

Case 1

A healthy 28-year-old female with no prior attempts at conception is considering attending medical school and presents to fertility specialist, referred by general practitioner with an AMH value of 2.0 ng/ml by an outside laboratory with normal range reported as 0–6.9 ng/ml.

Patient: “Will I still be able to have children in 8 years after I finish medical school and residency?”

Clinician: “With no family or medical history concerning for early loss of fertility, it would be wise to recheck this lab value before drawing any conclusions. In the meantime, let’s obtain an antral follicle count today by ultrasound”

AFC shows a total of ten follicles between 2 and 10 mm. Rechecking of the AMH at a different laboratory regularly used by the clinician with well-established value ranges returns value of 0.6 ng/ml which fell into a range that was consistent with the patient already being at high risk for poor egg supply. Discussion at next visit:

Clinician: “Rechecking your AMH shows you have a value that is low for your age and that you already are at risk for low egg supply. There are now several studies from several sources that show women with low AMHs are more likely to go into menopause sooner than women with high AMHs of the same age. While we can’t give you any specific prediction about your fertility window, you are at likely at higher risk than the average to have menopause earlier and thereby have a shorter fertility window. If having children right now is not want you want or are able to do, you may want to consider egg cryopreservation. While long term follow up data isn’t yet available we are cautiously optimistic about there not being significant difference between babies born through natural conception versus

(continued)

Case 1 (continued)

babies born through IVF using preserved oocytes.”

Patient: “Why was my AMH lab value so different when you retested it?”

Clinician: “It could be natural fluctuation or differences in the methods of measurement. Let’s reevaluate in the next several months and also try to figure out what you would like to do about planning your future family building goals.”

nario is presented with possible responses to better illustrate use of the recommendations. This scenario is not intended to represent consensus views or incontrovertible information.

5.7 Conclusion

Research over the past 10 years has demonstrated a wide variety of clinical utility for ORTs such as improving COS management, risk stratification for ART treatment success, identification of women at risk for infertility, more sensitive detection of diminished ovarian reserve, prediction of time until menopause, and adjunctive use to identify and/or diagnose PCOS. The ORTs, AMH and AFC, have each emerged as the two most predictive individual ORTs for responsiveness to COS for retrieval of oocytes as well as sensitive identifiers of diminished oocyte supply, proximity to menopause, and likelihood of PCOS. Many of these research findings are currently applied with clinical benefit.

While the potential advantages of ORT use in clinical medicine is clear, with the biological fluctuations in ORT results, the complexity of fertility assessment, and lack of standardization, consensus is not possible regarding most of the above utilities, and the risk of misguiding clinical care using an ORT result is high if appropriate steps are not taken by clinicians. This risk can be minimized by (1) recognizing that performance of an ORT is specific to the source of ORT and

the clinical environment in which it is applied, (2) identifying at least two different ORTs for use, (3) use a consistent source of ORT results where possible with an understanding of how the values relate to the clinical outcomes being managed, (4) inquiring periodically about assay change at a chosen ORT source which could change interpretation, (5) avoiding use cut points from publications without understanding how they apply to your source of ORT, (6) paying attention to exogenous hormone use, and (7) avoiding the use ORTs alone to make clinical decisions. This approach likely will reduce the risk of misinterpretation of results while simultaneously harnessing the information ORTs can provide to improve clinical care.

References

1. Broekmans FJ, Kwee J, Hendriks DJ, et al. A systematic review of tests predicting ovarian reserve and IVF outcome. *Hum Reprod Update.* 2006;12(6): 685–718.
2. Dzik A, Lambert-Messerlian G, Izzo VM, et al. Inhibin B response to EFORT is associated with the outcome of oocyte retrieval in the subsequent in vitro fertilization cycle. *Fertil Steril.* 2000;74(6):1114–7.
3. Gnoth C, Schuring AN, Friol K, et al. Relevance of anti-Mullerian hormone measurement in a routine IVF program. *Hum Reprod.* 2008;23(6):1359–65.
4. Jee BC, Ku SY, Suh CS, et al. Serum anti-Mullerian hormone and inhibin B levels at ovulation triggering day can predict the number of immature oocytes retrieved in in vitro fertilization cycles. *J Korean Med Sci.* 2008;23(4):657–61.
5. Kwee J, Elting ME, Schats R, et al. Ovarian volume and antral follicle count for the prediction of low and hyper responders with in vitro fertilization. *Reprod Biol Endocrinol.* 2007;5:9.
6. Riggs RM, Duran EH, Baker MW, et al. Assessment of ovarian reserve with anti-Mullerian hormone: a comparison of the predictive value of anti-Mullerian hormone, follicle-stimulating hormone, inhibin B, and age. *Am J Obstet Gynecol.* 2008;199(2):202. e1–8.
7. van Tilborg TC, Eijkemans MJ, Laven JS, Koks CA, de Bruin JP, Scheffer GJ, et al. The OPTIMIST study: optimisation of cost effectiveness through individualised FSH stimulation dosages for IVF treatment. A randomised controlled trial. *BMC Womens Health.* 2012;12:29.
8. Lee TH, Liu CH, Huang CC, et al. Serum anti-Mullerian hormone and estradiol levels as predictors of ovarian hyperstimulation syndrome in assisted

