Human Reproduction, Vol.25, No.1 pp. 221–227, 2010
Advanced Access publication on October 19, 2009 doi:10.1093/humrep/dep366

human reproduction

ORIGINAL ARTICLE *Reproductive endocrinology*

Comparison of inter- and intra-cycle variability of anti-Müllerian hormone and antral follicle counts

J. van Disseldorp[1,6], C.B. Lambalk[2], J. Kwee[3], C.W.N. Looman[4], M.J.C. Eijkemans[1,5], B.C. Fauser[1], and F.J. Broekmans[1]

[1]Department of Reproductive Medicine and Gynaecology, University Medical Center Utrecht, PO Box 85500, Room F.05.126, 3508 GA Utrecht, The Netherlands [2]Division of Reproductive Medicine, Department of Obstetrics and Gynaecology, VU University Medical Center (VUmc), PO Box 7057, 1007 MB Amsterdam, The Netherlands [3]Department of Obstetrics and Gynaecology, Sint Lucas Andreas Ziekenhuis, PO Box 9243, 1006 AE Amsterdam, The Netherlands [4]Department of Public Health, Erasmus Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands [5]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

[6]Correspondence address. Tel: +31-88-755-3629; Fax: +31-88-755-5433; E-mail: j.vandisseldorp-2@umcutrecht.nl

BACKGROUND: The antral follicle count (AFC) and anti-Müllerian hormone (AMH) both represent age-related follicular decline quite accurately, although long-term follow-up studies are still lacking. The best ovarian reserve test would need only a single, cycle-independent measurement to be representative.

METHODS: To compare the inter- and intra-cycle stability of AFC and AMH, we used age-adjusted intra-class correlation coefficients (ICCs). To measure inter-cycle stability across a number of up to four menstrual cycles, we used data, prospectively collected for the purpose of an other study, from 77 regularly cycling, infertile women aged 24–40 years. AMH and AFC values were measured on cycle day 3. To study intra-cycle variability, we used data from a prospective cohort study of 44 regularly cycling volunteers, aged 25–46 years and measured AMH and assessed the AFC (2–10 mm) every 1–3 cycle days.

RESULTS: Between menstrual cycles, AFC and AMH varied between 0 and 25 follicles (median 10), and 0.3 and 27.1 ng/ml (median 4.64). The difference in age-adjusted ICC between AMH [ICC, 0.89 (95% CI, 0.84–0.94)] and AFC [ICC, 0.71 (95% CI, 0.63–0.77)] was 0.18 (95% CI, 0.12–0.27). For the intra-cycle variation, 0–43 antral follicles (median 7) were counted per volunteer. The difference in age-adjusted ICC between AMH [ICC, 0.87 (95% CI, 0.82–0.91)] and AFC [ICC, 0.69 (95% CI, 0.46–0.82)] was 0.18 (95% CI, 0.034–0.42).

CONCLUSIONS: Serum AMH demonstrated less individual intra- and inter-cycle variation than AFCs and may therefore be considered a more reliable and robust means of assessing ovarian reserve in subfertile women.

Key words: antral follicle count / anti-Müllerian hormone / Müllerian inhibiting substance / menstrual cycle variability / measurement reliability

Introduction

Ovarian reserve tests aim to predict outcome of *in vitro* fertilization (IVF) treatment in terms of poor response and pregnancy. A recent meta-analysis has shown that multivariate models are comparable to single tests like the antral follicle count (AFC) in its capacity to predict ovarian response to stimulation for IVF (Verhagen *et al.*, 2008). It was concluded that the AFC may be considered the test of first choice when assessing diminished ovarian reserve.

Recent research reports that anti-Müllerian hormone (AMH) might be at least as good as the AFC in predicting response to controlled ovarian hyperstimulation in IVF (van Rooij *et al.*, 2005; McIlveen *et al.*, 2006; La Marca *et al.*, 2007; Broer *et al.*, 2008; Elgindy *et al.*, 2008; Kwee *et al.*, 2008). A similar accuracy of AMH and AFC in predicting ovarian response to hyperstimulation in IVF is not surprising, since AMH is produced by antral follicles up to the size of 6 mm (Weenen *et al.*, 2004). This follicle size class may well be associated with the AFC (2–5 or 2–10 mm in diameter) on ultrasound. Because of its production already in pre-antral follicle stages, AMH is suggested to represent the cohort of primordial follicles better (Durlinger *et al.*, 2002; Kevenaar *et al.*, 2006). With respect to age-related antral follicle decline in humans, AMH has been shown to be a better marker for the change in reproductive status over time, when compared with the AFC (van Rooij *et al.*, 2005).

As a laboratory test, AMH may have further advantages since assay variation is well documented, in contrast to the AFC (Broekmans et al., 2009). Moreover, most studies consider AMH to be cycle independent (Hehenkamp et al., 2006; La Marca et al., 2006; Tsepelidis et al., 2007), although others challenge this (Cook et al., 2000; Wunder et al., 2008). The AFC might be more prone to observer bias and show more variance between cycles in patients (Scheffer et al., 2002; Fanchin et al., 2005; van Rooij et al., 2005).

The ideal ovarian reserve test would only need one, preferably cycle independent, measurement to represent the ovarian reserve status. This study compares the AFC with the AMH both as to their inter-cycle variability across four subsequent cycles and their stability across a full cycle. For this purpose, the inter-class correlation coefficient and intra-class correlation coefficient (ICC) are calculated, which provide an estimate of the variation present both within the same woman and between different women.

Materials and Methods

Inter-cycle variation

To assess the inter-cycle variation for AFC and AMH, we used a study population previously described by Kwee et al. (2004). In brief, this study population is part of a prospective randomized study on the determination of ovarian reserve conducted at the VU University Medical Center, Amsterdam, The Netherlands. In the original study, patients were randomized to undergo a clomiphene citrate challenge test or an exogenous FSH ovarian reserve test in the early follicular phase of four menstrual cycles.

