Ovarian Primordial and Non-Growing Follicle Counts According to the Stages of Reproductive Aging Workshop (STRAW) Staging System

Karl R. Hansen, M.D., Ph.D.[1], **LaTasha B. Craig, M.D.**[1], **Michael T. Zavy, Ph.D.**[1], **Nancy A. Klein, M.D.**[2], and **Michael R. Soules, M.D.**[2]

[1]Section of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73190

[2]Seattle Reproductive Medicine, Seattle, WA 98109.

Abstract

Objective—To characterize the ovarian primordial and non-growing follicle number according to the Stages of Reproductive Aging Workshop (STRAW) staging system as defined by menstrual cycle characteristics.

Methods—Normal ovaries were collected from 63 women (age 26-52 years) undergoing oophorectomy for benign indications. Prior to surgery, each participant completed a detailed questionnaire collecting information regarding menstrual cycle characteristics and were classified by bleeding patterns into STRAW stages -4, -3, -2, and -1. A single ovary was selected for determination of the ovarian primordial and total non-growing follicle number utilizing a validated fractionator/optical disector method. A subset of the participants (n = 43) underwent transvaginal ultrasound examination for the determination of the ovarian antral follicle count and serum measurements of FSH, estradiol, anti-müllerian hormone and inhibin B. All measurements were obtained within two weeks of surgery, irrespective of cycle day.

Results—Significant differences were identified in ovarian primordial (p <0.0001) and non-growing follicle (p <0.0001) counts across the STRAW stages. In post-hoc testing, the differences in primordial follicle counts were significant between each of the STRAW stages. Significant differences were also identified in serum levels of anti-müllerian hormone, FSH and the ovarian antral follicle count across the STRAW stages.

Conclusions—Progression through the STRAW stages as defined by menstrual cycle characteristics is associated with progressive and significant decreases in the ovarian primordial follicle number.

Keywords

menopause; transition; primordial follicles; reproductive aging

Introduction

The development of a staging system for reproductive aging is of considerable interest to both researchers and clinicians alike in order to better predict the timing and duration of the menopausal transition and to standardize patient populations across studies. Ideally, such a system would include easily identifiable milestones for entry into progressively advanced stages of the menopausal transition and be highly reproducible across populations.[1,2]

Recognizing the need for such a staging system to address the medical and social consequences of reproductive aging, a sponsored workshop was held in 2001 to develop a reliable and useful staging system. The executive summary of this workshop created the Stages of Reproductive Aging Workshop (STRAW) staging system.[1] This system is anchored by the final menstrual period (FMP, Figure 1), and consists of seven stages. Stages -5 to -3 represent the reproductive years, whereas stages -2 to -1 represent the menopausal transition. Advancement from one stage to the next is indicative of a shorter interval of time prior to the FMP, and is categorized based on menstrual cycle characteristic changes (Figure 1). Stages +1 and +2 represent the early and late postmenopause, respectively. Only one biochemical parameter, serum FSH, was included in the initial STRAW proposal.

The participants of the STRAW conference recognized that the STRAW staging system was an initial proposal that would need to be validated and modified as additional information became available. Toward that end, the ReSTAGE collaboration[3-5] has considerably improved our understanding of the menstrual cycle changes associated with the progression from one STRAW stage to the next through the empiric review of menstrual calendars from population based cohorts including TEMIN, Melbourne Women's Midlife Health Project (MWMHP), Seattle Midlife Women's Health Study (SMWHS), and Study of Women's Health Across the Nation (SWAN).[6-9] Additionally, statistically significant differences in levels of hormones including FSH, inhibin B and anti-müllerian hormone (AMH) have been identified between the STRAW stages.[10-12] Unfortunately, because of the considerable overlap in serum levels of hormones, none of these endocrine markers are predictive of a specific STRAW stage.[11-13]

The underlying anatomical change associated with the progressive decline in fertility and ovarian endocrine function associated with aging is ovarian follicle depletion.[14,15] The basic ovarian follicle is the anatomic/functional unit of the ovary consisting of a primary oocyte and the surrounding granulosa cells. Many investigators consider the ovarian primordial follicle (PF) pool to constitute the ovarian reserve, whereas others consider the ovarian reserve to include the primordial as well as the intermediate and primary follicles.[16,17] Altogether, this cohort of resting follicles is known as the non-growing follicle (NGF) pool (Figure 2). Regardless of one's opinion on this issue, it is clear that the ovarian reserve encompasses one or all of these groups.

Given that the physiological basis for the reproductive aging process is ovarian follicle depletion, the ultimate validation of a staging system of reproductive aging would be the identification of significant differences in resting ovarian follicle numbers in women at different STRAW stages. Therefore, the goal of this investigation was to determine if significant differences in the ovarian PF and total NGF count exist between the STRAW stages as defined by menstrual cycle characteristics. To compare these observed differences in ovarian PF and total NGF count with commonly used clinical markers of ovarian reserve, we also examined endocrine and anatomical (the ovarian antral follicle count as determined by transvaginal ultrasound examination) differences between the STRAW stages in a subset (n = 43) of the participants.

