Antral follicle counts by transvaginal ultrasonography are related to age in women with proven natural fertility

Gabriëlle J. Scheffer, M.D.,* Frank J. M. Broekmans, Ph.D.,* Marinus Dorland, Ph.D.,* Johannes D. F. Habema, Ph.D.,† Caspar W. N. Looiman, MSc.,† and Egbert R. te Velde, Ph.D.*

University Hospital Utrecht, Utrecht, and Erasmus University Medical School, Rotterdam, the Netherlands

Objective: To investigate the relation between reproductive age and ultrasound (US)-based follicle counts and the reproducibility of follicle counts in regularly cycling women with proven fertility.

Design: Prospective observational study.

Setting: Tertiary fertility center.

Patient(s): Healthy female volunteers with proven fertility, recruited by advertisement in local newspapers.

Intervention(s): The number of antral follicles sized 2–10 mm and ovarian volume were estimated by transvaginal US in the early follicular phase of the menstrual cycle in 162 women. A subgroup of 81 women underwent transvaginal US at several times in three subsequent cycles.

Main Outcome Measure(s): Antral follicle count and total ovarian volume.

Result(s): Women aged 25–46 years (n = 162) were studied. The relation of age with the US indices was computed after natural log transformation. Antral follicle count showed the clearest correlation with age \( R = 0.67 \). A biphasic linear model gave the best fit to the data. Before the age of 37 years, the antral follicle count showed a mean yearly decline of 4.8%, compared with 11.7% thereafter. The reproducibility of the antral follicle count in two subsequent cycles was moderate.

Conclusion(s): The number of small antral follicles in both ovaries as measured by US is clearly related to reproductive age and could well reflect the size of the remaining primordial follicle pool. (Fertil Steril 1999; 72:845–51. ©1999 by American Society for Reproductive Medicine.)

Key Words: Transvaginal sonography, reproductive aging, antral follicles, proven fertility, ovarian reserve test

Demographic and clinical studies show that female fertility starts to decrease from the age of 30 years, and the ability to conceive has become almost zero at a median age of 41 years (1, 2). In infertility practice, the role of the age-related decrease in female fertility has been recognized more and more (3, 4). Although basal FSH may help to identify severe loss of ovarian function, an adequate test to assess the reproductive status of an individual woman has not yet been established. One reason for this is the limited understanding of the processes determining aging of the reproductive axis.

Aging of the ovary seems to play the major role in reproductive aging and is related to both a decrease in the quality of the oocytes and a gradual reduction in the number of primordial follicles (5). Although oocyte quality may be an important factor in the age-related decline in fertility (6–8), direct assessment of oocyte integrity is not possible to date. The number of primordial follicles declines exponentially through childhood and adult life, leading to ovaries that are almost devoid of follicles at the age of menopause (9). Likewise, the number of follicles that have entered the growing phase toward the antral stages of development decreases with age.

The number of primordial follicles in the ovary appears to be correlated with the number of growing follicles (10). It is almost certain, therefore, that the decline in primordial follicle reserve leads to a decreased size of the antral...
Antral follicles in the ovary can be readily visualized by the use of transvaginal US (12). It has been reported that with modern US equipment, antral follicles of ≥2 mm can be visualized with use of a transvaginal transducer (13). More important, transvaginal US provides an accurate and reproducible measurement of the total number of antral follicles throughout the menstrual cycle with favorable intraobserver variability (13).

In this study, we investigated the relation between the number of small antral follicles in both ovaries, as measured by US, and reproductive age in a large number of healthy, regularly cycling women with proven natural fertility. In addition, the reproducibility of US-based antral follicle counts was assessed. In such a specific group of women of various ages, reproductive age can be assumed to be strongly related to chronological age. Therefore, the age-related decline in antral follicle count can be considered to reflect the loss of reproductive potential.

**MATERIALS AND METHODS**

**Subjects**

This study was approved by the local ethics committee, and written informed consent was obtained from all participants. Healthy women (n = 162; age 25–46 years) were recruited by advertisement in the local newspapers. Volunteers were enrolled in the study protocol if they met all of the following criteria: [1] regular menstrual cycles varying from 21 to 35 days, [2] biphasic BBT, [3] proven natural fertility by having carried at least one pregnancy to term, [4] each of the pregnancies established within 1 year after the interruption of contraceptive methods, [5] no evidence of endocrinologic disease, [6] no history of ovarian surgery, [7] no ovarian abnormalities as assessed by vaginal US, and [8] cessation of hormonal contraception ≥2 months before entering the study protocol. The volunteers received monetary compensation for study participation.