From June 1997 to May 1999, 85 patients aged 18–39 years who were eligible for intrauterine insemination (IUI) entered the study. From the original 85 subjects, 77 completed two or more cycles which made them eligible for the current analysis. Their infertility was either idiopathic for >3 years and/or due to a male factor and/or cervical hostility (negative well-timed post-coital test). Patients had to have two ovaries and regular menstrual cycles (between 21 and 35 days with the next cycle predictable within 7 days). Excluded were patients with either polycystic ovary syndrome diagnosed according to the Rotterdam consensus criteria (The Rotterdam ESHRE and ASRM-sponsored PCOS consensus workshop group, 2004) or a severe male factor. Severe male factor was defined as (i) $<1 \times 10^6$ motile spermatozoa after Percoll centrifugation (gradient 40/90); (ii) >20% antibodies present on the spermatozoa after processing with Percoll centrifugation (gradient 40/90); and/or (iii) >50% of the spermatozoa without an acrosome. Other exclusion criteria were untreated or insufficiently corrected endocrinopathies, clinically relevant systemic diseases or a body mass index >28 kg/m^2.

During the first three cycles, patients were treated with IUI; in the fourth cycle, patients underwent an IVF treatment. The IVF treatment followed the first IUI treatment within 1 year. Patients did not use contraceptive pills before IUI or IVF treatment. As an integral part of the study, AFCs were performed on cycle day 3 of every treatment cycle, before initiating ovarian reserve testing and treatment. All antral follicles of 2–10 mm diameter present in both ovaries were measured by calculating the mean of two perpendicular measurements on an Aloka SSD-1700 with 5.0 MHz probe and were counted by the same author (J.K.) as described previously (Kwee et al., 2007). All data were recorded in the patient file, using a standard form. Also blood was drawn and serum was frozen at −80°C for subsequent per batch measurements of serum AMH using an enzyme-immunometric assay (Diagnostic Systems Laboratories, Webster, TX, USA). Inter- and intra-assay coefficients of variation were <5% at the level of 3 ng/ml and <11% at the level of 13 ng/ml. Repeated freezing and thawing of the samples or storage at 37°C for 1 h have been shown not to affect the results of the assay (van Rooij et al., 2002).

The study protocol was approved by the Ethics Committee of research involving human subjects of the VU University Medical Center, Amsterdam, The Netherlands. Informed consent was signed by all the couples participating in the study.

Intra-cycle variation

To assess the intra-cycle variation of the AFC, we used a study population previously described by Scheffer et al. (1999). A study describing the intra-cycle variation of AMH was recently published using data from the same study population (Hehenkamp et al., 2006). The AMH data from this publication were used for comparison purposes only. Briefly, the study was conducted at the University Medical Center Utrecht, The Netherlands. A group of 44 healthy, regularly cycling, fertile, Caucasian female volunteers aged 25–46 years was recruited through advertisements in local newspapers. Volunteers were enrolled in the study protocol if they met all of the following criteria: (i) regular menstrual cycles, with mean length varying from 21 to 35 days; (ii) biphasic basal body temperature (BBT); (iii) proven natural fertility by having carried at least one pregnancy to term; (iv) each of the pregnancies established within 1 year after the interruption of contraceptive methods; (v) no evidence of endocrine disease; (vi) no history of ovarian surgery; (vii) no ovarian abnormalities, as assessed by vaginal ultrasound; and (viii) cessation of hormonal contraception 2 months before entering the study protocol.

Serial transvaginal ultrasound scans were performed by the same observer with a 7.5-MHz transvaginal probe on a Toshiba Capasee SSA-220A (Toshiba Medical Systems Europe BV, Zoetermeer, The Netherlands) as described previously (Scheffer et al., 1999). Measuring and counting follicles 2–10 mm was started in the mid-luteal phase of the first study cycle. The luteal phase was assumed to have started when a temperature rise on the BBT chart, based on classical criteria (WHO, 1967), had been observed. From the seventh day after the temperature shift onwards, the volunteers visited the clinic every 2 or 3 days for antral follicle measurement and blood sampling until the occurrence of the subsequent ovulation. Ovulation was registered by daily ultrasound scans for at least 4 days when the dominant follicle had reached a mean diameter of at least 14 mm. Ovulation day was defined as the day at which a complete disappearance of the follicle or a reduction of its mean diameter by at least 5 mm was observed (Janssen-Caspers et al., 1986; Check et al., 1990). The Institutional Review Board approved the study, and written informed consent was obtained from all participants. The volunteers received monetary compensation for participating. Inter- and intra-observer variations of the AFC were found to be low with inter- and intra-class coefficients between 0.98 and 0.99, indicating high reproducibility (Scheffer et al., 2002).

Statistical analysis

For the analysis of inter-cycle fluctuations, we visualized the available AMH and AFC values per cycle in box plots and analysed if they differed significantly using linear mixed models. To assess within-subject reproducibility of the AFC and AMH results, we calculated the ICC and its 95% confidence intervals (CI) (Shrout and Fleiss, 1979). We chose to calculate ICCs, since they distinguish between variation within the same woman and between individual women. We adjusted the ICC for woman's age, since AMH and AFC decline are age-dependent. The difference in the ICCs between AFC and AMH and the 95% CI of this difference was assessed using a bootstrap procedure with 2000 replications.

For the analysis of intra-cycle fluctuations, we defined seven cycle phases as a range of days counted from either menstruation (M) or

from the ultrasound assessed ovulation (O) day in cycle 2 as described earlier (Hehenkamp *et al*., 2006). The seven cycle phases were defined as follows: mid-luteal, M−9 to M−5; late luteal, M−4 to M−1; early follicular, M to M+4; mid-follicular, O−9 to O−6; late follicular, O−5 to O−2; peri-ovulation, O−1 to O+1; and early luteal, O+2 to O+4.