Methods

Participants

As part of a series of investigations examining the age-related depletion of ovarian PF and NGF associated with aging, we collected ovaries from 119 women undergoing elective oophorectomy for benign gynecologic indications between 2001 and 2011. All participants completed a detailed questionnaire regarding menstrual cycle characteristics. Of the 119 women, 63 (20 from the University of Washington, 43 from the University of Oklahoma) could be classified into the original STRAW staging system by their menstrual cycle characteristics and were not taking any hormonal preparations that may have altered menstrual cycles. Participants were classified according to the STRAW staging system[1] based on *menstrual cycle characteristics only* (Figure 1) with the exception of Stages -4 and -3, wherein women with regular cycles and age ≤ 35 years were considered stage -4, and those ≥ 40 years were considered stage -3. Serum levels of FSH could not be utilized to classify participants into the respective STRAW stages, as these measurements were obtained irrespective of cycle day due to the usual short time frame between enrollment and surgery. Prior to surgery, each participant enrolled at the University of Oklahoma underwent a transvaginal ultrasound examination for the determination of the ovarian AFC, venipuncture for the determination of serum levels of AMH, estradiol, inhibin B and FSH, and completed the informed consent process. Participants were recruited from patients undergoing benign gynecologic surgery at the University of Oklahoma and the University of Washington. Exclusion criteria included gynecological malignancy, prior radiation or chemotherapy, autoimmune disease, and prior ovarian surgery. Additionally, ovarian pathology such as endometriomas, dermoid cysts, and other cystic masses of the ovary > 2 cm also excluded women from participation as did a solid ovarian mass of any size. Participants were required to be premenopausal (clinically defined). For the purpose of this investigation, women were not enrolled with amenonorrhea of greater than 3 months duration in order to avoid enrolling women that had experienced their final menstrual period. All participants underwent the ultrasound examination and venipuncture within two weeks of scheduled surgery. The authors were not involved in the decision to perform surgery, nor the surgery itself. If both ovaries were removed from a participant, the ovary best visualized on transvaginal ultrasound examination was selected for histological determination of the ovarian PF and NGF count. If a single ovary was removed, it was processed for the determination of the follicle count. Prior investigations have demonstrated a strong correlation in the ovarian PF and NGF count between the two ovaries constituting a pair.[18] At the University of Oklahoma, participants were compensated $50 for the time required to undergo the ultrasound examination, venipuncture, and to complete consent documents. This investigation was approved by the Universities of Oklahoma and Washington Institutional Review Boards.

Transvaginal Ultrasound Examinations for the determination of ovarian AFC

All ultrasound examinations were performed by a single investigator (K.R.H.) using Philips (Andover, MA) EnVisor or HD-7 ultrasound machine with a C8-4v vaginal transducer. All follicles 2-10 mm in size were considered to be antral follicles. The total number of antral follicles for each ovary was identified and recorded.

Hormone assays

Inhibin B assays were performed with a solid phase sandwich ELISA (Inhibin B Gen II ELISA, Beckman Coulter, Inc., Brea, CA) by Quest Diagnostics (San Juan Capistrano, CA). The limit of detection with this assay is 2.6 pg/mL and the limit of quantitation 10 pg/mL. At 19 pg/mL, the combined intra- and inter-assay coefficient of variation was 6.8%, and at 275 pg/mL, the combined intra- and inter-assay coefficient of variation was 4.3%. Values

below quantitative thresholds (the lower limit for reporting a result as set by the commercial laboratory) were given half of the threshold value in analyses.

The immunoassay for estradiol was performed using an Immulite autoanalyzer with reagents supplied by Siemens (Deerfield, IL). The lower limit of detection for the assay is 15 pg/mL. At 46 pg/mL, the intra-assay coefficient of variation was 15%, and at 116 pg/mL, the intra-assay coefficient of variation was 9.5%. At 56 pg/mL, the inter-assay coefficient of variation was 16%, and at 151 pg/mL, the inter-assay coefficient of variation was 9.3%.

The FSH assay is a solid-phase two-site chemiluminescent assay with reagents supplied by Siemens (Deerfield, IL) using an Immulite autoanalyzer. The lower limit of detection of the assay is 0.1mIU/mL. At 7.8 mIU/mL, the intra-assay coefficient of variation was 5.4%, and at 42.5 mIU/mL, the intra-assay coefficient of variation was 7.7%. At 8.3 mIU/mL, the inter-assay coefficient of variation was 8.1%, and at 42.9 mIU/mL, the inter-assay coefficient of variation was 7.9%.

AMH measurements were performed with a commercially available immunoenzymometric assay (AMH Gen II ELISA, Beckman Coulter, Inc., Brea, CA) by Quest Diagnostics (Valencia, CA). The limit of detection with this assay is 0.08 ng/mL and the limit of quantitation is 0.16 ng/mL. At 0.5 ng/mL the combined intra- and inter-assay coefficient of variation was 14%. At 4.42 ng/mL, the combined intra- and inter-assay coefficient of variation was 7.7%, and at 14 ng/mL, the combined intra- and inter-assay coefficient of variation was 5.8%. Values below quantitative thresholds (the lower limit for reporting a result as set by the commercial laboratory) were given half of the threshold value in analyses.