**Experimental Design**

All volunteers visited our research department in the early follicular phase of the menstrual cycle (cycle day 2, 3, or 4). Transvaginal US was performed to measure the size and number of follicles and the volume of both ovaries. In a subgroup of 81 women, repeated transvaginal US examinations were performed during three subsequent menstrual cycles, starting in the midluteal phase (according to a temperature shift in the BBT) of the first study cycle. From that point on, US scans were performed every 2 or 3 days.

When the dominant follicle in the midfollicular phase of the second study cycle reached a diameter of ≥14 mm, US scans were performed daily until ovulation occurred. Thereafter, US examination was performed only once in the midluteal phase of the second study cycle and on cycle day 3 of the subsequent third study cycle. All women documented a body temperature chart during one menstrual cycle, and the subgroup of 81 women recorded a body temperature chart during two subsequent menstrual cycles.

**Transvaginal US Measurements**

All transvaginal US measurements were performed by the same observer (G.J.S.) using a 7.5-MHz transvaginal probe on a Toshiba Capasee SSA-220A (Toshiba Medical Systems Europe BV, Zoetermeer, the Netherlands). The ovary was examined by scanning from the outer to the inner margin, as described previously (13). Round or oval echo-free structures in the ovaries were regarded as follicles and were counted and measured as such. The numbers of follicles in both ovaries were added for the antral follicle count.

Pache et al. (13) performed intraobserver reproducibility studies and concluded that transvaginal US permits precise determinations of the total number of small antral follicles.

In the statistical analysis, follicles with a diameter of up to 10 mm were included. This is based on the finding that follicle size can vary up to 10 mm before the dominant follicle is identified (13). By including follicles up to 10 mm, we determined the antral follicle cohort at its maximal size.

Follicle diameter was calculated from the mean of two perpendicular measurements if a follicle measured <6 mm and from the mean of three perpendicular measurements if a follicle measured >6 mm. The volume of each ovary was calculated by measuring the three perpendicular diameters and applying the formula for an ellipsoid: \( V = \frac{4}{3} \pi \times a \times b \times c \), where \( a, b, \) and \( c \) are the perpendicular measurements if a follicle measured \( >6 \) mm. The volumes of both ovaries were added for the total ovarian volume.

**Statistical Analysis**

Statistical analysis of the data was performed with the Generalized Linear Interactive Modelling (GLIM) package (N.A.G., Oxford, United Kingdom) (14) and SPSS (Statistical Package for Social Sciences; SPSS, Inc., Chicago, IL) for Windows (release 7.5.2). Correlation between the study variables and age was analyzed by linear and nonlinear regression. A normalizing log transformation was applied for the correlation and regression analysis. Goodness-of-fit analysis was used to test the applicability of the simple linear model to the logarithmic data. The \( \beta \) of the regression formula was then transformed into the percentage change in the corresponding variable per year of age. Biphasic patterns were analyzed by nonlinear regression using the same log-transformed numbers and (in principle) the same least-squares approximation.
The reproducibility of antral follicle counts in two subsequent cycles was analyzed by calculating limits of agreement as described by Bland and Altman in 1986 (15) and 1995 (16). The objectives of this specific method are to describe the degree of agreement of measurements and to assess any trend in the extent of agreement throughout the range of measurements. The limits-of-agreement method is based on calculating the mean difference between two measurements as well as the SD of the differences. The limit of agreement is defined as 1.96 times the SD above or below the mean difference and indicates to what extent two measurements can vary.

### RESULTS

The median age of the 162 women was 38.0 years, and the median cycle length was 28 days. There was no even distribution of the women over the age range of 25–46 years. The low number of women aged <31 years was the cause of this uneven distribution. There were no weight restrictions for this study. The median body mass index of the 162 women was 23.5 kg/m²; the 25th percentile was 21.8 kg/m² and the 75th percentile was 26.9 kg/m².

