Prospective study into the value of the automated Elecsys antimüllerian hormone assay for the assessment of the ovarian growing follicle pool

Citation for published version:

Digital Object Identifier (DOI):
10.1016/j.fertnstert.2015.01.004

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published in:
Fertility and Sterility

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Prospective study into the value of the automated Elecsys antimüllerian hormone assay for the assessment of the ovarian growing follicle pool

Richard A. Anderson, M.D., Ph.D., Ellen Anckaert, M.D., Ph.D., Ernesto Bosch, M.D., Didier Dewailly, M.D., Cheryl E. Dunlop, M.D., Daniel Fehr, M.D., Luciano Nardo, M.D., Johan Smitz, M.D., Ph.D., Kelton Tremellen, M.D., Ph.D., Barbara Denk, Ph.D., Andrea Geistanger, Ph.D., and Martin Hund, Ph.D.

Objective: To evaluate a new fully automated assay measuring antimüllerian hormone (AMH; Roche Elecsys) against antral follicle count in women of reproductive age.

Design: Prospective cohort study.

Setting: Hospital infertility clinics and academic centers.

Patient(s): Four hundred fifty-one women aged 18 to 44 years, with regular menstrual cycles.

Intervention(s): None.

Main Outcome Measure(s): AMH and antral follicle count (AFC) determined at a single visit on day 2–4 of the menstrual cycle.

Result(s): There was a statistically significant variance in AFC but not in AMH between centers. Both AFC and AMH varied by age (overall Spearman rho = −0.50 for AFC and −0.47 for AMH), but there was also significant between-center variation in the relationship between AFC and age but not for AMH. There was a strong positive correlation between AMH and AFC (overall spearman rho 0.68), which varied from 0.49 to 0.87 between centers. An agreement table using AFC cutoffs of 7 and 15 showed classification agreement in 63.2%, 56.9% and 74.5% of women for low, medium, and high groups, respectively.

Conclusion(s): The novel fully automated Elecsys AMH assay shows good correlations with age and AFC in women of reproductive age, providing a reproducible measure of the growing follicle pool. (Fertil Steril® 2015;103:1074–80. ©2015 by American Society for Reproductive Medicine.)

Key Words: Antimüllerian hormone, antral follicle count, ovarian follicle, ovarian reserve, reproductive life span

Discuss: You can discuss this article with its authors and with other ASRM members at http://fertstertforum.com/anderson-elecsys-amb-ovarian-follicle/
biomarkers that reflect follicular activity and thus the 
response to gonadotropins during subsequent ovarian stimula-
tion. To achieve maturity during the time course of stimula-
tion, ovarian follicles must already be at the small antral stage of
development, so clinically valuable biomarkers will accu-
rately quantify the number of such growing follicles, often 
termed the ovarian reserve. Although several biomarkers have 
been investigated over the years (including follicle-stimulating hormone [FSH], estradiol, and inhibin B), two 
have shown markedly greater accuracy and are in widespread 
use: antral follicle count (AFC) and antimüllerian hormone 
(AMH) (1–4). These have been extensively investigated in 
women undergoing assisted conception but have much 
broader application in women's health across the 
reproductive life span.

Transvaginal ultrasound is used to determine AFC, identi-
fying and counting the number of small antral follicles. 
Developments in technology have improved image resolution, 
allowing follicles ≥2 mm to be readily visualized; follicles up to 10 mm are generally included in the analysis. 
The normal range is currently a matter of controversy (5).
Although prediction of pregnancy is poor, AFC shows 
good prediction of the number of oocytes that will be 
retrieved after stimulation (3, 6). Its immediacy and wide 
availability are significant advantages. However, although 
automated determination of AFC is being developed and 
protocols for standardization are described (7), it remains 
an essentially subjective measure affected by the operator, 
the equipment used, and the patient because factors such 
as high body mass index (BMI) and pelvic pathology may 
impact the result.

Measurement of circulating AMH also reflects the num-
ber of small antral follicles and is predictive of ovarian response (1, 6). Although AMH is produced by the 
granulosa cells of follicles from the earliest stages of growth 
through the preantral and early antral stages, production declines abruptly at the stage at which follicles are selected for 
dominance, 8–10 mm (8). In normal women, the population of follicles of 5–8 mm diameter produces most 
circulating AMH (9). In comparison with AFC, AMH shows 
good prediction of oocyte number and similarly limited 
prediction of pregnancy and live birth (4, 6, 10). As a 
biochemical test, there are potential advantages in 
standardization and thus consistency in results both within 
and between centers, but this has yet to be realized. The 
currently available assays are manual, plate-based enzyme-linked immunosorbent assays (ELISAs); although these have 
led to a wealth of understanding of the value of AMH measurement across a wide range of physiologic and clinical situations (11), there are issues of lack of standardization among those produced by different manufacturers and concerns about data reliability (12).