Tissue preparation, follicle counting and stereology

Tissue preparation, follicle identification (PF and NGF), and counting were performed utilizing a validated technique combining systematic random sampling and the optical disector as previously described[18,19] Briefly, the fractionator/optical disector method is based on directly counting the particles of interest (in this case, the oocyte nucleoli) in a known fraction of the original structure. The total number of nucleoli encountered in this fraction is then multiplied by the inverse of a hierarchy of systematic random sampling fractions in order to generate an estimate of the total number in the original specimen.

The first sample fraction (F1) consisted of the original ovary minus the small portion previously removed for pathological examination. Each ovary was cut into approximately 1 mm slabs perpendicular to the long axis of the ovary. Approximately 8 slabs were selected out of the total generated (yielding a second fraction, F2) using systematic random sampling rules. The selected slabs were dehydrated and embedded as a group in one or two large (2" × 3") blocks of glycol methacrylate (GMA, Technovit 8100, Energy Beam Sciences, Inc., Agawam, MA). The blocks were sectioned at a thickness of 25 μm using a rotary microtome. Every 10th section (the third fraction, F3) was collected in the order generated on glass slides for staining. Sections were stained with Richardson's stain and then mounted with cover slips using Cytoseal 280 (Stephens Scientific, Kalamazoo, MI).

Sections representing the largest 2-dimensional profile of each slab were then selected for counting with the optical disector. The fraction that this section represented from the entire collected stack of sections from each slab (F4) was determined by placing a point grid over the section and summing the points that fell over the sections. This value was then divided by the total number of points landing over all collected sections (including the initial and trailing partial slab fragments encountered at the beginning and end of the sectioning run across each slab).

Optical disector counting frames were placed over the selected stained sections using systematic random sampling rules.[20] Placement of optical disectors and delineation of the areas of interest was accomplished by use of StereoInvestigator software (MicroBright Field, Colchester, MA) operating on a PC style computer coupled to a Nikon (Univ. of Oklahoma) or Zeiss Photomicroscope II (University of Washington). Sequential placement of optical disector frames was performed by a motor driven microscope stage directed by the StereoInvestigator software.

The entire cortex of each section in the counting sample was outlined under low magnification for placement of the disector frames. The area of the disector frame divided by the area of the steps between placements (representing a grid) represented a fifth sampling fraction (F5). The next sampling fraction (F6) consisted of the height of the optical disector divided by the height of the tissue section. This fraction accounts for the portion of the tissue section represented by the guard area, in which no counting was performed.

Follicle identification

All follicles were classified according to the morphologic criteria as described by Gougeon.[21] The population of NGFs consisted of primordial (PF), intermediate, and primary follicles (Figure 2). Primordial follicles were defined as containing a single layer of flattened granulosa cells; intermediate follicles were defined as a single layer of granulosa cells with at least one cuboidal and one flattened granulosa cell; and primary follicles were defined as containing a single layer of cuboidal granulosa cells without any flattened granulosa cells (Figure 2). Raw counts (Q^-) for each class of NGFs were then converted to an estimate of the total number (N) of NGFs in the entire ovary by the following equation (where "Q^-" = number of each class of NGF identified in the fraction of tissue counted):

$$N = \Sigma Q - \times 1/(F1 \times F2 \times F3 \times F4 \times F5 \times F6)$$

Statistical Analysis

Baseline characteristics, endocrine levels, the ovarian antral follicle count as determined by transvaginal ultrasound examination (AFC), and log-transformed ovarian PF and total NGF counts were compared between the stages of the STRAW staging system with one-way analysis of variance. Post-hoc comparisons between STRAW stages, endocrine tests, age and the AFC were performed with the Bonferroni/Dunn method to adjust for multiple comparisons. Trends in ovarian follicle counts and the individual biomarkers of ovarian reserve were determined with linear regression. Statistical analyses were performed with StatView version 5.0.1. A p-value of less than 0.05 was considered statistically significant for the initial comparisons and a p-value less than 0.0083 was considered significant for the post-hoc tests.

RESULTS

Endocrine and AFC results of the participants (n = 63 total; n = 43 for endocrine and AFC measures) are summarized in Table 1. Box plots of endocrine parameters, the ovarian AFC as determined by transvaginal ultrasound examination, and log10-transformed ovarian primordial and total NGF counts are illustrated in Figure 3. Although age was significantly different across the STRAW stages (p < 0.0001, Table 1), there was no significant difference in age between stages -3, -2 and -1 in post-hoc testing (p = 0.68, 0.15, and 0.20 between stages -3 and -2, -3 and -1, and -2 and -1, respectively).

Ovarian PF counts were significantly different across the STRAW stages (p < 0.0001, Figure 3a), with significant decreases in PF count noted between each stage in post-hoc

testing (p < 0.0001 between each stage with the exception of the difference between stages -2 and -1, where p = 0.0074; test for trend p < 0.0001). Similarly, NGF counts decreased significantly with advancing STRAW stage (p < 0.0001, Figure 3b; test for trend p < 0.0001). Differences in ovarian NGF counts were significant between each stage in pair-wise comparisons (p < 0.001) with the exception of the difference between stages -2 and -1, where p = 0.015.