To visualize the intra-cycle variability of AMH and AFC, all available AMH and AFC values per cycle phase were averaged and box plots were constructed using these data.

We again calculated the age-adjusted ICC and its 95% CI to assess within-subject reproducibility of the AFC and AMH results. The relationship between two different continuous variables was assessed by correlation coefficient.

As the size classes of follicles may affect their clinical significance (Pohl *et al*., 2000; Haadsma *et al*., 2007), all intra-cycle variability analyses were done for AFC 2–5 mm as well as for AFC 2–10 mm.

For the intra-cycle variability, we evaluated the size of the effect of intra-individual fluctuations by classifying values of the AFC in five quintiles and registering how often two paired measurements of an individual were located in the same quintile, in adjacent quintiles or in non-adjacent quintiles. The cut-off levels used for the analysis of quintile categories for AFC 2–5 mm were 0.5, 2, 4, 6, 9 and 37, respectively (corresponding to 0%, 20%, 40%, 60%, 80% and 100%). Cut-off levels for AFC 2–10 mm were 0.5, 3, 6, 9, 12 and 43.

Statistical analyses were performed by using the linear mixed-effects model in SPlus (version 6.0; Mathsoft Inc., Seattle, WA, USA) and with SPSS version 15.1 (SPSS Inc., Chicago, IL, USA).

Results

The 77 women eligible for the inter-cycle variation analysis had a median age of 33 (24–40) years and completed on average 3.73 cycles (83% completed 3 cycles, 77% completed 4 cycles). The AFC and AMH levels in the population varied between 0 and 25 follicles (median 10 follicles), and 0.3 and 27.1 ng/ml (median 4.64 ng/ml), respectively. The mean AMH and AFC levels per cycle were not statistically different over the four consecutive cycles and ranged between 5.7 and 6.0 ng/ml, and 9.1 and 10.4 follicles, respectively (Fig. 1). AMH correlated positively and significantly with the AFC with correlation coefficients of 0.70, 0.66, 0.68 and 0.63 in each of the four respective cycles ($P < 0.001$ for all four cycles). The age-adjusted ICCs for AMH and AFC across four cycles were 0.89 (95% CI, 0.84–0.94) and 0.71 (95% CI, 0.63–0.77), respectively. The difference was shown to be statistically significant [ICC_{diff}, 0.18 (95% CI, 0.12–0.27)]. In other words, since 89% (95% CI, 84–94%) of the variation in AMH can be attributed to between-subject variation, only 11% (95% CI, 6–16%) is true individual cycle fluctuation. For the AFC, 71% (95% CI, 63–77%) of the variation can be attributed to between-subject variation and 29% (95% CI, 23–37%) is individual cycle variation.

The 44 females volunteering in the intra-cycle variation analysis had a median age of 38.3 years (25.6–46.2 years), a median AFC 2–5 mm of 5 follicles (range 0–37 follicles) and a median AFC 2–10 mm of 7 follicles (range 0–43 follicles). On average, 9.4 (range 5–16) AFCs were made per volunteer. The correlation of the AFCs between the different cycle phases was modest, with a correlation coefficient of 0.43 and 0.58 for AFCs 2–5 and 2–10 mm, respectively (Fig. 2). The age-adjusted ICCs for AFC across the seven cycle phases were 0.66 (95% CI, 0.41–0.82) and 0.69 (95% CI, 0.46–0.82) for AFCs 2–5 and 2–10 mm, respectively. In other words, respectively, 66% (95% CI, 41–82%) and 69% (95% CI, 46–82%) of AFC variation can be attributed to between-subject variation. Within the same subject, AFCs varied 34% (95% CI, 18–59%) and 31% (95% CI, 18–54%) for AFCs 2–5 and 2–10 mm, respectively.

