Predicting Reproductive Age with Biomarkers of Ovarian Reserve—How (and What) Are We Measuring?

Karl R. Hansen, MD, PhD

Section of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma

Address for correspondence Karl R. Hansen, MD, PhD, Section of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, University of Oklahoma Health Sciences Center, P.O. Box 26901, WP 2410, Oklahoma City, OK 73126 (e-mail: karl-hansen@ouhsc.edu).

Abstract

Predicting the reproductive lifespan of an individual woman remains an elusive, yet clinically important, goal. The development of models and staging systems that accurately determine the end of natural fertility and the anticipated age of menopause will represent a significant advance in our ability to counsel women regarding family planning issues and in the individualization of risk assessment. Recent histological and longitudinal investigations have demonstrated a significant relationship between commonly used clinical markers of ovarian reserve and the true ovarian reserve, as assessed by the ovarian nongrowing follicle count and the age of menopause, respectively. Models and staging systems that have been developed based on these findings represent important advances in the field of women’s health and promise to provide additional insights into the process of reproductive aging in general. Although the models developed to date appear to improve the prediction of the age of menopause as compared with chronological age alone, wide confidence intervals in the predicted age of menopause and less accurate predictions at ages remote from menopause limit their clinical utility for the individual woman. Future longitudinal and histological investigations are necessary to improve the accuracy of models of reproductive aging.

Keywords
- ovarian reserve
- nongrowing follicles
- primordial follicles
- anti-Müllerian hormone
- antral follicle count

Reproductive aging in women refers to the progressive loss of fertility and ovarian endocrine function which ultimately results in the menopause. Although the process is certainly not new, events over the last few decades have brought renewed interest in developing a better understanding of the reproductive aging process and its clinical prediction. First, socioeconomic and demographic trends have resulted in many women delaying childbearing until their late 30s and early 40s, when natural fertility is compromised relative to younger reproductive-aged women.1–3 Educational campaigns have increased the awareness of the general public to this phenomenon and have fueled a desire to better predict the end of natural fertility.4 Second, a broader understanding has emerged regarding the relationship between an early age of menopause and an increased risk for medical problems including colorectal cancer, osteoporosis, and cardiovascular and urogenital disease.5–8 Conversely, a delayed age of menopause may be associated with an increased risk of cancers of the breast, endometrium, and ovary.9,10 Finally, large numbers of women are passing through the menopausal transition as the “baby boom” population ages, with the attendant symptoms that may significantly impact quality of life. In sheer population numbers, the prediction of the reproductive lifespan in now compelling—a relatively modern occurrence given that only in the last approximately 100 years has the average woman lived to age 50 and actually experienced the menopause.11 Thus, understanding and predicting the reproductive aging process is much more than an academic exercise—it has real-world applications and implications.
The average age of menopause is 51 years; however, there is a wide variation between individuals, with some women entering menopause by the age of 40 and others having menstrual cycles into their mid to late 50s. Unfortunately, the progressive loss of ovarian function and fertility is largely a silent process. Changes in menstrual cycle characteristics such as shortening of the menstrual cycle length are subtle initially. By the time a woman experiences oligomenorrhea associated with elevated gonadotropins, she is near the end of the reproductive aging process and her fertility is severely compromised. In reality though, her fertility was compromised many years earlier. With the understanding that chronological age is a crude predictor of the reproductive aging process, considerable interest has been directed toward identifying biomarkers as well as developing models and staging systems that might better characterize and predict the reproductive lifespan. Such models would have many potential applications. It has been hypothesized that the interval of time between the end of natural fertility and menopause is fixed, regardless of the exact age at which menopause occurs. Therefore, an accurate model that forecasts the age of menopause may also predict the end of natural fertility. Beyond family planning, such a model may also be useful for women with a history of prior exposure to chemotherapy, radiation, or ovarian surgery. This article reviews our current understanding of the reproductive aging process and the utility of common clinical tests of ovarian reserve in assessing reproductive age.