Similarly, serum FSH levels (Figure 3c) were significantly different among the STRAW stages (p = 0.024), with increasing values associated with more advanced STRAW stage (test for trend p = 0.015). In post-hoc testing, only the difference between STRAW stages -4 and -1 reached statistical significance (p = 0.0036). In contrast, estradiol and Inhibin B levels were not significantly different across the STRAW stages (p = 0.81 and 0.23, respectively, Table 1 and Figure 3f).

The ovarian AFC was significantly different between STRAW stages (p < 0.0001, Figure 3d), with a progressive decrease noted with advancing stage (test for trend p < 0.0001). In pair-wise comparisons, only the differences between STRAW stages -4 and -3, -4 and -2, and -4 and -1 were statistically significant (p < 0.0001). AMH levels also decreased significantly with advancing STRAW stage (p < 0.0001, Figure 3e, test for trend p <0.0001). In pair-wise comparisons; however, these differences were only significant between STRAW stages -4 and -3, -4 and -2, and -4 and -1 (p < 0.0001). Of note, all of the participants in STRAW stage -1 had AMH levels below the limit of quantitation (n = 5), as did 7 of 12 participants in STRAW stage -3 and 14 of 20 in STRAW stage -2.

DISCUSSION

The creation and validation of a staging system for reproductive aging is of paramount importance in our desire to better counsel women regarding future reproduction and health. For most women, the assertion that menopause will occur at age 51 ± 10 years is simply not adequate for family planning and risk assessment. Our desire to better predict menopause is rooted in the understanding that some women will experience an early menopause with the corresponding increased risks for infertility, cardiovascular disease and osteoporosis.[15,22,23] Conversely, others will experience a late menopause, with the possibility of an increased risk for breast and endometrial cancer.[24,25] In order to conduct meaningful studies on reproductive aging, it is necessary to have a valid means by which to characterize study populations. Ten years ago, the members of the STRAW workshop proposed a staging system of reproductive aging to address this need with the understanding that future modifications and validation would be necessary.

In this investigation we have identified significant differences in the ovarian PF count between STRAW stages as defined by menstrual cycle characteristics. Given that Richardson et al.[26] identified lower ovarian PF counts in women with cycle irregularity compared to those with regular menstrual cycles, the decrease in PF count associated with irregular cycles themselves was anticipated. The important aspect of this investigation is the demonstration that progressively decreasing PF counts are associated with advancement through the STRAW stages as defined by menstrual cycle characteristics, each representing a window of time prior to the menopause. Thus, this investigation represents an important step in the validation of the STRAW staging system. Additionally, ovarian PF counts were significantly different between stages -3,-2, and -1, whereas age was not significantly different between these groups. Therefore, STRAW stage as defined by menstrual cycle characteristics provides information regarding the ovarian reserve above and beyond age alone.

We were unable in include serum levels of FSH to categorize women into the specific STRAW stages, as levels were obtained irrespective of cycle day due to the short time frame between enrollment and surgery. As a result, only menstrual cycle characteristics were utilized to classify women into the STRAW stages. Given that the only distinction between STRAW stages -4 and -3 is an elevated early follicular phase FSH level, we cannot exclude the possibility that some of the participants categorized as stage -3 could have had normal early follicular phase FSH levels and should have been categorized as stage -4. Nevertheless, the average age of the women in stage -3 was 45. 4 years, an age range that would typically be considered beyond the "peak" reproductive years (stage -4).

Although we were able to detect significant differences in the ovarian PF count between the STRAW stages in this investigation, determining ovarian PF counts with modern morphometric techniques is strictly a research tool, not a clinical assay. Therefore, considerable efforts have been put forth into identifying biomarkers which would be useful in predicting the age of menopause and could be incorporated into a clinical staging system of reproductive aging. Perhaps the most promising of these markers are serum levels of AMH and the ovarian AFC as assessed by transvaginal ultrasound examination. Both are strongly correlated with the ovarian PF and NGF count[27], the true measures of ovarian reserve, and both are known to decline with chronological age in cross-sectional and some longitudinal investigations.[28-32] Furthermore, unlike serum levels of FSH, inhibin B, and estradiol, serum levels of AMH and the ovarian AFC appear to be largely cycle-phase independent.[33,34]

In order to contrast the observed differences in the ovarian PF and NGF count with biomarkers of ovarian reserve, we measured serum levels of AMH, inhibin B, FSH, and estradiol, and determined the ovarian AFC in a subset of the participants. Although we detected significant differences in serum levels of AMH, FSH, and the ovarian AFC across the STRAW stages, we were unable to detect significant differences in pair-wise analyses for any of the biomarkers between stages -3, -2 and -1. Other investigations; however, have identified significant differences in AMH[11] and FSH[13] levels between all STRAW stages. Possible explanations for the discrepancy between our findings and those of other investigators include our smaller sample size and, in the case of FSH, the lack of cycle-phase specific measurements. Additionally, the currently available commercial assay for AMH (AMH Gen II ELISA, Beckman Coulter, 0.08 ng/mL), has a much higher limit of detection than the prior assay utilized in the Hale et al.[11] investigation (Diagnostic Systems Laboratories, Inc., Webster, TX, 0.017 ng/mL). The lack of sensitivity of the current assay suggests AMH levels may be less useful in models forecasting the age of the FMP when an individual is at more advanced stages of reproductive age (e.g. STRAW stages -2 and -1). Conversely, other longitudinal investigations have demonstrated that AMH exhibits more consistent changes over time than the ovarian AFC and serum levels of FSH and inhibin B in younger reproductive-aged women.[30] Therefore, serum levels of AMH may be more useful in predicting the age of menopause when an individual is at a less advanced stage of reproductive age. Indeed, a recent longitudinal investigation has demonstrated reasonable agreement between the predicted and observed age of menopause with serum levels of AMH drawn ~ 6 years prior to the FMP.[35] Additional prospective longitudinal studies are needed to confirm these findings.