- reproduction technology cycles. *Hum Reprod.* 2008;23(1):160–7.
9. Yates AP, Rustamov O, Roberts SA, et al. Anti-Mullerian hormone-tailored stimulation protocols improve outcomes whilst reducing adverse effects and costs of IVF. *Hum Reprod.* 2011;26(9):2353–62.
 10. Freeman EW, Sammel MD, Lin H, et al. Anti-Mullerian hormone as a predictor of time to menopause in late reproductive age women. *J Clin Endocrinol Metab.* 2012;97(5):1673–80.
 11. Sowers MR, Eyvazzadeh AD, McConnell D, et al. Anti-Mullerian hormone and inhibin B in the definition of ovarian aging and the menopause transition. *J Clin Endocrinol Metab.* 2008;93(9):3478–83.
 12. Tehrani FR, Shakeri N, Solaymani-Dodaran M, et al. Predicting age at menopause from serum antimullerian hormone concentration. *Menopause.* 2011;18(7):766–70.
 13. Tehrani FR, Solaymani-Dodaran M, Azizi F. A single test of antimullerian hormone in late reproductive-aged women is a good predictor of menopause. *Menopause.* 2009;16(4):797–802.
 14. Nelson LM. Clinical practice. Primary ovarian insufficiency. *N Engl J Med.* 2009;360(6):606–14.
 15. Practice Committee of the American Society for Reproductive Medicine. Testing and interpreting measures of ovarian reserve: a committee opinion. *Fertil Steril.* 2012;98(6):1407–15.
 16. Rosen MP, Johnstone E, Addaun-Andersen C, et al. A lower antral follicle count is associated with infertility. *Fertil Steril.* 2011;95(6):1950–4. e1.
 17. Steiner AZ, Herring AH, Kesner JS, et al. Antimullerian hormone as a predictor of natural fecundability in women aged 30–42 years. *Obstet Gynecol.* 2011;117(4):798–804.
 18. White YA, Woods DC, Takai Y, et al. Oocyte formation by mitotically active germ cells purified from ovaries of reproductive-age women. *Nat Med.* 2012;18(3):413–21.
 19. Harlow SD, Gass M, Hall JE, et al. Executive summary of the Stages of Reproductive Aging Workshop + 10: addressing the unfinished agenda of staging reproductive aging. *Fertil Steril.* 2012;97(4):843–51.
 20. Hansen KR, Knowlton NS, Thyer AC, et al. A new model of reproductive aging: the decline in ovarian non-growing follicle number from birth to menopause. *Hum Reprod.* 2008;23(3):699–708.
 21. Riggs R, Kimble T, Oehninger S, et al. Anti-Mullerian hormone serum levels predict response to controlled ovarian hyperstimulation but not embryo quality or pregnancy outcome in oocyte donation. *Fertil Steril.* 2011;95(1):410–2.
 22. Hazout A, Bouchard P, Seifer DB, et al. Serum antimullerian hormone/mullerian-inhibiting substance appears to be a more discriminatory marker of assisted reproductive technology outcome than follicle-stimulating hormone, inhibin B, or estradiol. *Fertil Steril.* 2004;82(5):1323–9.
 23. Knauff EA, Eijkemans MJ, Lambalk CB, et al. Anti-Mullerian hormone, inhibin B, and antral follicle count in young women with ovarian failure. *J Clin Endocrinol Metab.* 2009;94(3):786–92.
 24. Kunt C, Ozaksit G, Keskin Kurt R, et al. Anti-Mullerian hormone is a better marker than inhibin B, follicle stimulating hormone, estradiol or antral follicle count in predicting the outcome of in vitro fertilization. *Arch Gynecol Obstet.* 2011;283(6):1415–21.
 25. Forman EJ, Hong KH, Treff NR, et al. Comprehensive chromosome screening and embryo selection: moving toward single euploid blastocyst transfer. *Semin Reprod Med.* 2012;30(3):236–42.
 26. Forman EJ, Tao X, Ferry KM, et al. Single embryo transfer with comprehensive chromosome screening results in improved ongoing pregnancy rates and decreased miscarriage rates. *Hum Reprod.* 2012;27(4):1217–22.
 27. Treff NR, Su J, Kasabwala N, et al. Robust embryo identification using first polar body single nucleotide polymorphism microarray-based DNA fingerprinting. *Fertil Steril.* 2010;93(7):2453–5.
 28. Luke B, Brown MB, Wantman E, et al. Cumulative birth rates with linked assisted reproductive technology cycles. *N Engl J Med.* 2012;366(26):2483–91.
 29. Stern JE, Hickman TN, Kinzer D, et al. Can the Society for Assisted Reproductive Technology Clinic Outcome Reporting System (SART CORS) be used to accurately report clinic total reproductive potential (TRP)? *Fertil Steril.* 2012;97(4):886–9.
 30. Broekmans FJ, de Ziegler D, Howles CM, et al. The antral follicle count: practical recommendations for better standardization. *Fertil Steril.* 2010;94(3):1044–51.
 31. Nelson SM. Biomarkers of ovarian response: current and future applications. *Fertil Steril.* 2013 Jan 8.
 32. Jayaprakasan K, Campbell B, Hopkisson J, et al. A prospective, comparative analysis of anti-Mullerian 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(3):855–64.
 33. Johnson NP, Bagrie EM, Coomarasamy A, et al. Ovarian reserve tests for predicting fertility outcomes for assisted reproductive technology: the International Systematic Collaboration of Ovarian Reserve Evaluation protocol for a systematic review of ovarian reserve test accuracy. *BJOG.* 2006;113(12):1472–80.
 34. Broekmans FJ, Soules MR, Fauser BC. Ovarian aging: mechanisms and clinical consequences. *Endocr Rev.* 2009;30(5):465–93.
 35. Klein NA, Houmar BS, Hansen KR, et al. Age-related analysis of inhibin A, inhibin B, and activin a relative to the intercycle monotropic follicle-stimulating hormone rise in normal ovulatory women. *J Clin Endocrinol Metab.* 2004;89(6):2977–81.
 36. Scott RT, Toner JP, Muasher SJ, et al. Follicle-stimulating hormone levels on cycle day 3 are pre-