Fully automated platform-based AMH assays are being 
developed, with the characteristics of one developed by Roche Diagnostics recently described (13). Our study was designed to 
investigate the value of AMH measurement using this novel 
Elecsys AMH assay in the assessment of ovarian reserve as 
expressed by AFC.

MATERIALS AND METHODS

Study Design

Our prospective observational study, undertaken in seven 
infertility centers, enrolled a subject cohort of healthy female 
volunteers and patients and measured the AMH concentra-
tion in relation to the AFC determined via transvaginal 
sonography on days 2–4 of a menstrual cycle. The women 
were recruited from the general population and from infert-
ility clinics. In both cases, the inclusion criteria included 
ages 18–44 years inclusive, regular self-reported menstrual cycles from 24 to 35 days in length, and informed consent 
given in writing. The exclusion criteria included pregnancy, 
major uterine or ovarian abnormalities detected by transvagi-
nal sonography, polycystic ovary syndrome (PCOS), endo-
crine or metabolic abnormalities (i.e., diabetes type I or II, 
or pituitary, adrenal, pancreas, liver, or kidney disturbances), 
ovidian surgery in the past 6 months, hormone therapy in the 
preceding 3 months (hormonal contraceptives, gonadotropin-
releasing hormone [GnRH] agonists, FSH), or current or past 
smoking. The sample size calculation was based on accurate 
estimation of the Spearman correlation coefficient as well as 
on an accurate estimation of AMH cutoffs for the AFC 
<7 and AFC ≥15 group.

We calculated that at least 400 patients should be enrolled 
to obtain accurate estimates for the Spearman correlation co-
efficient as well as AMH cutoffs. With this number of patients, 
a presumed correlation coefficient of 0.6 can be estimated 
with the width of its confidence interval smaller than 0.2 
(14). In addition, the assumed prevalences for the AFC ≤7, 
AFC 8–15, AFC >15 groups were 10%, 40%, and 50%, result-
ing in the determination of the 10% and 50% quantile for 
AMH cutoffs. With the determined sample size, the width of 
the confidence intervals (CI) of these cutoffs will be smaller 
than 0.56 ng/mL. The study received ethics committee 
approval in all centers.

The study intervention consisted of a single visit, at which a blood sample was taken for later hormone measurements and a transvaginal ultrasound examination was performed to determine AFC. Antral follicles were classified as those measuring 2–10 mm in diameter, and AFC was determined as the overall number of antral follicles counted in both ovaries. In all centers, AFC was determined by as few individ-
uals as possible, and a consistent methodology was followed (7). In all centers, two-dimensional (2D) transvaginal ultra-
sound equipment with a probe ≥6 MHz, minimal resolution 
2 mm was used (Supplemental Table 1, available online). 
The sonographic evaluation of all study participants was per-
fomed between January 2013 and January 2014.

Blood was allowed to clot, and the serum separated by centrifugation and stored at −80°C (or at −20°C for a 
maximum of 6 months) until analysis after shipping to a cen-
tral laboratory (Free University Brussels, Belgium). All serum 
markers (AMH, FSH, E2) were determined in single measure-
ment on the e601-module of the fully automated cobas®6000 
system. The measurements were split over 10 independent 
runs which were performed on different days. For each of the 
markers, a two-level control sample set was determined
in each run (AMH 0.93 and 4.8 ng/mL, FSH 17.5 and 44.5 IU/L, and E2 368 and 1,999 pmol/L).

The new Elecsys AMH assay is a sandwich assay based on electrochemiluminescence technology. The total duration of the assay is 18 minutes, and the sample volume is 50 μL. The assay is calibrated against the Beckman Coulter AMH Gen II ELISA (unmodified version without predilution) assay with a measuring range of 0.01–23 ng/mL. The limit of quantitation (functional sensitivity) is 0.03 ng/mL. The coefficients of variation as determined for the control samples during the study measurements were ≤3.3%, ≤2.2%, and ≤3.7% for the intermediate precision for AMH, FSH, and E2, respectively. The statistical analysis was performed using R 3.0.1 software.

The influence of categorical covariates on the AMH or AFC value was based on analysis of variance (ANOVA) F tests, and box plots show the distribution of these values in the different categories. Individual group comparisons were based on the t test. The correlation between AMH or AFC and other continuous variables as well as between AMH and AFC was assessed via Spearman’s ρ correlation coefficient and tested whether this correlation coefficient is different from zero. The relationship between AFC and age, as well AMH and age, is exemplified by the Passing-Bablok regression, a robust regression analysis. The agreement between AFC and AMH, and AFC and age, a logistic regression model was used for the classification of low and high AFC by age, a logistic regression model was used for the classification in low or high ovarian reserve.