Serum inhibin B and estradiol concentrations were not significantly different across the STRAW stages in this investigation. Others investigations; however, have identified significant differences in these hormones across the STRAW stages, although not between all stages in pair-wise comparisons.[11,13] As with serum measurements of FSH, it seems likely that the lack of significant differences in inhibin B and estradiol levels between the

STRAW stages in this investigation are due to the smaller sample size and the lack of cycle-day specific measurements.

In considering the relative value of determining ovarian PF and NGF counts to validate the STRAW staging system as well as to investigate fundamental aspects of reproductive aging, it is useful to review the strengths and limitations of the study design. The strength of an investigation of this nature includes the fact that the major dependent variable of interest, the true ovarian reserve, is what is being measured. Studies utilizing the FMP as the outcome of interest rely on another organ, the uterus, for a lack of uterine bleeding. In effect, the FMP is a surrogate marker for ovarian follicle depletion. Although the FMP would appear to be a reasonable marker in most cases, some women may cease having menstrual cycles for reasons other than follicle depletion (e.g. due to thyroid, hypothalamic or pituitary dysfunction). Similarly, vaginal bleeding may persist beyond follicle depletion in some women (e.g. due to estrogen production from adipose tissue). Additionally, longitudinal studies investigating the age of the FMP can only identify this event one year after it has passed. Therefore, longitudinal studies are time-consuming and expensive relative to investigations utilizing ovarian follicle counts as the outcome of interest. Given that counts of ovarian PFs and NGFs are direct measures of the ovarian reserve, it is not unexpected that these measures would have more discriminatory power between the STRAW stages than would any of the proposed biomarkers.

Limitations of utilizing PF and NGF counts as the best indicator of ovarian aging include that these studies can only be cross-sectional by their very nature. Furthermore, the argument can be made that women undergoing surgical removal of the uterus and ovaries cannot be considered "normal", and thus it may not be appropriate to extrapolate the results to the general population. Although processing ovaries for determination of the ovarian PF and NGF count with modern morphometric techniques is relatively quick and inexpensive, obtaining appropriate clinical specimens is time consuming and difficult. This is particularly true at the more advanced stages of reproductive age, wherein indications for surgery are rare. Finally, our study design relies upon the participants' recall of menstrual cycle characteristics, which may be inaccurate.[36,37] Nevertheless, the recollection of menstrual cycle characteristics would seem to generally be adequate in a clinical situation wherein a staging system of reproductive aging would be utilized.

Although this investigation makes progress towards the validation of the STRAW staging system, the lack of specific "cut points" for categorizing an individual into one stage versus another based on biomarkers highlights the limitations of our current knowledge. Examining endocrine, ultrasound, and even advanced anatomical data is useful for categorizing groups, but not individuals into a specific stage. Future large prospective studies examining changes in AMH and the ovarian AFC associated with the distinct stages of the STRAW staging system and the development of more sensitive AMH assays may be useful in this regard.

CONCLUSIONS

The present study demonstrates that significant differences in the ovarian PF count exist between the different stages of the STRAW staging system as defined by menstrual cycle characteristics. This finding suggests that the progressive STRAW stages do represent distinct and progressively more advanced stages of reproductive age. Considerable overlap exists between the STRAW stages in the levels of the ovarian and peripheral biomarkers of reproductive age evaluated in this investigation. As such, these markers currently have limited utility in categorizing individuals into a particular STRAW stage and ultimately forecasting the age of depletion of the ovarian reserve. From a practical standpoint, models and staging systems that forecast the age of the FMP would be most useful if they could

identify the age of the FMP at a relatively young age rather than when ovarian failure is imminent. AMH may be the most promising biomarker in this regard. Further longitudinal research is needed to characterize the decline in this hormone and other biomarkers associated with reproductive aging.