- dictive of in vitro fertilization outcome. *Fertil Steril.* 1989;51(4):651–4.
37. Sharif K, Elgendy M, Lashen H, et al. Age and basal follicle stimulating hormone as predictors of in vitro fertilisation outcome. *Br J Obstet Gynaecol.* 1998;105(1):107–12.
 38. Leader B, Hegde A, Baca Q, et al. High frequency of discordance between antimullerian hormone and follicle-stimulating hormone levels in serum from estradiol-confirmed days 2 to 4 of the menstrual cycle from 5,354 women in U.S. fertility centers. *Fertil Steril.* 2012;98(4):1037–42.
 39. Nilsson EE, Schindler R, Savenkova MI, Skinner MK. Inhibitory Actions of Anti-Mullerian Hormone (AMH) on Ovarian Primordial Follicle Assembly. *PLoS One.* 2011;6(5):e20087.
 40. Hehenkamp WJ, Looman CW, Themmen AP, et al. Anti-Mullerian hormone levels in the spontaneous menstrual cycle do not show substantial fluctuation. *J Clin Endocrinol Metab.* 2006;91(10):4057–63.
 41. La Marca A, Giulini S, Tirelli A, et al. Anti-Mullerian hormone measurement on any day of the menstrual cycle strongly predicts ovarian response in assisted reproductive technology. *Hum Reprod.* 2007;22(3):766–71.
 42. La Marca A, Sighinolfi G, Giulini S, et al. Normal serum concentrations of anti-Mullerian hormone in women with regular menstrual cycles. *Reprod Biomed Online.* 2010;21(4):463–9.
 43. La Marca A, Stabile G, Arsenio AC, et al. Serum anti-Mullerian hormone throughout the human menstrual cycle. *Hum Reprod.* 2006;21(12):3103–7.
 44. Sowers M, McConnell D, Gast K, et al. Anti-Mullerian hormone and inhibin B variability during normal menstrual cycles. *Fertil Steril.* 2010;94(4):1482–6.
 45. Overbeek A, Broekmans FJ, Hehenkamp WJ, et al. Intra-cycle fluctuations of anti-Mullerian hormone in normal women with a regular cycle: a re-analysis. *Reprod Biomed Online.* 2012;24(6):664–9.
 46. Seifer DB, Baker VL, Leader B. Age-specific serum anti-Mullerian hormone values for 17,120 women presenting to fertility centers within the United States. *Fertil Steril.* 2011;95(2):747–50.
 47. Almog B, Shehata F, Suissa S, et al. Age-related nomograms of serum antimullerian hormone levels in a population of infertile women: a multicenter study. *Fertil Steril.* 2011;95(7):2359–63. 63 e1.
 48. Nelson SM, Messow MC, McConnachie A, et al. External validation of nomogram for the decline in serum anti-Mullerian hormone in women: a population study of 15,834 infertility patients. *Reprod Biomed Online.* 2011;23(2):204–6.
 49. Nelson SM, Messow MC, Wallace AM, et al. Nomogram for the decline in serum antimullerian hormone: a population study of 9,601 infertility patients. *Fertil Steril.* 2011;95(2):736–41. e1–3.
 50. Spath MA, Feuth TB, Allen EG, et al. Intra-individual stability over time of standardized anti-Mullerian hormone in FMR1 premutation carriers. *Hum Reprod.* 2011;26(8):2185–91.
 51. Gleicher N, Weghofer A, Kim A, et al. The impact in older women of ovarian FMR1 genotypes and sub-genotypes on ovarian reserve. *PLoS One.* 2012;7(3):e33638.
 52. Haadsma ML, Groen H, Fidler V, et al. The predictive value of ovarian reserve tests for miscarriage in a population of subfertile ovulatory women. *Hum Reprod.* 2009;24(3):546–52.
 53. Kwee J, Schats R, McDonnell J, et al. The clomiphene citrate challenge test versus the exogenous follicle-stimulating hormone ovarian reserve test as a single test for identification of low responders and hyper-responders to in vitro fertilization. *Fertil Steril.* 2006;85(6):1714–22.
 54. Kwee J, Schats R, McDonnell J, et al. Evaluation of anti-Mullerian hormone as a test for the prediction of ovarian reserve. *Fertil Steril.* 2008;90(3):737–43.
 55. Nelson SM, Lawlor DA. Predicting live birth, preterm delivery, and low birth weight in infants born from in vitro fertilisation: a prospective study of 144,018 treatment cycles. *PLoS Med.* 2011;8:e1000386.
 56. Oudendijk JF, Yarde F, Eijkemans MJ, et al. The poor responder in IVF: is the prognosis always poor?: a systematic review. *Hum Reprod Update.* 2012;18(1):1–11.
 57. Al-Azemi M, Killick SR, Duffy S, et al. Multi-marker assessment of ovarian reserve predicts oocyte yield after ovulation induction. *Hum Reprod.* 2011;26(2):414–22.
 58. Kelsey TW, Wright P, Nelson SM, Anderson RA, Wallace WH. A validated model of serum anti-Mullerian hormone from conception to menopause. *PLoS One.* 2011;6(7):e22024.
 59. Khader A, Lloyd SM, McConnachie A, et al. External validation of anti-Mullerian hormone based prediction of live birth in assisted conception. *J Ovarian Res.* 2013;6(1):3.
 60. La Marca A, Nelson SM, Sighinolfi G, et al. Anti-Mullerian hormone-based prediction model for a live birth in assisted reproduction. *Reprod Biomed Online.* 2011;22(4):341–9.
 61. Nelson SM, Yates RW, Fleming R. Serum anti-Mullerian hormone and FSH: prediction of live birth and extremes of response in stimulated cycles—implications for individualization of therapy. *Hum Reprod.* 2007;22(9):2414–21.
 62. Broer SL, Dolleman M, Opmeer BC, et al. AMH and AFC as predictors of excessive response in controlled ovarian hyperstimulation: a meta-analysis. *Hum Reprod Update.* 2011;17(1):46–54.
 63. Broer SL, Mol BW, Hendriks D, et al. The role of antimullerian hormone in prediction of outcome after IVF: comparison with the antral follicle count. *Fertil Steril.* 2009;91(3):705–14.
 64. Broer SL, van Disseldorp J, Broeze KA, et al. 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.
 65. Nelson SM, Anderson RA, Broekmans FJ, Raine-Fenning N, Fleming R, La Marca A. *Hum Reprod.*