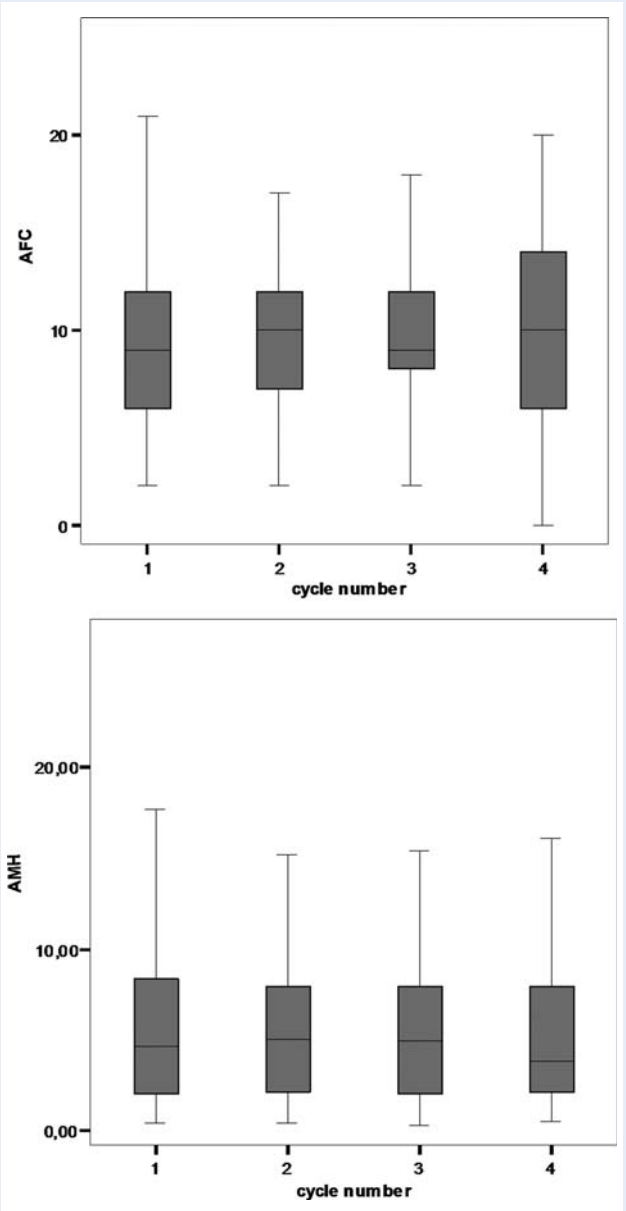

Figure 1 Box plots depicting distribution of AMH levels (in ng/ml) and AFCs (of all follicles 2–10 mm in both ovaries) in the early follicular phase across four cycles.

In the analysis of the clinical effect of intra-individual cycle fluctuations, it appeared that (Table I) the paired AFCs were located in the same quintile in 41% and 45% of the cases for the AFCs 2–5 and 2–10 mm, respectively. Paired measurements of AFCs 2–5 and 2–10 mm crossed over two or more quintiles in 21% and 16% of the cases, respectively.

To compare the intra-cycle variability of AFC with AMH, we have added the previously published box plot depicting AMH levels over the seven cycle phases in Fig. 2 (Hehenkamp *et al*., 2006). From

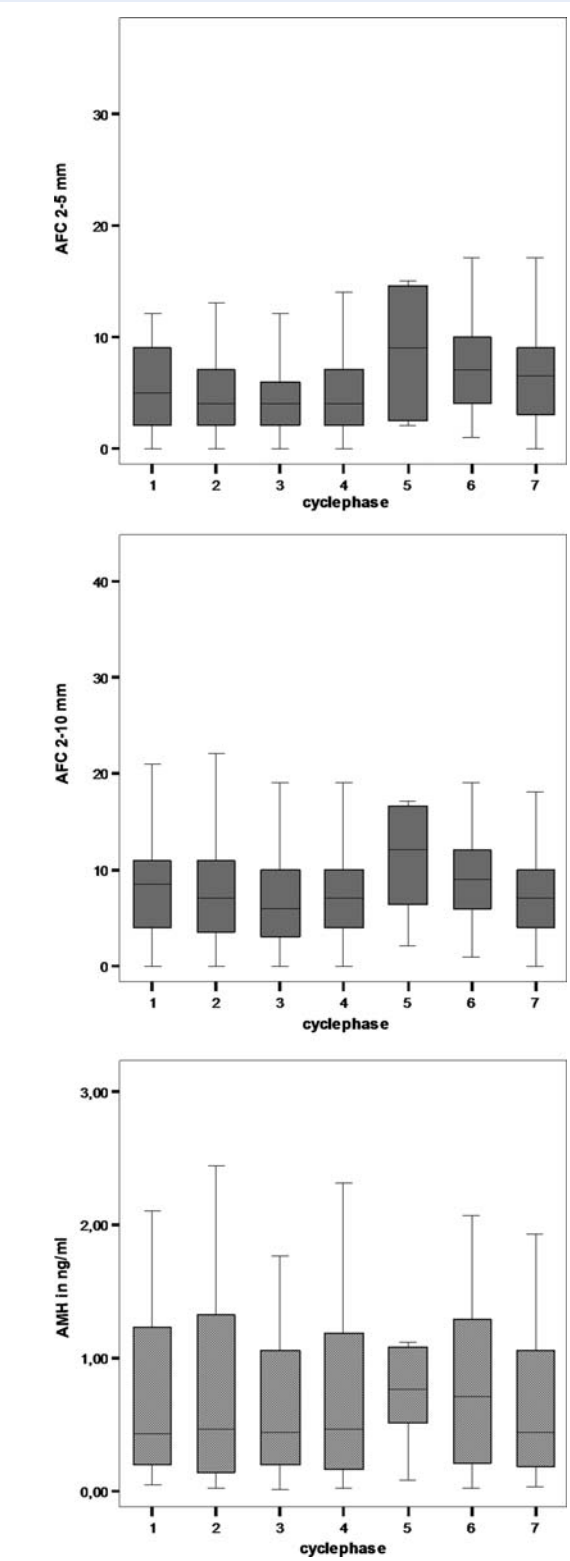

Figure 2 Distribution of AFCs 2–5 and 2–10 mm with correlation coefficients across the seven cycle phases of 0.43 and 0.58, respectively. Cycle phases are depicted as (1) early follicular ($n = 58$), (2) mid-follicular ($n = 32$), (3) late follicular ($n = 81$), (4) peri-ovulatory ($n = 106$), (5) early luteal ($n = 8$), (6) mid-luteal ($n = 42$) and (7) late luteal ($n = 66$). Mean AFCs did not differ significantly between cycle phases. Please note that the variation in the early luteal phase (5) is quite large due to few observations, with an overrepresentation of younger volunteers. The AMH distribution previously published is presented for comparison. (Reprint with permission from Hehenkamp *et al.*, 2006.)

Table I Intra-individual fluctuations of the AFC, showing the percentage of random paired measurements that fall in the same (0), adjacent (1) or non-adjacent (2–4) quintile categories

Quintile	0	1	2	3	4
AFC 2–5 mm	41.2%	38.3%	13.0%	5.5%	2.0%
AFC 2–10 mm	45.5%	38.7%	12.3%	2.6%	1.0%

these data, we have calculated the ICC for AMH [ICC, 0.87 (95% CI, 0.82–0.91)], demonstrating 13% (95% CI, 9–18%) within-subject variation. The difference in ICC was shown to be statistically significant both for the AFC 2–5 mm [ICC_{diff}, 0.21 (95% CI, 0.046–0.47)] and for the AFC 2–10 mm [ICC_{diff}, 0.18 (95% CI, 0.034–0.42)]. Furthermore, for AMH, intra-individual fluctuations were shown to fall in the same quintile in 72% of the cases and to cross two quintiles in only 1% of the cases (Hehenkamp *et al.*, 2006).