Acknowledgments

References

1. Soules MR, Sherman S, Parrott E, et al. Executive Summary: Stages of reproductive aging workshop (STRAW). Fertil Steril. 2001; 76:874–878. [PubMed: 11704104]

2. Sherman S. Defining the menopausal transition. Am J Med. 2005; 118:3s–7s.

3. Harlow SD, Mitchell ES, Crawford S, et al. The ReSTAGE Collaboration: defining optimal bleeding criteria for onset of early menopausal transition. Fertil Steril. 2008; 89:129–140. [PubMed: 17681300]

4. Harlow SD, Cain K, Crawford S, et al. Evaluation of four proposed bleeding criteria for the onset of late menopausal transition. J Clin Endocrinol Metab. 2006; 91:3432–3438. [PubMed: 16772350]

5. Taffe JR, Cain KC, Mitchell ES, et al. "Persistence" improves the 60 days amenorrhea marker of entry to late stage menopausal transition for women aged 40-44. Menopause. 2010; 17:191–193. [PubMed: 19713872]

6. Treloar AE, Boynton RE, Behn BG, et al. Variation of human menstrual cycle through reproductive life. Int J Fertil. 1967; 12:77–126. [PubMed: 5419031]

7. Mitchell ES, Woods NF, Mariella A. Three stages of the menopausal transition from the Seattle midlife women's health study: toward a more precise definition. Menopause. 2000; 7:334–349. [PubMed: 10993033]

8. Sowers, MF.; Crawford, S.; Sternfeld, B., et al. Design, survey sampling and recruitment methods of SWAN: a multi-center, multi-ethnic, community-based cohort study of women and the menopausal transition.. In: Wren, J.; Lobo, RA.; Kelsey, J.; Marcus, R., editors. Menopause: biology and pathobiology. Academic Press; San Diego, CA: 2000. p. 175-188.

9. Taffe J, Dennerstein L, MacLennan A. Menstrual diary data and the menopausal transition: methodological issues. Acta Obstet Gynecol Scand. 2001; 81:588–594. [PubMed: 12190832]

10. Randolf, Jf; Crawford, S.; Dennerstein, L., et al. The value of follicle-stimulating hormone concentration and clinical findings as markers of the late menopausal transition. J Clin Endocrinol Metab. 2006; 91:3034–3040. [PubMed: 16720656]

11. 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; 92:3060–3067. [PubMed: 17550960]

12. Yang YS, Hur MH, Kim SY, Oh KY. Correlation between sonographic and endocrine markers of ovarian aging as predictors for late menopausal transition. Menopause. 2011; 18:138–145. [PubMed: 21341396]

13. Gracia CR, Sammel MD, Freeman EW, et al. Defining menopause status: creation of a new definition to identify the early changes of the menopausal transition. Menopause. 2005; 12:128–135. [PubMed: 15772558]

14. te Velde ER, Pearson PL. The variability of female reproductive aging. Hum Reprod Update. 2002; 8:141–54. [PubMed: 12099629]

15. Broekmans FJ, Soules MR, Fauser BC. Ovarian aging: mechanisms and clinical consequences. Endocrinol Rev. 2009; 30:465–93.

16. Lambalk CB, van Disseldorp J, de Koning CH, Broekmans FJ. Testing ovarian reserve to predict age at menopause. Maturitas. 2009; 63:280–291. [PubMed: 19631481]

17. Gougeon, A. The early stages of follicular growth.. In: Trounson, AO.; Gosden, RG., editors. Biology and pathology of the oocyte: role in fertility and reproductive medicine. Cambridge University Press; Cambridge: 2003. p. 29-41.

18. Charleston JS, Hansen KR, Thyer AC, et al. Estimating human ovarian non-growing follicle number-the application of modern stereology techniques to an old problem. Hum Reprod. 2007; 22:2103–2110. [PubMed: 17548367]

19. Hansen KR, Knowlton NS, Thyer AC, Charleston JS, Soules MR, Klein NA. A new model of reproductive aging: the decline in ovarian non-growing follicle number from birth to menopause. Hum Reprod. 2008; 23:699–708. [PubMed: 18192670]

20. Gundersen JJ, Bagger TF, Bendtsen SM, et al. The new stereological tools: Disector, fractionator, nucleator and point sampled intercepts and their use in pathological research and diagnosis. APMIS. 1988; 96:857–881. [PubMed: 3056461]

21. Gougeon A. Regulation of ovarian follicular development in primates: facts and hypotheses. Endocr Rev. 1996; 17:121–155. [PubMed: 8706629]

22. Atsma F, Bartelink ML, Grobbee DE, van der Schouw YT. Postmenopausal status and early menopause as independent risk factors for cardiovascular disease: a meta-analysis. Menopause. 2006; 13:265–279. [PubMed: 16645540]

23. van Der Voort DJ, van der Weijer PH, Barentsen R. Early menopause: increased fracture risk at older age. Osteoporosis Int. 2003; 14:525–530.

24. Trichopoulos D, MacMahon B, Cole P. Menopause and breast cancer risk. J Natl Cancer Inst. 1972; 48:605–613. [PubMed: 5058966]

25. MacMahon B. Risk factors for endometrial cancer. Gynecol Oncol. 1974; 2:122–129. [PubMed: 4457398]

26. Richardson S, Senikas V, Nelson J. Follicular depletion during the menopausal transition: evidence for accelerated loss and ultimate exhaustion. J Clin Endocrinol Metab. 1987; 65:1231–1237. [PubMed: 3119654]

27. Hansen KR, Hodnett GM, Knowlton N, Craig LB. Correlation of ovarian reserve tests with histologically determined primordial follicle number. Fertil Steril. 2011; 95:170–175. [PubMed: 20522327]