- 2012;27(3):631-6. doi: [10.1093/humrep/der446](https://doi.org/10.1093/humrep/der446). Epub 2012 Jan 11.
66. van Disseldorp J, Lambalk CB, Kwee J, et al. Comparison of inter- and intra-cycle variability of anti-Mullerian hormone and antral follicle counts. *Hum Reprod*. 2010;25(1):221-7.
 67. Majumder K, Gelbaya T, Laing I, Nardo LG. The use of anti-Müllerian hormone and antral follicle count to predict the potential of oocytes and embryos. *Eur J Obstet Gynecol Reprod Biol*. 2010;150:166-70.
 68. Scott Jr RT, Elkind-Hirsch KE, Styne-Gross A, et al. The predictive value for in vitro fertility delivery rates is greatly impacted by the method used to select the threshold between normal and elevated basal follicle-stimulating hormone. *Fertil Steril*. 2008;89(4):868-78.
 69. Seifer DB, Lambert-Messerlian G, Hogan JW, Gardiner AC, Blazar AS, Berk CA. Day 3 serum inhibin-B is predictive of assisted reproductive technologies outcome. *Fertil Steril*. 1997;67(1):110-4.
 70. Seifer DB, Scott Jr RT, Bergh PA, et al. Women with declining ovarian reserve may demonstrate a decrease in day 3 serum inhibin B before a rise in day 3 follicle-stimulating hormone. *Fertil Steril*. 1999;72(1):63-5.
 71. Eldar-Geva T, Ben-Chetrit A, Spitz IM, Rabinowitz R, Markowitz E, Mimoni T, et al. Dynamic assays of inhibin B, anti-Mullerian hormone and estradiol following FSH stimulation and ovarian ultrasonography as predictors of IVF outcome. *Hum Reprod*. 2005;20:3178-83.
 72. Fanchin R, de Ziegler D, Olivennes F, et al. Exogenous follicle stimulating hormone ovarian reserve test (EFORT): a simple and reliable screening test for detecting 'poor responders' in in-vitro fertilization. *Hum Reprod*. 1994;9(9):1607-11.
 73. Elting MW, Kwee J, Schats R, et al. The rise of estradiol and inhibin B after acute stimulation with follicle-stimulating hormone predict the follicle cohort size in women with polycystic ovary syndrome, regularly menstruating women with polycystic ovaries, and regularly menstruating women with normal ovaries. *J Clin Endocrinol Metab*. 2001;86(4):1589-95.
 74. Jayaprakasan K, Chan Y, Islam R, et al. Prediction of in vitro fertilization outcome at different antral follicle count thresholds in a prospective cohort of 1,012 women. *Fertil Steril*. 2012;98(3):657-63.
 75. Nardo LG, Gelbaya TA, Wilkinson H, et al. Circulating basal anti-Mullerian hormone levels as predictor of ovarian response in women undergoing ovarian stimulation for in vitro fertilization. *Fertil Steril*. 2009;92(5):1586-93.
 76. Ebner T, Sommergruber M, Moser M, et al. Basal level of anti-Mullerian hormone is associated with oocyte quality in stimulated cycles. *Hum Reprod*. 2006;21(8):2022-6.
 77. Sabatini L, Zosmer A, Hennessy EM, et al. Relevance of basal serum FSH to IVF outcome varies with patient age. *Reprod Biomed Online*. 2008;17(1):10-9.