Discussion

This study describes both the intra- and inter-individual variability of AMH and AFC across a number of up to four consecutive cycles. Although both AMH and AFC on a group level did not differ significantly between the various cycles, the intra-individual variation in AMH levels appeared significantly smaller than for the AFC. These findings are consistent with a previous comparative study (Fanchin *et al.*, 2005), where ICCs for AMH and AFC across three cycles were 0.89 (95% CI, 0.83–0.94) and 0.73 (95% CI, 0.66–0.86), respectively. Our validation across four cycles renders these results even more robust. Others (McIlveen *et al.*, 2006; Streuli *et al.*, 2008) studied the inter-cycle variability of AMH across two cycles and concluded independently that AMH varies little between cycles. Inter-cycle variability of AFC alone has also been studied previously and was shown to be moderate in most studies (Scheffer *et al.*, 1999; Hansen *et al.*, 2003; Bancsi *et al.*, 2004; Elter *et al.*, 2005).

During the four menstrual cycles studied, AMH levels showed a positive correlation with AFC levels, indicating that a substantial proportion of antral follicles must contribute to AMH serum levels (de Vet *et al.*, 2002; van Rooij *et al.*, 2002; Fanchin *et al.*, 2003). Moreover, coefficients of correlation were reasonably constant over the four cycles ranging from 0.63 to 0.70, which is in concordance with the current concept that AMH is produced by a steady pool of small (pre)antral follicles (Baarends *et al.*, 1995; Weenen *et al.*, 2004).

Not all patients completed four cycles. All patients that dropped out after cycle 2 or 3 (18 out of 77) were pregnant. Age, AMH and AFC were not significantly different between those that dropped out or completed four cycles. Because of this, we believe that the influence on the final result is small.

A possible explanation for the higher variability of AFC is the reproducibility and standardization of the AFC itself (Broekmans *et al*., 2009). Jayaprakasan *et al*. have shown that the 2D real-time ultrasound method does not generate significant differences in AFCs compared with 3D ultrasound *post hoc* processing. However, 3D ultrasound may allow for improvement in measurement reproducibility (Jayaprakasan *et al*., 2007, 2008). Moreover, ICC and inter-class correlation coefficient found for intra- and inter-observer variations have varied between 0.96 and 0.99 for 2D and 3D AFC measurements (Scheffer *et al*. 2002; Jayaprakasan *et al*. 2008). However, limits of agreement in intra- and inter-observer variations are wide, exceeding the variation commonly found in AMH assays, as reported in the Materials and methods section. Since all AFC measurements were made by the same experienced observer on the same machine in both parts of the present study, observer variation in our data has thereby been minimized. With an expected increase in variation among different observers, the variation in the AFC may further increase.

The higher stability of AMH measurements may also be explained by assuming that AMH levels are also determined by a cohort of invisible pre-antral or small antral follicles, whereas the number of larger and visible antral follicles, expressed by the AFC, may be more prone to short-term variation. Possible explanations for a varying cohort of antral follicles might be cyclic differences in decay or growth rate which may depend on the presence of larger follicles in the early follicular phase (Grynberg *et al*., 2008; Hansen and Soules, 2008).

For the whole study population, the AFC did not differ significantly between the various cycle phases. However, intra-individual variation appeared to be moderate as demonstrated by the ICC, determining individual intra-cycle consistency of AFC. Compared with the intra-individual, intra-cycle variation of AMH, calculated from data previously published (Hehenkamp *et al*., 2006), the AFC was indeed less consistent. This demonstrates that AMH has superior performance compared with both AFC size classes, when individual cycle consistency is concerned.

The intra-cycle stability of AMH has already been described in several studies (Cook *et al*., 2000; Hehenkamp *et al*., 2006; La Marca *et al*., 2006; Tsepelidis *et al*., 2007; Wunder *et al*., 2008). Most of these studies found very stable levels of AMH throughout the cycle, except two studies, which found a late follicular or periovulatory rise in AMH (Cook *et al*., 2000; Wunder *et al*., 2008). An obvious explanation for this difference, such as study setup or patient age, could not justify this difference. Our relative (non-significant) AFC and AMH rise (Fig. 2) in the early luteal phase is caused by relatively few measurements in young individuals (Hehenkamp *et al*., 2006). The intra-cycle variability of AFC has not been extensively studied before (Pache *et al*., 1990). For the first time, this feature has now been analysed for two size classes of AFC over seven cycle phases. The box plots in Fig. 2 show modest intra-cycle differences in mean AFC and ICCs demonstrate that this variation is primarily based on inter-individual differences. In agreement with the variation depicted in Fig. 2, multilevel analysis of variance failed to show large differences between the seven cycle phases (unpublished data). This suggests that the intra-cycle fluctuations of AFC may be of limited biological importance and may be considered chance findings or caused by observer bias, as discussed earlier.

For the purpose of clinical application as a test for ovarian reserve, we analysed the amount of quintile transgression. Only 41.2% and 45.5% of AFCs of 2–5 and 2–10 mm, respectively, remained in the same quintile, whereas 21% and 16%, respectively, crossed two or more quintiles. In a recent study, AMH was shown to cross two or more quintiles in only 1% of the cases (Hehenkamp *et al*., 2006). This quintile transgression is probably due to the large deviations from the mean AFC per cycle phase, caused by high amounts of variation in patients, particularly at low AFCs. This difference in quintile transgression between AMH and AFC, together with the lower intra-individual intra- and inter-cycle variability of AMH, indicates that AMH may be the better cycle-independent test for prediction of ovarian reserve. Assessed irrespective of the cycle phases, misclassification in relevant diagnostic categories for AMH seems unlikely. This has already been shown in clinical studies in ART populations (La Marca *et al*., 2007). Obviously, this study was not designed to study the performance of AMH and AFC to predict poor response or pregnancy. Future prospective studies specifically designed for this purpose are needed to test this hypothesis. As for the AFC, a switch towards random chosen moments of ultrasound-based counting of antral follicles seems not justified, especially in view of the moderate consistency even at lower counts.