Median values and ranges of the various US indices are summarized in Table 1. The majority of follicles visualized by US in the early follicular phase were 2–5 mm in size. There was a wide range for all the variables examined. Table 2 shows that there was no difference between the median number of follicles or between volumes of the right and left ovary.

The relation between age and the various US indices after natural log transformation was computed by the use of linear regression and is summarized in Table 3. All correlations were highly significant. Total ovarian volume showed only a moderate correlation with age. The correlation between the volume of the smallest or the largest ovary and age was less marked ($R = -0.23$, $P = .005$ and $R = -0.25$, $P = .002$, respectively) than the correlation between the total ovarian volume and age. The number of follicles measuring 2–10 mm in the early follicular phase showed the clearest correlation with age ($R = -0.67$, $P < .001$).

In each chronological year, the number of antral follicles declined by 8.2% (95% confidence interval [CI], 5.2%–11.2%). Figure 1 shows that there was a considerable variation in the antral follicle count per age group. For example, at the age of 30 years, the follicle number varied from 7 to 22, whereas for 40-year-old women, this range was 2 to 7. Goodness-of-fit analysis revealed a biphasic linear model as a better descriptive fit of the log-transformed data. Estimation of breakpoint and different slopes before and after was performed by fitting a nonlinear model. The $F$ test revealed that this model was significantly better than a model with a single slope ($P = .02$, 2 df). The breakpoint appeared to be at 38 years of age. Because the 95% CI for the breakpoint (34–42 years) included 37, as found by Faddy et al. (9), we decided to use the model with 37 years for the rest of this study for reasons of comparability. Before the age of 37, the number of follicles declined at 4.8% per year (95% CI, 2.1%–7.4%), but after that the rate was 11.7% per year (95% CI, 5.2%–11.2%).

### TABLE 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total study group (n = 162)</th>
<th>Subgroup (n = 81)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Range</td>
</tr>
<tr>
<td>No. of follicles 2–10 mm</td>
<td>7</td>
<td>1–34</td>
</tr>
<tr>
<td>No. of follicles 2–5 mm</td>
<td>5</td>
<td>0–30</td>
</tr>
<tr>
<td>No. of follicles &gt;5–10 mm</td>
<td>2</td>
<td>0–13</td>
</tr>
<tr>
<td>Total ovarian volume (cm³)</td>
<td>10.3</td>
<td>4.5–40.3</td>
</tr>
</tbody>
</table>


### TABLE 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total study group (n = 162)</th>
<th>Subgroup (n = 81)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Range</td>
</tr>
<tr>
<td>Right ovary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of follicles 2–10 mm</td>
<td>4</td>
<td>0–16</td>
</tr>
<tr>
<td>Ovarian volume (cm³)</td>
<td>4.7</td>
<td>1.6–11.8</td>
</tr>
<tr>
<td>Left ovary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of follicles 2–10 mm</td>
<td>4</td>
<td>0–23</td>
</tr>
<tr>
<td>Ovarian volume (cm³)</td>
<td>4.4</td>
<td>1.2–19.7</td>
</tr>
</tbody>
</table>


### TABLE 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pearson’s coefficient</th>
<th>Mean decline per year of age (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Follicles 2–10 mm</td>
<td>−0.67*</td>
<td>8.2</td>
</tr>
<tr>
<td>Follicles 2–5 mm</td>
<td>−0.58*</td>
<td>9.1</td>
</tr>
<tr>
<td>Follicles 5–10 mm</td>
<td>−0.24†</td>
<td>2.9</td>
</tr>
<tr>
<td>Total ovarian volume (cm³)</td>
<td>−0.29*</td>
<td>2.4</td>
</tr>
</tbody>
</table>

* $P < .001$.
† $P < .01$.

In the subgroup of 81 women who underwent serial transvaginal US studies, the age distribution, cycle characteristics, and results of US measurements (Table 1) were not different from those in the study group as a whole. Table 4 shows the results of regression analysis of the relation between follicle number (2–10 mm) and age at various times in the menstrual cycle. It appears that the correlation between age and antral follicle count is independent of the moment in the menstrual cycle at which the number of follicles is counted. If the mean number of follicles measured in the two subsequent cycles is analyzed, the correlation with age appears to be slightly improved (\( R = 0.73\), \( P < .001\)).