 78. Tremellen K, Kolo M. Serum anti-Mullerian hormone is a useful measure of quantitative ovarian reserve but does not predict the chances of live-birth pregnancy. *Aust N Z J Obstet Gynaecol*. 2010;50(6):568-72.
 79. Gleicher N, Weghofer A, Barad DH. Anti-Mullerian hormone (AMH) defines, independent of age, low versus good live-birth chances in women with severely diminished ovarian reserve. *Fertil Steril*. 2010;94(7):2824-7.
 80. Massie JA, Burney RO, Milki AA, et al. Basal follicle-stimulating hormone as a predictor of fetal aneuploidy. *Fertil Steril*. 2008;90(6):2351-5.
 81. Sahu B, Oztutk O, Serhal P, Jayaprakasan K. Do ovarian reserve tests predict miscarriage in women undergoing assisted reproduction treatment? *Eur J Obstet Gynecol Reprod Biol*. 2010;153(2):181-4.
 82. Katz-Jaffe MG, Surrey ES, Minjarez DA, et al. Association of abnormal ovarian reserve parameters with a higher incidence of aneuploid blastocysts. *Obstet Gynecol*. 2013;121(1):71-7.
 83. Pigny P, Jonard S, Robert Y, et al. Serum anti-Mullerian hormone as a surrogate for antral follicle count for definition of the polycystic ovary syndrome. *J Clin Endocrinol Metab*. 2006;91(3):941-5.
 84. Dewailly D, Gronier H, Poncelet E, et al. Diagnosis of polycystic ovary syndrome (PCOS): revisiting the threshold values of follicle count on ultrasound and of the serum AMH level for the definition of polycystic ovaries. *Hum Reprod*. 2011;26(11):3123-9.
 85. Freeman EW, Sammel MD, Lin H, et al. Contribution of the rate of change of antimullerian hormone in estimating time to menopause for late reproductive-age women. *Fertil Steril*. 2012;98(5):1254-9.
 86. Sowers MR, Zheng H, McConnell D, et al. Follicle stimulating hormone and its rate of change in defining menopause transition stages. *J Clin Endocrinol Metab*. 2008;93(10):3958-64.
 87. Tehrani FR, Solaymani-Dodaran M, Tohidi M, et al. Modeling age at menopause using serum concentration of anti-Mullerian hormone. *J Clin Endocrinol Metab*. 2013;98:729-35.
 88. Schuh-Huerta SM, Johnson NA, Rosen MP, et al. Genetic variants and environmental factors associated with hormonal markers of ovarian reserve in Caucasian and African American women. *Hum Reprod*. 2012;27(2):594-608.
 89. The Practice Committees of the American Society for Reproductive Medicine, The Society for Assisted Reproductive Technology. Mature oocyte cryopreservation: a guideline. *Fertil Steril*. 2013;99(1):37-43.
 90. Hill KA, Nadler T, Mandel R, et al. Experience of young women diagnosed with breast cancer who undergo fertility preservation consultation. *Clin Breast Cancer*. 2012;12(2):127-32.
 91. Fanchin R, Schonauer LM, Righini C, et al. Serum anti-Mullerian hormone dynamics during controlled ovarian hyperstimulation. *Hum Reprod*. 2003;18(2):328-32.
 92. Li HW, Wong CY, Yeung WS, et al. Serum anti-mullerian hormone level is not altered in women using