When comparing inter- with intra-cycle variation of AMH, a recent study found the inter-cycle variability of AMH to be 28% (95% CI, −23.2 to 80.3%), thus larger (Streuli *et al*., 2008) than the amplitude of intra-cycle AMH fluctuations from previous studies, which all fell below this 28% (Hehenkamp *et al*., 2006; Wunder *et al*., 2008; Streuli *et al*., 2009). This suggests that in clinical practice, AMH can be measured independent of the cycle, as shown by La Marca *et al*. (2007). However, caution seems justified when AMH is used for estimating ovarian reserve status in young women, as fluctuations may be more considerable and misclassifications more likely (Hehenkamp *et al*., 2006). We feel the conclusions regarding intra-cycle variation in the fertile population can be extrapolated to an infertility population. First, we corrected for female age, allowing the information to be applied to a population with a different age distribution. Second, the alleged differences between the fertile and infertile population may be small, as the proven fertility of the fertile population could be dated years earlier and consequently did not always specify the fertility status at the time of assessment. Also, AMH levels may vary according to the cause of subfertility, especially for endometriosis and PCOS cases (Pigny *et al*., 2003; Al-Qahtani *et al*., 2005; Falconer *et al*., 2009). However, whether the intra-cycle variability of AMH and AFC differs for the various causes of subfertility is currently unknown.

There is still substantial variation in literature concerning the size of antral follicles to be measured and used for clinical decision-making. Current data indicate that the AFC of 2–5 mm shows more fluctuation than the AFC of 2–10 mm. Moreover, the AFC of 2–5 mm showed substantially more quintile transgression than did that of 2–10 mm. A possible explanation might be that the accuracy of follicular measurements has been shown to increase with follicular diameter (Pache *et al*., 1990). Therefore, we advocate the use of the AFC of 2–10 mm over that of 2–5 mm. Earlier studies regarding size classes of follicles have shown some conflicting results. In one study, the follicle size class 2–6 mm appeared clearly related to female age, whereas the follicle size class 7–10 mm did not (Haadsma *et al*., 2007). In an other study, the follicle size class 6–10 mm correlated very clearly with oocyte yield after IVF which is highly correlated with age, whereas the follicle size class 2–5 mm did not (Pohl *et al*., 2000).

The current study is a secondary data analysis of prospective data, collected for the purpose of two other studies, with inherent limitations. First, the current study was not designed to study the clinical implications of cycle variation on the accuracy of poor response and pregnancy prediction. Second, the current design does not allow the comparison of inter- and intra-cycle variations, because both populations have a different fertility history. We feel, however, that the current design is appropriate to test our hypothesis concerning cycle fluctuation.

Future research should focus on improving reliable measurement techniques such as operator independent automated AFCs (Deb *et al*., 2009). Furthermore, the performance of random AFCs to classify expected poor responders or hyper-responders should also be investigated. Finally, daily measurements of AMH and AFC, aligned with the LH peak, could give more insight into subtle fluctuations within the menstrual cycle. This might show that possibly only younger women exhibit relevant cycle variation.

In conclusion, although AFC and AMH do not seem to differ between and within cycles at first sight (Figs 1 and 2), we demonstrated that AMH displays less intra-individual fluctuation than the AFC both within and between cycles. This suggests AMH to be the better cycle-independent parameter to assess ovarian reserve, but future prospective studies specifically designed for this purpose are needed to confirm this hypothesis. If the AFC is used for screening ovarian reserve, it is advocated to count follicles in the size range 2–10 mm, in view of better cycle stability. Previous studies have shown intra- and inter-observer variations to be small for AFC, but with wide limits of agreement, exceeding the variation commonly found for AMH assays.

Author's role

J.v.D.: (i) substantial contributions to conception and design, analysis and interpretation of data, (ii) drafting the article and (iii) final approval of the version to be published. C.B.L.: (i) substantial contributions to analysis and interpretation of data, (ii) revising the manuscript critically for important intellectual content and (iii) final approval of the version to be published. J.K.: (i) substantial contributions to acquisition of data, analysis and interpretation of data, (ii) revising the manuscript critically for important intellectual content, and (iii) final approval of the version to be published. C.W.N.L. and M.J.C.E.: (i) substantial contributions to analysis and interpretation of data, (ii) revising the manuscript critically for important intellectual content, and (iii) final approval of the version to be published. B.C.F.: (i) substantial contributions to conception and design, interpretation of data, (ii) revising the manuscript critically for important intellectual content, and (iii) final approval of the version to be published. F.J.B.: (i) substantial contributions to conception and design, acquisition of data and interpretation of data, (ii) drafting the article and revising it critically for important intellectual content, and (iii) final approval of the version to be published.

Conflict of interest: B.C.F. has received fees and grant support from the following companies: Andromed, Ardana, Ferring, Merck Serono, Organon, Pantharei Bioscience, PregLem, Schering Plough, Schering, Serono and Wyeth. C.B.L. has received fees and grant support from the following companies: Ferring, Merck Serono and Schering Plough. F.J.B. is a member of the external advisory board for Ferring Pharmaceuticals, Hoofddorp, The Netherlands. He receives no monetary compensation.