Reproducibility was assessed by calculating limits of agreement between two early follicular measurements, as shown in Figure 2. The mean difference between repeated measurements was \(-0.28\). The upper and lower limits of agreement of the measurements were 8.37 (95% CI, 7.41–9.33) and \(-8.93\) (95% CI, \(-7.41\) to \(-9.89\)), respectively. This implies only moderate agreement between the measurements. There was no relation between the mean and the difference of the two counts, indicating that the degree of agreement was not different for lower or higher follicle numbers.

The two midluteal measurements also showed a moderate agreement, with a mean difference of \(-1.09\), an upper limit of agreement of \(8.03\) (95% CI, 6.97–9.09), and a lower limit of agreement of \(-10.1\) (95% CI, \(-9.05\) to \(-11.17\)). Finally, two early follicular measurements of the total ovarian volume also showed moderate agreement, with a mean difference of \(0.24\), an upper limit of agreement of 6.38 (95% CI, 5.62–7.14), and a lower limit of agreement of \(-5.9\) (95% CI, \(-5.14\) to \(-6.66\)).
DISCUSSION

In this study, we showed that the number of antral follicles assessed by transvaginal US is clearly correlated with chronological age. A relation between antral follicle count and age was also found in women with regular cycles but with unproven fertility (17). The women in this study were selected on the basis of strict criteria for normal reproductive performance during their lives. Therefore, our assumption seems justified that chronological age approximates reproductive age in this group of women and that they illustrate the normal gradual decline in reproductive capacity.

The use of other approximations of ovarian functional decline with age, such as FSH, was also considered. Basal FSH levels, however, seemed not to be a good marker of the gradual loss of ovarian functional capacity with age and presumably only indicate advanced deterioration in ovarian function (18, 19). Likewise, one might have chosen to confine the study to women in several age categories who had delivered an infant within the 2 years before enrollment in the study. We would then have selected women who represent only the upper part of the reproductive performance scale. Still, one may wonder how the older women in our study would have performed if they had tested their natural fertility in a later stage of life. Because questions like these cannot be answered, it can be concluded that the study group assembled here is a better representation of the natural decline in reproductive capacity than earlier published cohorts.

The biphasic pattern of decline in antral follicle counts determined by the use of transvaginal US in this study is concordant with the results of autopsy studies. By the combination of histologic data, a model of follicle disappearance from birth to menopause was obtained (9). This model shows that total follicle numbers decline biexponentially with age. Whether this biphasic pattern is best described by two distinct linear relations or by one multiexponential curve has been debated recently (20). If the histologic data of the number of primordial follicles (21–23) are restricted to women of the same age group, as in our study (25–46 years), and reanalyzed by applying the same goodness-of-fit analysis for biphasic patterns as applied to our data set, a striking analogy is observed in the disappearance rate of follicles before the age of 37 years. In fact, the mean yearly disappearance rate in our study was 4.0% for primordial follicles and 4.8% for antral follicles. Thereafter, loss rates of both antral and primordial follicles increase, but the change is much smaller for the antral follicle disappearance pattern (Fig. 3).

This latter finding can be explained by the fact that the proportion of follicles leaving the pool of primordial follicles to develop into growing and antral follicles increases with advancing age, as shown in primates (9, 23–26) and mice (27, 28). As the total number of primordial follicles in the pool of so-called “resting” follicles declines with age, the absolute number of follicles entering the growth phase to become antral follicles decreases to a much lesser extent. From the point of view of fecundity, this may be considered as a salvage operation.

Our data showed that there was no difference in the median number of follicles between the right and left ovary. It is questionable whether antral follicle counts could be applicable in clinical testing in women with only one ovary. First, women who have had reductive surgery of the ovaries start menopause earlier, and the age at menopause is influenced by the age at the time of the operation (29). Second, to our knowledge there is no literature concerning changes in the disappearance rates of follicles in ovariectomized women.