**The Biology of Reproductive Aging: What Are We Attempting to Measure?**

Although neuroendocrine changes have been described late in the reproductive aging process, there is broad consensus that the primary mechanism behind reproductive aging is the progressive loss of microscopic ovarian follicles. Therefore, modeling or predicting the reproductive lifespan is, in essence, an attempt to predict the number of remaining ovarian follicles in a given woman and the point at which this follicular reserve will be exhausted. The ovarian follicle is the functional/anatomical unit of the ovary, consisting of a primary oocyte surrounded by granulosa cells. The ovarian primordial follicle (PF) consists of a primary oocyte surrounded by a single layer of flattened granulosa cells. Intermediate follicles consist of a primary oocyte surrounded by some flattened and some cuboidal granulosa cells, and primary follicles consist of a primary oocyte surrounded by an entirely cuboidal layer of granulosa cells. Collectively, this group of the three follicle types is referred to as the pool of resting or nongrowing follicles (NGFs). Some investigators consider the PF pool to represent the ovarian reserve, whereas others believe the reserve includes all NGFs. Differences of opinion notwithstanding, it is clear that the true ovarian reserve is represented by one or both of these groups, with investigations utilizing PF counts or total NGF counts as the outcome of interest reaching similar conclusions.

Follicles are selected or recruited from this pool of NGFs, progressively moving through the preantral, early antral, antral, and ultimately the Graafian follicle stage. The earliest stages of this process are gonadotropin independent, with the vast majority of follicles undergoing atresia rather than ovulation. Beyond the quantitative loss of NGFs associated with reproductive aging, a considerable loss in quality of the remaining oocytes also occurs during this process, as demonstrated by an increased rate of aneuploid oocytes and derived embryos in the setting of the assisted reproductive technologies (ART). Given that the reproductive lifespan is dictated by the initial size of the ovarian follicle pool in a given woman and its rate of loss, considerable efforts have been put forth to develop a better understanding of this process.

**Modeling the Depletion of NGFs: Histological Investigations**

Direct measures of ovarian PFs and total NGFs in women are relatively rare due to the challenges associated with the execution of these types of investigations. Appropriate tissue for histological examination is exceedingly difficult to obtain, and the processing of ovarian tissue and the counting of ovarian follicles is demanding and time-intensive. Older model-based techniques involved embedding the ovaries in paraffin followed by serial sectioning and counting between 1 in 10 and 1 in 200 sections. The number of follicles counted was then multiplied by the inverse of the sampling fraction to obtain a raw estimate of the total NGF number within a given ovary. Raw counts were then frequently multiplied by a correction factor to obtain a “corrected” follicle count. The application of these correction factors (applied to counts of some, but not all ovaries in many investigations) attempts to correct for the observation that small particles (follicles) tend to be underrepresented in tissue sections and undercounted, whereas larger particles (follicles at more advanced stages) tend to be over counted.

![Figure 1](image-url) Distribution curves for observed age at last birth (proxy variable for natural sterility, left line) and age at menopause (right line). Reprinted with permission from Lambalk et al.15
due to their appearance in multiple thin tissue sections. Whether the incorporation of these correction factors into estimates of ovarian follicle counts improves their accuracy cannot be ascertained, as the model cannot be tested by the model itself. As a result of these challenges, to our knowledge, only three investigations have ever estimated NGF counts for more than forty ovaries.\(^{21,22,24}\) Only one of these investigations has utilized modern stereology techniques which do not require the incorporation of correction factors.\(^{24}\) These modern stereology methods, combining the fractionator and optical dissector tools, have become the standard in structural analysis.\(^{25,26}\)

Attempts to model ovarian NGF decay have frequently incorporated follicle counts from multiple studies. Limitations of this approach include the systematic bias introduced with the variable use of correction factors as described above, the lack of appropriate interobserver validation studies, and the combination of NGF count estimates from investigations that used similar, but not identical counting techniques. Earlier investigations using these methodologies suggested that the decline in the log-transformed ovarian NGF count exhibited a biphasic exponential decline, with approximately 1 to 2 million follicles at birth and a break point in the mid-30s at approximately 25,000 NGFs when the decay curve suddenly developed a more negative slope.\(^{23}\) A more recent investigation which also combined NGF counts from multiple studies suggested that the change in NGF counts from conception to menopause was best characterized by a five-parameter asymmetric double Gaussian cumulative curve.\(^{27}\) As described above, models derived from the combined data from multiple studies must be interpreted with caution. In essence, they are the histological equivalent of determining the U.S. census using different methods in different regions and combining the results. In spite of the questionable validity of this approach, it remains common due to the scarcity of histological investigations.