References

Al-Qahtani A, Muttukrishna S, Appasamy M, Johns J, Cranfield M, Visser JA, Themmen AP, Groome NP. Development of a sensitive enzyme immunoassay for anti-Mullerian hormone and the evaluation of potential clinical applications in males and females. *Clin Endocrinol (Oxf)* 2005;**63**:267–273.

Baarends WM, Uilenbroek JT, Kramer P, Hoogerbrugge JW, van Leeuwen EC, Themmen AP, Grootegoed JA. Anti-mullerian hormone and anti-mullerian hormone type II receptor messenger ribonucleic acid expression in rat ovaries during postnatal development, the estrous cycle, and gonadotropin-induced follicle growth. *Endocrinology* 1995;**136**:4951–4962.

Bancsi LF, Broekmans FJ, Looman CW, Habbema JD, te Velde ER. Impact of repeated antral follicle counts on the prediction of poor ovarian response in women undergoing in vitro fertilization. *Fertil Steril* 2004; **81**:35–41.

Broekmans FJ, de Ziegler D, Howles CM, Gougeon A, Trew G, Olivennes F. The antral follicle count: practical recommendations for better standardization. *Fertil Steril* 2009 (in press).

Broer SL, Mol BW, Hendriks D, Broekmans FJ. The role of antimullerian hormone in prediction of outcome after IVF: comparison with the antral follicle count. *Fertil Steril* 2009;**91**:705–714.

Check JH, Adelson HG, Dietterich C, Stern J. Pelvic sonography can predict ovum release in gonadotrophin-treated patients as determined by pregnancy rate. *Hum Reprod* 1990;**5**:234–236.

Cook CL, Siow Y, Taylor S, Fallat ME. Serum mullerian-inhibiting substance levels during normal menstrual cycles. *Fertil Steril* 2000;**73**:859–861.

de Vet A, Laven JS, de Jong FH, Themmen AP, Fauser BC. Antimullerian hormone serum levels: a putative marker for ovarian aging. *Fertil Steril* 2002;**77**:357–362.

Deb S, Jayaprakasan K, Campbell BK, Clewes JS, Johnson IR, Raine-Fenning NJ. Intraobserver and interobserver reliability of automated antral follicle counts made using three-dimensional ultrasound and SonoAVC. *Ultrasound Obstet Gynecol* 2009;**33**:477–483.

Durlinger AL, Visser JA, Themmen AP. Regulation of ovarian function: the role of anti-Mullerian hormone. *Reproduction* 2002;**124**:601–609.

Elgindy EA, El-Haieg DO, El-Sebaey A. Anti-Mullerian hormone: correlation of early follicular, ovulatory and midluteal levels with ovarian response and cycle outcome in intracytoplasmic sperm injection patients. *Fertil Steril* 2008;**89**:1670–1676.

Elter K, Sismanoglu A, Durmusoglu F. Intercycle variabilities of basal antral follicle count and ovarian volume in subfertile women and their relationship to reproductive aging: a prospective study. *Gynecol Endocrinol* 2005;**20**:137–143.

Falconer H, Sundqvist J, Gemzell-Danielsson K, von SB, D'Hooghe TM, Fried G. IVF outcome in women with endometriosis in relation to tumour necrosis factor and anti-Mullerian hormone. *Reprod Biomed Online* 2009;**18**:582–588.

Fanchin R, Schonauer LM, Righini C, Guibourdenche J, Frydman R, Taieb J. Serum anti-Mullerian hormone is more strongly related to ovarian follicular status than serum inhibin B, estradiol, FSH and LH on day 3. *Hum Reprod* 2003;**18**:323–327.

Fanchin R, Taieb J, Lozano DH, Ducot B, Frydman R, Bouyer J. High reproducibility of serum anti-Mullerian hormone measurements suggests a multi-staged follicular secretion and strengthens its role in

the assessment of ovarian follicular status. *Hum Reprod* 2005; **20**:923–927.

Grynberg M, Feyereisen E, Scheffer JB, Koutroubis P, Frydman R, Fanchin R. Early follicle development alters the relationship between antral follicle counts and inhibin B and follicle-stimulating hormone levels on cycle day 3. *Fertil Steril* 2008 (in press).

Haadsma ML, Bukman A, Groen H, Roeloffzen EM, Groenewoud ER, Heineman MJ, Hoek A. The number of small antral follicles (2–6 mm) determines the outcome of endocrine ovarian reserve tests in a subfertile population. *Hum Reprod* 2007;**22**:1925–1931.

Hansen KR, Soules MR. Counting ovarian follicles is not without its challenges. *Fertil Steril* 2008;**89**:1028–1029.

Hansen KR, Morris JL, Thyer AC, Soules MR. Reproductive aging and variability in the ovarian antral follicle count: application in the clinical setting. *Fertil Steril* 2003;**80**:577–583.

Hehenkamp WJ, Looman CW, Themmen AP, de Jong FH, te Velde ER, Broekmans FJ. Anti-Mullerian hormone levels in the spontaneous menstrual cycle do not show substantial fluctuation. *J Clin Endocrinol Metab* 2006;**91**:4057–4063.

Janssen-Caspers HA, Kruitwagen RF, Wladimiroff JW, de Jong FH, Drogendijk AC. Diagnosis of luteinized unruptured follicle by ultrasound and steroid hormone assays in peritoneal fluid: a comparative study. *Fertil Steril* 1986;**46**:823–827.

Jayaprakasan K, Walker KF, Clewes JS, Johnson IR, Raine-Fenning NJ. The interobserver reliability of off-line antral follicle counts made from stored three-dimensional ultrasound data: a comparative study of different measurement techniques. *Ultrasound Obstet Gynecol* 2007; **29**:335–341.