Studies concerning the ovarian response to controlled ovarian hyperstimulation in IVF patients with only one ovary in terms of number of follicles, number of oocytes,
and pregnancy rates have reported conflicting results. One study showed that the ovarian response in women with one ovary was only slightly lower than the response in women with two ovaries, whereas pregnancy rates were similar in both groups (30). Finally, a histologic study performed in unilaterally ovariectomized mice revealed that the frequency of follicle death declined and the number of large preantral or antral follicles rose to approach those in pairs of age-matched control ovaries. The authors suggested that follicles otherwise undergoing atresia were being rescued (27). The same compensatory mechanisms are observed after unilateral ovariectomy in the bovine model (31).

We assumed that the number of antral follicles counted by transvaginal US reflects the remaining pool of primordial follicles within the pool of so-called “resting” follicles. This assumption is supported by the fact that both the number of primordial follicles and the age of the woman directly influence the number of antral follicles of >1 mm in women aged ≥38 years (21, 23, 24). In humans, the number of large growing follicles decreases with age, and a clear relation was found between the number of growing follicles and the primordial follicle pool (10). Finally, the analogy between the annual relative decline in primordial and US-based antral follicle count as shown in Figure 3 is another indication of this hypothesis. So, although there is no direct evidence that the number of antral follicles assessed by US reflects the pool of primordial follicles, it seems highly plausible.

The variation in the number of antral follicles within a certain age group appeared to be high, especially at the age of <40 years (Fig. 1). Still, the findings in this study allow further evaluation of follicle counting as a prognosticator of the probability of pregnancy in couples with unexplained infertility or of the outcome in couples referred for assisted reproductive procedures. In a study of IVF patients, it was found that the number of 2–5-mm antral follicles predicted the outcome of treatment (number of recovered oocytes) better than ovarian volume or age (32). Likewise, a good correlation was found between the antral follicle count and oocyte yield in a group of IVF and GIFT treatment cycles, but this study failed to show any additional predictive value with regard to pregnancy probability over FSH and age in a multivariate model (33). Published reports on the value of the antral follicle count for patient counseling during initial infertility workup are lacking at the moment. Studies addressing this question are currently under way.

The total volume of both ovaries was only weakly correlated with reproductive age in the study group \( R = -0.29, P = .003 \). This is in accordance with a study in IVF patients (24–46 years) that showed that ovarian size was weakly related to age \( R = -0.17 \) and that small ovaries (<3 cm³) in IVF patients could predict a poor ovarian response to gonadotropin stimulation (34). A significant negative correlation was found between ovarian volume and age \( R = -0.46 \) and between follicular density and age \( R = -0.43 \) in a group of 60 infertile women aged 19–45 years (35).

Two studies (32, 36) were not able to show an age-related decrease in the volume of both ovaries in IVF patients. The total ovarian volume did, however, correlate with outcomes after controlled ovarian hyperstimulation (peak E₂, number of oocytes, and number of embryos). So, although the relation between total ovarian volume and reproductive age is not clearly established, further study of the predictive value in infertility patients is warranted. It is often difficult to estimate the margins of the ovary reliably by the use of transvaginal US. Because total ovarian volume is calculated by using the formula for an ellipsoid \( [L \times B \times D \times \pi]/6 \), a small measurement error has a great influence on the computed volume.

Repeated cycle-to-cycle measurements of antral follicle number and total ovarian volume revealed only moderate agreement in every range of counts. It is not likely that the variation in antral follicle count in different cycles has its origin in measurement errors because it has been shown that the antral follicle count by transvaginal US is an accurate and reproducible measurement (13). A plausible explanation could be that the size of the growing follicle cohort varies among different menstrual cycles. This finding is important because it may influence the predictive capacity of the antral follicle count or ovarian volume with regard to the probability of pregnancy. In clinical studies addressing this question, repeated measurements should be performed to account for this problem. Measurements of follicle number and ovarian volume can, however, be done at various times of the cycle without changing the relation with reproductive status, as shown by the finding in this study that early follicular, midluteal, and late luteal estimates were all correlated with reproductive age to the same extent.

In summary, we conclude that the antral follicle count is clearly related to reproductive status, presumably because it represents the size of the remaining number of primordial follicles. For this reason, the use of antral follicle counting as a prognosticator of the probability of pregnancy in infertile patients and of the ovarian response to controlled ovarian hyperstimulation should be studied further.

References