In contrast to models derived from combined data, the largest study (\(n = 122\)) from a single group of investigators utilizing modern morphometric techniques has suggested that the decline in the log-transformed ovarian NGF count from birth to menopause is best described by a simple power function (Fig. 2):\(^{24}\)

\[
\log(\text{NGF count}) = (-0.00019) \times \text{(age in years)}^{2.452} + 5.717
\]

The power model is a robust fit to the observed data \((R^2 = 0.84)\), and suggests that there is no sudden change in the rate of ovarian NGF loss associated with aging, but rather a smooth change. Additionally, the model predicts, on average, approximately 520,000 NGFs present in both ovaries at birth and approximately 750 follicles remaining at the time of menopause.\(^{28}\) An independent dataset developed with modern stereology techniques is necessary to validate the power model of ovarian NGF decline associated with aging. Although log-transformation of ovarian NGF counts is necessary from a modeling standpoint, this approach can easily obscure the tremendous loss of NGFs associated with aging: a loss readily appreciated when reviewing the non-log transformed data (Fig. 3). By the age of 33 on average, approximately 90% of the ovarian NGFs are depleted.\(^{24}\)

**Figure 2** Power-model of ovarian nongrowing follicle (NGF) decay. The log of the ovarian NGF number is plotted versus age (years). The solid line indicates the fitted model with dashed lines representing the 95% confidence interval \((n = 122)\). Reprinted with permission from Hansen et al.\(^{24}\)

**Figure 3** Histogram representation of the average nongrowing follicle (NGF) count for subjects in different age groups \((n = 122)\). The height of the bar represents the average NGF count for the given age group \(\pm\) SEM. Adapted with permission from Hansen et al.\(^{24}\)

### “Ovarian Reserve” Tests—What are We Measuring?

The concept of “ovarian reserve” testing is deeply rooted in the ARTs. With the understanding that chronological age is a relatively poor predictor of how an individual woman might respond to gonadotropin stimulation in the setting of in vitro fertilization (IVF), multiple tests have been developed to predict the ovarian response. Although the ideal outcome of interest would be live-births in these investigations, most have used surrogate markers, including the number of oocytes retrieved or the peak estradiol level following gonadotropin stimulation. Commonly used clinical markers of ovarian reserve include basal serum levels of follicle-stimulating hormone (FSH), basal estradiol, basal inhibin B, the
ovarian antral follicle count (AFC) as determined by transvaginal ultrasound examination, and serum levels of anti-Müllerian hormone (AMH).28–41 Other less frequently utilized tests include the clomiphene citrate challenge test and the gonadotropin-releasing hormone (GnRH) agonist challenge test.42–45

Although all of these tests have some predictive value with regards to the outcome of gonadotropin stimulation from a quantitative standpoint (e.g., number of oocytes obtained following treatment), they appear to have much less power at predicting pregnancy and live-births.46 Most investigations have suggested that serum levels of AMH and the ovarian AFC have the most predictive power from a quantitative standpoint, and a recent meta-analysis has suggested there is no benefit to obtaining multiple markers of ovarian reserve rather than a single marker to predict responsiveness.46,47 Few investigations have incorporated chronological age as a variable to improve the predictive value of a given biomarker, although the construction of efficiency curves, which include the impact of chronological age and ovarian reserve markers, may be useful in the identification of threshold levels of a given marker beyond which pregnancy is unlikely.48,49

While these clinical tests are commonly referred to as “ovarian reserve” tests, they are more accurately measures of ovarian responsiveness. In other words, their utility is to predict response in the artificial setting of IVF. Conversely, the ovarian reserve, in its truest sense, is referring to the size of the remaining ovarian NGF pool. This distinction between ovarian responsiveness and the true ovarian reserve is critical. Both are clinically relevant and are definitely related. However, the pool of recruitable or selectable follicles (i.e., the ovarian responsiveness) at any given point in time may be influenced by factors other than the size of the ovarian NGF pool, such as the dose of gonadotropins utilized in ART, the specific stimulation protocol, and likely many other unknown factors. This distinction may help explain why commonly used clinical markers of “ovarian reserve” may outperform age in the clinical prediction of ovarian responsiveness in the setting of IVF whereas chronological age is the best single predictor of the true ovarian reserve (i.e., the ovarian NGF pool).29,37,39,50

With the understanding of the distinction between ovarian responsiveness and the true ovarian reserve, what is the evidence that these clinical biomarkers are actually reflective of the ovarian NGF pool, and thus, the reproductive lifespan? Multiple lines of evidence have supported a correlation between ovarian reserve markers and the true ovarian reserve. However, only recently have histological and longitudinal investigations confirmed this relationship.