- hormonal contraceptives. *Contraception*. 2011;83(6): 582–5.
93. Deb S, Campbell BK, Pincott-Allen C, Clewes JS, Cumberpatch G, Raine-Fenning NJ. Quantifying effect of combined oral contraceptive pill on functional ovarian reserve as measured by serum anti-Müllerian hormone and small antral follicle count using three-dimensional ultrasound. *Ultrasound Obstet Gynecol*. 2012;39(5):574–80.
 94. Fabregues F, Castelo-Branco C, Carmona F, et al. The effect of different hormone therapies on anti-mullerian hormone serum levels in anovulatory women of reproductive age. *Gynecol Endocrinol*. 2011;27(4):216–24.
 95. van den Berg MH, van Dulmen-den BE, Overbeek A, et al. 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(6):1520–7.
 96. Kallio S, Puurunen J, Ruokonen A, Vaskivuo T, Piltonen T, Tapanainen JS. *Fertil Steril*. 2012. doi: [pii: S0015-0282\(12\)02446-6](https://doi.org/10.1016/j.fertnstert.2012.02.046).
 97. Deb S, Campbell BK, Clewes JS, Pincott-Allen C, Raine-Fenning NJ. The intra-cycle variation in the number of antral follicles stratified by size and in the endocrine markers of ovarian reserve in women with normal ovulatory menstrual cycles. *Ultrasound Obstet Gynecol*. 2012;41:216–22.
 98. 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;23:411–20.
 99. Rustamov O, Smith A, Roberts SA, et al. Anti-Mullerian hormone: poor assay reproducibility in a large cohort of subjects suggests sample instability. *Hum Reprod*. 2012;27(10):3085–91.
 100. Streuli I, Fraisse T, Chapron C, et al. Clinical uses of anti-Mullerian hormone assays: pitfalls and promises. *Fertil Steril*. 2009;91(1):226–30.