28. Rosen MP, Sternfeld B, Schuh-Huerta SM, Reijo Pera RA, McCulloch CE, Cedars MI. Antral follicle count: absence of significant midlife decline. Fertil Steril. 2010; 94:2182–2185. [PubMed: 20149366]

29. Scheffer GJ, Broekmans FJM, Dorland M, et al. Antral follicle counts by transvaginal ultrasonography are related to age in women with proven natural fertility. Fertil Steril. 1999; 72:845–851. [PubMed: 10560988]

30. van Rooij IA, Broekmans FJ, Scheffer GJ, et al. Serum antimullerian hormone levels best reflect the reproductive decline with age in normal women with proven fertility: a longitudinal study. Fertil Steril. 2005; 83:979–987. [PubMed: 15820810]

31. 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; 93:3478–3483. [PubMed: 18593767]

32. van Rooij IA, Tonkelaar I, Broekmans FJ, et al. Anti-mullerian hormone is a promising predictor for the occurrence of the menopausal transition. Menopause. 2004; 11:601–606. [PubMed: 15545787]

33. 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; 91:4057–4063. [PubMed: 16804046]

34. Pache TD, Wladimiroff JW, de Jong FH, Hop WC, Fauser BCJM. Growth patterns of nondominant ovarian follicles during the normal menstrual cycle. Fertil Steril. 1990; 54:638–42. [PubMed: 2209884]

35. Tehrani FR, Shakeri N, Solaymani-Dodaran M, Azizi F. Predicting age at menopause from serum antimüllerian hormone concentration. Menopause. 2011; 18:766–770. [PubMed: 21451424]

36. Bean JA, Leeper JD, Wallace RB, Sherman BM, Jagger H. Variations in the reporting of menstrual histories. Am J Epidemiol. 1979; 109:181–185. [PubMed: 425957]

37. Taffe J, Dennerstein L. Retrospective self-reporting compared with menstrual diary data prospectively kept during the menopausal transition. Climacteric. 2000; 3:183–191. [PubMed: 11910620]

Stages:	-5	-4	-3	-2	-1	0	+1	+2
Terminology:	Reproductive			Menopausal Transition			Postmenopause	
	Early	Peak	Late	Early	Late*		Early*	Late
				Perimenopause				
Duration of Stage:	variable			variable		(a) 1 yr	(b) 4 yrs	until demise
Menstrual Cycles:	variable to regular	regular		variable cycle length (>7 days different from normal)	≥2 skipped cycles and an interval of amenorrhea (≥60 days)	Amen x 12 mos	none	
Endocrine:	normal FSH		↑ FSH	↑ FSH		↑ FSH		