Examining the Circumstantial Evidence

**Follicle-Stimulating Hormone**

FSH is a dimeric glycoprotein produced by the anterior pituitary known to regulate the recruitment and growth of ovarian follicles from the antral stage to the Graafian follicle. FSH has been studied extensively as a marker of ovarian responsiveness, with some investigations suggesting it outperforms chronological age in the prediction of ART outcome from a quantitative standpoint in univariate analyses.29,37,39 Multiple cross-sectional investigations have demonstrated an increase in serum levels of FSH associated with increasing chronological age.51,52 Although these increases can be demonstrated across the menstrual cycle, they are most prominent in the early follicular phase.53 Longitudinal investigations have confirmed these findings, with variable increases in FSH being associated with subtle shortening of the follicular phase of the menstrual cycle, and more dramatic increases associated with menstrual cycle changes characteristic of the late perimenopause (Stages of Reproductive Aging Workshop [STRAW] stage-1) (Fig. 4).54–59 Gradual increases in FSH are observed within 10 years of menopause.59 These observations strongly suggest that serum levels of FSH are correlated with reproductive age, but the relatively late onset of these changes relative to the reproductive aging process and the cycle-specific nature of its measurement highlight its limitations as a marker of true ovarian reserve.

**Inhibin B**

Inhibin B is a dimeric polypeptide produced by the granulosa and theca cells of the developing cohort of antral follicles, whereas inhibin A is primarily a product of the developing dominant follicle and the corpus luteum.60 Both are known to exert a negative feedback at the level of the hypothalamus and pituitary, ultimately resulting in the decrease in FSH secretion.61,62 Like FSH, the inhibins have been evaluated extensively as a marker of ovarian responsiveness, although they have never enjoyed the popularity of FSH due to a variety of reasons. Early assays for the inhibins were not as specific for inhibin A versus B, and levels of inhibin B associated with decreased responsiveness in the setting of ART were near the limits of detection of the early assays.63,64 Additionally, some investigations suggested its predictive value was no greater than that of FSH.46,64,65 Inhibin B levels are also cycle phase-dependent, with the highest levels observed in the early follicular phase and mid-cycle. Cross sectional investigations have demonstrated lower early follicular phase levels of inhibin B in older as opposed to younger reproductive-aged women.53 These findings have been confirmed in longitudinal investigations.59,66 Inhibin B concentrations fall below the limits of detection within 4 to 5 years of the menopause, and significant decreases are variably detected by the late reproductive years (STRAW stages-3a).55–59 Although these findings would suggest that inhibin B levels may reflect the size of the remaining pool of NGFs, some investigations have suggested the decline in its level is not gradual, which would limit its value as a biomarker of true ovarian reserve.67

**The Ovarian Antral Follicle Count**

The ovarian AFC as determined by transvaginal ultrasound examination has been widely evaluated as a marker of ovarian responsiveness.34–38 The AFC is the number of antral follicles between 2 and 10 mm in size within both ovaries observed on transvaginal ultrasound examination, and is reflective of the pool of potentially recruitable follicles following...
gonadotropin stimulation. The AFC is highly predictive of the response of an individual patient to gonadotropins in the setting of IVF, and along with serum levels of AMH, is one of the best single predictors of stimulation outcome from a quantitative perspective. \cite{37,39,46} Multiple cross-sectional investigations have demonstrated a decrease in the ovarian AFC associated with increasing age. \cite{68,69} In addition to the age-related decrease in the ovarian AFC, the AFC is also noted to decrease with progressive changes in menstrual cycle characteristics associated with the menopausal transition. \cite{70,71} The gradual decline in the ovarian AFC associated with aging and its cycle-phase independence makes it a more attractive marker of reproductive aging as compared with serum levels of FSH and inhibin B. \cite{69}