Jayaprakasan K, Campbell BK, Clewes JS, Johnson IR, Raine-Fenning NJ. Three-dimensional ultrasound improves the interobserver reliability of antral follicle counts and facilitates increased clinical work flow. *Ultrasound Obstet Gynecol* 2008;**31**:439–444.

Kevenaar ME, Meerasahib MF, Kramer P, van de Lang-Born BM, de Jong FH, Groome NP, Themmen AP, Visser JA. Serum anti-mullerian hormone levels reflect the size of the primordial follicle pool in mice. *Endocrinology* 2006;**147**:3228–3234.

Kwee J, Schats R, McDonnell J, Lambalk CB, Schoemaker J. Intercycle variability of ovarian reserve tests: results of a prospective randomized study. *Hum Reprod* 2004;**19**:590–595.

Kwee J, Elting ME, Schats R, McDonnell J, Lambalk CB. Ovarian volume and antral follicle count for the prediction of low and hyper responders with in vitro fertilization. *Reprod Biol Endocrinol* 2007;**5**:9.

Kwee J, Schats R, McDonnell J, Themmen A, de Jong F, Lambalk C. Evaluation of anti-Mullerian hormone as a test for the prediction of ovarian reserve. *Fertil Steril* 2008;**90**:737–743.

La Marca A, Stabile G, Artenisio AC, Volpe A. Serum anti-Mullerian hormone throughout the human menstrual cycle. *Hum Reprod* 2006; **21**:3103–3107.

La Marca A, Giulini S, Tirelli A, Bertucci E, Marsella T, Xella S, Volpe A. Anti-Mullerian hormone measurement on any day of the menstrual cycle strongly predicts ovarian response in assisted reproductive technology. *Hum Reprod* 2007;**22**:766–771.

McIlveen M, Skull JD, Ledger WL. Evaluation of the utility of multiple endocrine and ultrasound measures of ovarian reserve in the prediction of cycle cancellation in a high-risk IVF population. *Hum Reprod* 2006;**22**:778–785.

Pache TD, Wladimiroff JW, de Jong FH, Hop WC, Fauser BC. Growth patterns of nondominant ovarian follicles during the normal menstrual cycle. *Fertil Steril* 1990;**54**:638–642.

Pigny P, Merlen E, Robert Y, Cortet-Rudelli C, Decanter C, Jonard S, Dewailly D. Elevated serum level of anti-mullerian hormone in patients with polycystic ovary syndrome: relationship to the ovarian follicle excess and to the follicular arrest. *J Clin Endocrinol Metab* 2003; **88**:5957–5962.

Pohl M, Hohlagschwandtner M, Obruca A, Poschalko G, Weigert M, Feichtinger W. Number and size of antral follicles as predictive factors in vitro fertilization and embryo transfer. *J Assist Reprod Genet* 2000; **17**:315–318.

Scheffer GJ, Broekmans FJ, Dorland M, Habbema JD, Looman CW, te Velde ER. Antral follicle counts by transvaginal ultrasonography are related to age in women with proven natural fertility. *Fertil Steril* 1999; **72**:845–851.

Scheffer GJ, Broekmans FJ, Bancsi LF, Habbema JD, Looman CW, te Velde ER. Quantitative transvaginal two- and three-dimensional sonography of the ovaries: reproducibility of antral follicle counts. *Ultrasound Obstet Gynecol* 2002;**20**:270–275.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;**2**:420–428.

Streuli I, Fraisse T, Pillet C, Ibecheole V, Bischof P, de Ziegler D. Serum antimullerian hormone levels remain stable throughout the menstrual cycle and after oral or vaginal administration of synthetic sex steroids. *Fertil Steril* 2008;**90**:395–400.

Streuli I, Fraisse T, Chapron C, Bijaoui G, Bischof P, de Ziegler D. Clinical uses of anti-Mullerian hormone assays: pitfalls and promises. *Fertil Steril* 2009;**91**:226–230.

The Rotterdam ESHRE and ASRM-sponsored PCOS consensus workshop group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril* 2004; **81**:19–25.

Tsepelidis S, Devreker F, Demeestere I, Flahaut A, Gervy C, Englert Y. Stable serum levels of anti-Mullerian hormone during the menstrual cycle: a prospective study in normo-ovulatory women. *Hum Reprod* 2007;**22**:1837–1840.

van Rooij IA, Broekmans FJ, te Velde ER, Fauser BC, Bancsi LF, de Jong FH, Themmen AP. Serum anti-Mullerian hormone levels: a novel measure of ovarian reserve. *Hum Reprod* 2002;**17**:3065–3071.

van Rooij IA, Broekmans FJ, Scheffer GJ, Looman CW, Habbema JD, de Jong FH, Fauser BJ, Themmen AP, te Velde ER. 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.

Verhagen TE, Hendriks DJ, Bancsi LF, Mol BW, Broekmans FJ. The accuracy of multivariate models predicting ovarian reserve and pregnancy after in vitro fertilization: a meta-analysis. *Hum Reprod Update* 2008;**14**:95–100.

Weenen C, Laven JS, Von Bergh AR, Cranfield M, Groome NP, Visser JA, Kramer P, Fauser BC, Themmen AP. Anti-Mullerian hormone expression pattern in the human ovary: potential implications for initial and cyclic follicle recruitment. *Mol Hum Reprod* 2004;**10**: 77–83.

World Health Organisation (WHO). Biology of fertility control by periodic abstinence. WHO Technical Report Series 360. Geneva; World Health Organisation; 1967.

Wunder DM, Bersinger NA, Yared M, Kretschmer R, Birkhauser MH. Statistically significant changes of antimullerian hormone and inhibin levels during the physiologic menstrual cycle in reproductive age women. *Fertil Steril* 2008;**89**:927–933.

Submitted on June 9, 2009; resubmitted on September 1, 2009; accepted on September 17, 2009