Final Menstrual Period (FMP) 0

Figure 1.
The STRAW workshop stages of reproductive aging adapted from Soules et al.[1] *Stages most likely to be characterized by vasomotor symptoms.
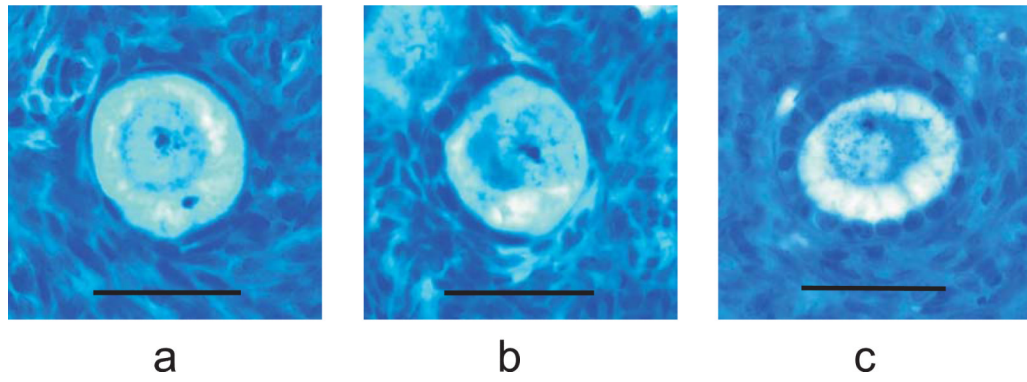
Figure 2.
Human ovarian non-growing follicles. The ovarian non-growing, or resting follicle pool consists of primordial (a), intermediate (b), and primary (c) follicles. Bar = 45 μm.
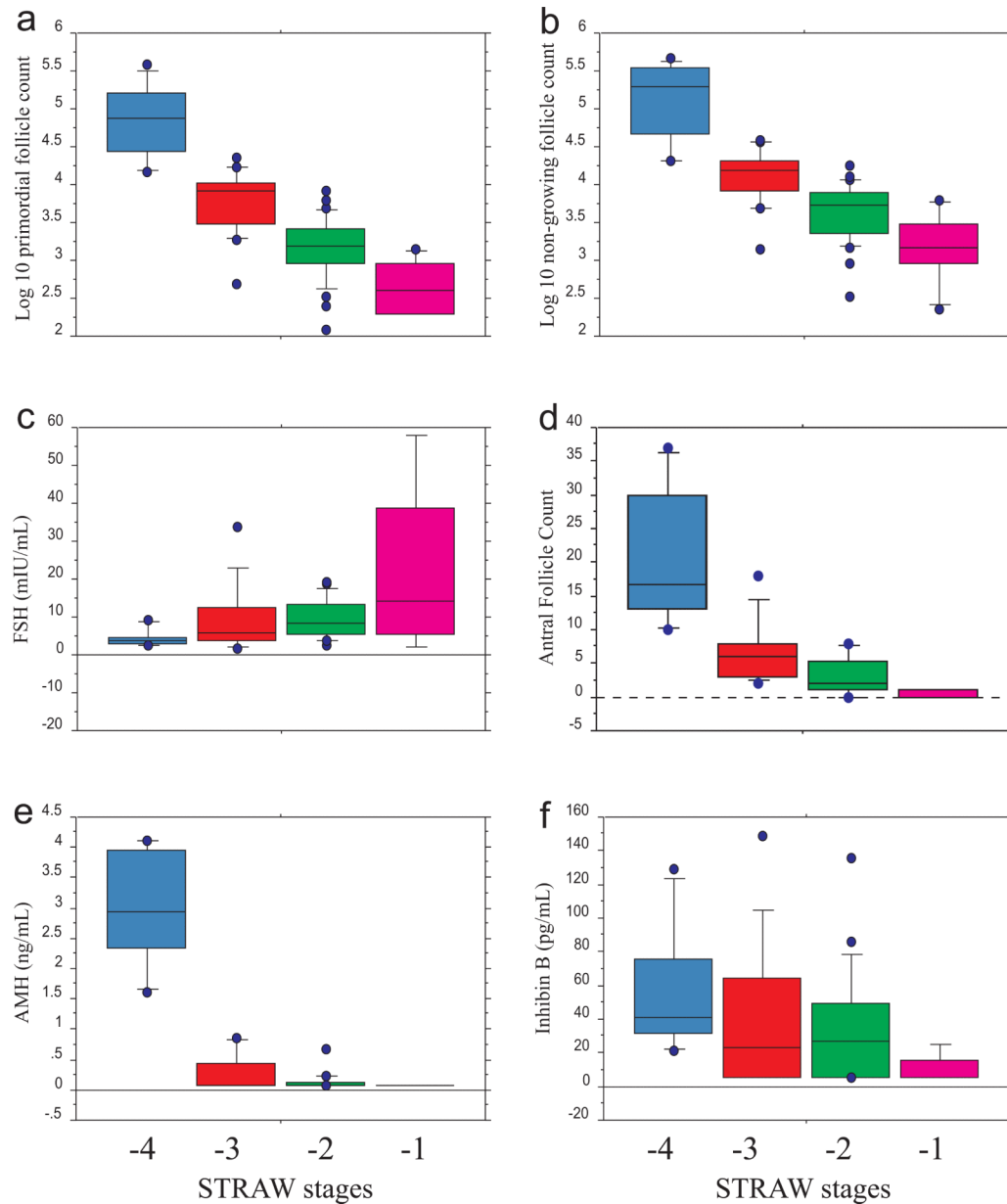
Figure 3.
Box plots of log10-transformed ovarian primordial and non-growing follicle counts, total ovarian antral follicle counts, and biomarkers of ovarian reserve for STRAW stages -4 through -1. a) Log 10 primordial follicle count: p < 0.0001 between all stages except -2/-1, where p = 0.0074; b) Log 10 non-growing follicle count: p < 0.0001 between all stages except -2/-1 where p = 0.015; c) FSH: p = 0.0036 between stages -4/-1, all others are not significantly different; d) Antral follicle count: p < 0.0001 between all stages except -3/-2, -3/-1, and -2/-1 which are not significantly different; e) AMH: p < 0.0001 between all stages except -3/-2, -3/-1, and -2/-1 which are not significantly different; f) Inhibin B: ANOVA not significant (p = 0.23).

Table 1

Summary characteristics of study participants (n = 63 total, 43 for endocrine and ultrasound measures). Results are reported as means ± 1SD. P-values refer to one-way ANOVA. Endocrine tests and ultrasound examinations were performed irrespective of cycle day.

STRAW stage	-4	-3	-2	-1	p-value
n (total, endocrine)	8(6)	16(12)	33(20)	6(5)	<0.0001
Age (years)	32.0 ± 3.5	45.4 ± 3.1	45.8 ± 3.1	47.5 ± 1.9	<0.0001
AMH (ng/mL)	2.98 ± 1.01	0.27 ± 0.29	0.13 ± 0.14	.08 ± 0 *	<0.0001
Inhibin B (pg/mL)	56.5 ± 40.0	39.6 ± 44.9	35.1 ± 33.7	10.6 ± 8.8	0.23
FSH (mIU/mL)	4.52 ± 2.5	9.32 ± 9.3	9.56 ± 5.3	22.56 ± 22.9	0.024
Estradiol (pg/mL)	97.1 ± 65.7	118.8 ± 79.9	95.8 ± 89.9	79.6 ± 79.5	0.81
Total AFC	20.5 ± 10.7	6.9 ± 4.7	3.2 ± 2.6	0.5 ± 1.0	<0.0001

*
AMH levels in all participants in the STRAW -1 stage were below the limits of quantitation