**Anti-Müllerian Hormone**

AMH is produced by the granulosa cells of preantral and early antral follicles, and serum levels fluctuate minimally throughout the menstrual cycle. \cite{72,73} Because it is secreted by follicles at the transition point between resting and growing follicles, it is one of the best single markers of ovarian responsiveness. Many reports in the literature indicate serum levels of AMH outperform all other ovarian reserve tests in the prediction of ART outcome with the possible exception of the AFC, the two of which are highly correlated. \cite{39,41,46,74} Multiple cross-sectional and longitudinal investigations have demonstrated a gradual decline in serum AMH level associated with increasing chronological age. \cite{39,59,75-77} As with the ovarian AFC, AMH is a prime candidate as a marker for reproductive age due to its gradual decline with advancing age, although some investigations have suggested this gradual decline does not begin until the age of approximately 25 to 30 years. \cite{78,79} Additionally, these decreases have been noted in longitudinal investigations wherein significant changes were not observed in other ovarian reserve markers. \cite{80} These observations suggest that AMH may be a useful marker in younger (at least at age 30 and older) reproductive-aged women. Progressive decreases in serum levels of AMH are also noted with advancing stage of the STRAW staging system, with levels below the limits of detection reached within 5 years of menopause. \cite{59,71,75} Although AMH appears promising as a measure of the reproductive lifespan, the lack of standardized assays and sufficient assay sensitivity remain problematic. Furthermore, since serum AMH levels appear to be reflective of the early growing follicular pool, pathological states associated with abnormal follicular development (e.g., hypogonadotropic hypogonadism) may result in conditions in which the relationship between serum AMH and the NGF pool may be altered. \cite{81}

**The Association between Early Menopause and Poor Stimulation in the Setting of IVF**

Since commonly used clinical test of ovarian reserve are highly predictive of ovarian responsiveness in the setting of ART, an earlier menopause in women with a poor response to gonadotropins would serve as indirect evidence of an association between ovarian reserve tests results and the reproductive lifespan. Several case control and retrospective cohort studies have supported this hypothesis. \cite{82-84} The largest of these investigations, which included 4601 subjects with a median follow-up of 5.5 years, suggested that the adjusted odds ratio for having entered the menopausal transition or natural menopause for women with a poor response to stimulation as compared with normal responders was 3.1.
(95% confidence interval [CI] 2.4–3.8).82 A second smaller investigation with similar follow-up times by Lawson et al reached similar conclusions.83 Although these studies suggest ovarian reserve markers are related to the reproductive life-span, this evidence must be considered circumstantial and preliminary.

Examining the Evidence: Histological Investigations

Investigations into the relationship between commonly used clinical markers of ovarian reserve and the actual ovarian PF or NGF pool are exceedingly rare. This scarcity is partially explained by the reasons outlined above for histological studies in general. However, an additional challenge faced in these investigations is obtaining ovarian reserve markers for study participants prior to histological examination of the ovary. In other words, tissue obtained from autopsy or organ donation, while appropriate for some investigations, is unsuitable to address the question at hand. Only women undergoing elective oophorectomy are appropriate participants, and ovarian reserve assessments must be performed shortly prior to surgery.

Only two prior studies have directly addressed the relationship between ovarian reserve markers and ovarian PF count in women. The first investigation did not identify a significant relationship between the GnRH agonist stimulation test, the clomiphene citrate challenge test or basal FSH levels and the number of ovarian follicles per cubic centimeter.85 Limitations of this investigation included its small sample size (n = 22) and the use of older counting techniques. Furthermore, because the ovarian AFC and serum levels of AMH were not in use at the time as ovarian reserve tests, the two best markers were not evaluated.

The only other investigation of the relationship between ovarian reserve markers and the ovarian PF and NGF counts demonstrated a significant correlation between follicle counts and serum levels of AMH, inhibin B, FSH and the ovarian AFC (Table 1).50 Partial correlations controlling for chronological age, the marker most strongly correlated with the ovarian PF count, demonstrated that the ovarian AFC and serum levels of AMH add predictive power in the assessment of the ovarian PF count beyond the contributions of age alone (Table 1). Although the AFC slightly outperformed AMH in this investigation, the difference between the two markers was small.50

Characteristic menstrual cycle changes associated with the menopausal transition process can also be considered biomarkers of ovarian reserve. Cross-sectional and longitudinal investigations have characterized these changes.86–90 Most commonly, a woman initially experiences a shortening in menstrual cycle length with advancing age due to a shortening of the follicular phase.86 Carefully designed studies have demonstrated that this shortening of the follicular phase is due to the advanced selection of the dominant follicle in older as compared with younger reproductive-aged women.91 In the latter transition, intervals of oligomenorrhea and amenorrhea occur due to anovulation. The STRAW staging system is the most widely utilized staging system of the reproductive aging process, adding significantly to the characterization of this process.55–58,92 Recent histological investigations have validated STRAW stage 4 through stage 1 as defined by these characteristic menstrual cycle changes, with significant decreases in the ovarian PF count noted with advancing STRAW stage (Fig. 5).71 However, as with biomarkers of ovarian reserve, there is considerable overlap in the ovarian PF count between the stages, and the PF count alone is inadequate to assign a given individual to a single stage.

While histological investigations have provided important insights into the process of reproductive aging and the relationship between biomarkers of ovarian reserve and the true ovarian reserve, it is important to have a full understanding of their strengths and limitations. Histological investigations measure the actual outcome of interest, the ovarian PF and NGF pool, rather than relying on surrogate markers of reproductive age. Although the menopause is certainly an outcome of interest, the lack of vaginal bleeding in a late reproductive-aged woman may be due to factors other than ovarian failure, such as hyperprolactinemia or thyroid disorders. Conversely, vaginal bleeding may occur in the absence of significant ovarian follicular activity due to pathology, or due to the peripheral conversion of androgens to estrogens, a problem more prevalent in our increasingly obese society. Limitations include the cross-sectional nature of all histological studies. Additionally, participants in these investigations may not be considered “normal,” as they all have some

Table 1

Univariate and partial correlations of endocrine parameters, ultrasound determined AFC and log 10 primordial follicle count (n = 42)

<table>
<thead>
<tr>
<th></th>
<th>Step 1: Adjusting for age</th>
<th>Step 2: Adjusting for age</th>
<th>Step 3: Adjusting for age + AFC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>p-Value</td>
<td>R</td>
</tr>
<tr>
<td>Age</td>
<td>−0.80</td>
<td>&lt; 0.0001</td>
<td>AMH</td>
</tr>
<tr>
<td>AMH</td>
<td>0.72</td>
<td>&lt; 0.0001</td>
<td>Inhibin B</td>
</tr>
<tr>
<td>Inhibin B</td>
<td>0.40</td>
<td>0.0100</td>
<td>FSH</td>
</tr>
<tr>
<td>FSH</td>
<td>−0.32</td>
<td>0.0402</td>
<td>Estradiol</td>
</tr>
<tr>
<td>Estradiol</td>
<td>0.12</td>
<td>0.4575</td>
<td>AFC</td>
</tr>
<tr>
<td>AFC</td>
<td>0.78</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: AFC, antral follicle count; AMH, anti-Müllerian hormone; FSH, follicle-stimulating hormone. Reprinted with permission from Hansen et al.50
indication for a surgical procedure. Therefore, it is uncertain if the findings can be more broadly applied to women without such indications.

Finally, it is important to note that the histological assessment of ovarian reserve is a research tool only. Because of the irregular distribution of ovarian primordial and NGFs within the ovarian cortex, ovarian biopsies are not an adequate technique for the assessment of ovarian reserve, largely intact ovaries are required. To state the obvious—it is not practical to remove the ovary(ies) of a women to determine the status of her ovarian reserve. Ultimately the development of clinically useful tools and models depends upon the longitudinal investigation of readily available clinical markers utilizing the outcome of the final menstrual period.

**Examining the Evidence: Longitudinal Investigations**

Several recent investigations with longitudinal follow-up have evaluated the relationship between clinical markers of ovarian reserve and the occurrence of menopause. Broekmans et al initially demonstrated a significant agreement between the predicted and observed age at menopause in a model developed using the ovarian AFC in a cohort of 163 regularly cycling women. However, the predictive value of the model was limited except in the setting of low AFC for chronological age. Additionally, follow-up studies by the same group of investigators have demonstrated that the AFC was only predictive of the age of menopause in a univariate analysis; with chronological age and AMH the only significant predictors in a multivariate analysis. Sowers et al, following a cohort of 50 women through the menopause, demonstrated that baseline AMH levels were highly associated with the age of menopause, with serum levels falling below the limits of detection within 5 years of the final menstrual period. Although inhibin B levels were also low or undetectable within 5 years of menopause, they were less predictive of the final menstrual period than was AMH.

More recent and larger investigations have confirmed the relationship between serum AMH concentrations and the reproductive life-span. Tehrani et al followed a group of 266 fertile women aged between 20 and 50 for an average of 6 years. In this time frame, 63 women experienced their final menstrual period. This investigation demonstrated a reasonable agreement between the observed and predicted age of menopause based on a single AMH measurement, with lower accuracy of the prediction noted at the extremes of age of anticipated menopause. An investigation of a larger cohort of fertile women (1,015 women) with 277 occurrences of menopause by the same group confirmed these findings utilizing a model that included both age and serum AMH as predictors of the final menstrual period. AMH and chronological age were also significant predictors of the age of menopause in a recent investigation of 401 women followed through the menopausal transition in the Penn Ovarian Aging Study. A more recent study by the same group of investigators has suggested that the rate of change of AMH, in addition to chronological age, may improve the precision of

---

**Figure 5** Box plots of log10-transformed ovarian primordial follicle counts, total ovarian antral follicle counts, and biomarkers of ovarian reserve for STRAW stage 4 through stage 1. (A) Log 10 primordial follicle count: \( p < 0.0001 \) between all stages except \(-2/1\), where \( p = 0.0074 \); (B) FSH: \( p = 0.0036 \) between stages \(-4/-1\), all others are not significantly different; (C) antral follicle count: \( p < 0.0001 \) between all stages except \(-3/-2\), \(-3/-1\), and \(-2/-1\) which are not significantly different; (D) AMH: \( p < 0.0001 \) between all stages except \(-3/-2\), \(-3/-1\), and \(-2/-1\) which are not significantly different. Adapted with permission from Hansen et al. AMH, anti-Müllerian hormone; FSH, follicle-stimulating hormone; STRAW, Stages of Reproductive Aging Workshop.
estimates of the final menstrual period as compared with models utilizing a single AMH measurement. Similar to the Tehrani et al investigations, less precision in the estimate of the final menstrual period was observed at more remote intervals from the menopause.

Although these clinical investigations have provided powerful evidence of the relationship between ovarian reserve biomarkers (particularly AMH) and the true ovarian reserve, limitations of their clinical applicability should be noted. Currently there is no international standard for AMH, and the commercially available assays lack the adequate sensitivity to be useful near the end of the reproductive lifespan. As discussed by the authors of many of these studies, the clinical prediction of the age of menopause tends to be more accurate at times closer to the menopause, and in some cases, at more advanced chronological age as compared with younger reproductive-aged women. For example, at a similar AMH level for a given woman (< 0.20 ng/ml), the 95% CI for the time to menopause is 4.20 to 6.33 years (median 5.99 years) if she is 45 to 48 years old, whereas it is 3.31 to 12.73 years (median 9.94 years) if she is 35 to 39 years old. Beyond the wide confidence intervals associated with some model-based predictions, it is important to note that even at the same AMH level, the time, on average, until the final menstrual period is significantly influenced by chronological age itself. In essence, the addition of AMH measurements to the derived models adds to the predictive power of age in the assessment of the reproductive lifespan, but it does not replace age. These observations are entirely consistent with the power model of ovarian NGF depletion, wherein 84% of the decline in NGF counts associated with increasing age can be explained by age alone. Strikingly, Tehrani et al. reached similar conclusions, with age alone having an adequacy of 84% to predict age at menopause correctly, increasing to 92% when AMH was added to their model. Although histological investigations have demonstrated that clinical ovarian reserve tests add to the predictive power of chronological age in partial correlation analyses, whether or not they can improve the predictive value of the power model of NGF decay awaits further histological studies.

In addition to the above considerations, it is also important to note the population of women included in these longitudinal investigations. Although “normal” women may be an ideal population in which to study the reproductive aging process, it is unclear if the resulting models can be generalized to those commonly seen in clinical practice (e.g., infertility patients). Finally, it is also important to have an understanding of when these models may be more likely to give less accurate predictions. As described above, AMH levels may only experience a relatively predictable decline beginning at approximately 25 to 30 years of age. Therefore, applying the above models to women in younger age groups may be problematic. Similar challenges may also be encountered in pathological states which may affect the relationship between the pool of growing and resting ovarian follicles. Given that previous histological investigations have suggested that a larger percentage of follicles are in the growing stages in older as compared with younger reproductive-aged women, some changes in these relationships should also be anticipated to naturally occur at the extremes of the reproductive lifespan.

Conclusions

Predicting the age of the final menstrual period for the individual woman remains an important goal for clinicians and patients alike. In addition to risk assessment, the prediction of the age of menopause may well predict the age of subfertility and the end of natural fertility. Ideally the prediction of the reproductive lifespan would occur at a younger age, allowing for effective interventions to be undertaken. Recent histological and longitudinal investigations have demonstrated a significant relationship between biomarkers of ovarian reserve and the true ovarian reserve, as assessed by the ovarian PF count and the final menstrual period, respectively. Models predicting the age of menopause incorporating ovarian reserve markers, particularly serum AMH levels, have shown promise in the prediction of the reproductive lifespan. However, the relatively wide confidence intervals associated with the predictions generated by these models, particularly at the extremes of the age of anticipated menopause, currently limit their clinical utility. Further limiting the application of these models is the lack of international standards for the AMH assay and its low sensitivity. In summary, predicting the reproductive lifespan remains a goal that we have not yet reached. Future longitudinal and histological investigations, including those aimed at identifying new biomarkers, are necessary to improve the accuracy of models of reproductive aging.

Acknowledgments

The author would like to thank Dr. Michael Soules and Dr. Nancy Klein for their critical review of the manuscript. Additionally, we would like to acknowledge the NIH and the Oklahoma Center for the Advancement of Science and Technology for funding much of the work presented in this review.

References

3. Matthews TJ, Hamilton BE. Delayed childbearing: more women are having their first child later in life. NCHS Data Brief 2009; 21(21):1–8
5. van Wayenburg CA, van der Schouw YT, van Noord PA, Peeters PH. Age at menopause, body mass index, and the risk of colorectal cancer mortality in the Dutch Diagnostisch Onderzoek Mammacarcinoom (DOM) cohort. Epidemiology 2000; 11(3):304–308
van Der Voort DJ, van Der Weijer PH, Barentsen R. Early menopause: increased fracture risk at older age. Osteoporos Int 2003;14(6):525–530


Block E. Quantitative morphological investigations of the follicular system in women; variations at different ages. Acta Anat (Basel) 1952;14(1–2, Suppl 16):108–123


Hall JE, Welt CK, Cramer DW. Inhibin A and inhibin B reflect ovarian function in assisted reproduction but are less useful at predicting outcome. Hum Reprod 1999;14(2):409–415


Frattarelli JL, Lauria-Costab DF, Miller BT, Bergh PA, Scott RT. Basal antral follicle number and mean ovarian diameter predict cycle cancellation and ovarian responsiveness in assisted reproductive technology cycles. Fertil Steril 2000;74(3):512–517


Verhagen TE, Hendriks DJ, Bancsi LF, Mol BW, Broekmans FJ. The accuracy of multivariate models predicting ovarian reserve and
72 Hehenkamp WJ, Looman CW, Themmen AP, de Jong FH, te Velde ER, Broekmans FJ. Anti-Müllerian hormone levels in the spontaneous menstrual cycle do not show substantial fluctuation. J Clin Endocrinol Metab 2006;91(10):4057–4063
75 Hale GE, Zhao X, Hughes CL, Burger HG, Robertson DM, Fraser IS. Endocrine features of menstrual cycles in middle and late reproductive age and the menopausal transition classified according to the Staging of Reproductive Aging Workshop (STRAW) staging system. J Clin Endocrinol Metab 2007;92(8):3060–3067
77 Seifer DB, Baker VL, Leader B. Age-specific serum anti-Müllerian hormone values for 17,120 women presenting to fertility centers within the United States. Fertil Steril 2011;95(2):747–750
81 Tran ND, Cedars MI, Rosen MP. The role of anti-müllerian hormone (AMH) in assessing ovarian reserve. J Clin Endocrinol Metab 2011;96(12):3609–3614


Taffe JR, Dennerstein L. Menstrual patterns leading to the final menstrual period. Menopause 2002;9(1):32–40


Taffe JR, Cain KC, Mitchell ES, Woods NF, Crawford SL, Harlow SD. “Persistence” improves the 60-day amenorrhea marker of entry to late-stage menopausal transition for women aged 40 to 44 years. Menopause 2010;17(1):191–193

Klein NA, Harper AJ, Houmard BS, Sluss PM, Soules MR. Is the short follicular phase in older women secondary to advanced or accelerated dominant follicle development? J Clin Endocrinol Metab 2002;87(12):5746–5750


van Rooij IA, Tonkelaar Id, Broekmans FJ, et al. Anti-müllerian hormone is a promising predictor for the occurrence of the menopausal transition. Menopause 2004;11(6, Pt 1):601–606