RESULTS

For our study, 487 eligible women met the inclusion and exclusion criteria. From these women, 36 had to be excluded due to unforeseen reasons. As for the participating women, 92% were Caucasian, their mean BMI was 24.3 kg/m² (range: 16.7–52.7), and 68.5% were not infertile. Their ages were approximately evenly distributed between groups aged 18–29, 30–34, and 35–39 years, with fewer aged 40–45 years.

To minimize interobserver variation, 91% of AFC determinations were performed by 13 clinicians/sonographers. The distribution of AFC and AMH values is shown in Supplemental Table 2 (available online) overall and for each site, together with FSH and E2 measurements. There were statistically significant differences in mean AFC values among the centers (P<.001) whereas the AMH mean values did not vary by center (P=.30). As expected, both AFC and AMH varied by age (overall Spearman’s ρ = 0.50 for AFC and 0.47 for AMH, P<.001; Fig. 1). There was also statistically significant between-center variation for age-adjusted AFC (P<.001) but not for age-adjusted AMH (P=.19).

The primary aim of this study was to investigate AMH (using the Elecsys AMH assay) as a biomarker of the ovarian reserve determined by AFC. The study confirmed that there is a strong positive correlation between AMH and AFC (Spearman’s ρ = 0.68, P<.001). This relationship was present in each center (all P<.001; Supplemental Table 3, available online), although Spearman’s ρ varied from 0.49 to 0.87.

As AFC is used clinically to define poor responders and those at risk of OHSS, further analysis explored AMH by AFC grouping of ≤7, 8–15, and >15 (15). This showed highly significant differences in mean AMH between the AFC groups (P<.001; Supplemental Fig. 1, available online). An agreement table was constructed using the same AFC cutoffs of 7 and 15, with an AFC of 7 corresponding to the 15th percentile and AFC of 15 to the 52nd percentile. The equivalent percentiles for AMH were 0.68 ng/mL and 2.72 ng/mL. This analysis (Table 2) shows classification agreement in 63.2%, 56.9%, and 74.5% of women for the low, medium, and high groups, respectively.

We also analyzed AMH in comparison with FSH and E2 concentrations, with all samples being taken on days 2–4 of the menstrual cycle. The FSH concentrations ranged from 0.51 to 45.9 IU/L and E2 from 18.4 to 684 pmol/L. There was a negative relationship between AMH and FSH (Spearman’s ρ = −0.42, P<.001) but no statistically significant relationship with E2 (Spearman’s ρ = −0.04, P=.45).

The ROC curve analysis was performed for the classification of low AFC (≤7), which included 66 women versus 385 with AFC >7 and high AFC >15 (216 women vs. 235 women with lower AFC) (Fig. 2). For both low and high AFC classifications, AMH showed good discrimination with AUC of 91.1% (95% CI, 87.1%–95.2%) and 82.7% (95% CI, 79.0%–86.5%), respectively (both P<.001). This was statistically significantly better than for age, and markedly so than for the hormones FSH and E2. The sensitivity and specificity calculated for the classification of low and high AFC by means of the AMH quantiles as derived for the agreement tables were 65.2% and 93.5%, respectively, for the low AMH quantile (0.68 ng/mL), and 74.5% and 76.2% for the high AMH quantile (2.72 ng/mL). Youden’s indices (sensitivity + specificity −1) were 0.59 and 0.51, respectively. The combination of AMH with age did not statistically significantly improve the clinical performance for the low and high AFC classifications: AUC 91.7% (95% CI, 87.8%–95.6%) and 83.6% (95% CI, 80.0%–87.2%), respectively.

DISCUSSION

This study shows the value of a novel fully automated Elecsys AMH assay in the analysis of ovarian reserve as defined by AFC. Both AFC and AMH have become widely used biomarkers for what is widely termed the ovarian reserve in the context of prediction of assisted reproductive treatment outcome, which is of key importance to patients and their
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Overall (N = 451)</th>
<th>1 (n = 37)</th>
<th>2 (n = 55)</th>
<th>3 (n = 59)</th>
<th>4 (n = 111)</th>
<th>5 (n = 57)</th>
<th>6 (n = 115)</th>
<th>7 (n = 17)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (y)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (95% CI)</td>
<td>32.84 (32.31–33.38)</td>
<td>34.16 (32.54–35.79)</td>
<td>32.76 (31.43–34.10)</td>
<td>33.12 (32.15–34.08)</td>
<td>30.17 (28.89–31.45)</td>
<td>31.16 (29.79–32.53)</td>
<td>35.23 (34.25–36.21)</td>
<td>36.24 (33.67–38.80)</td>
</tr>
<tr>
<td>SD</td>
<td>5.78</td>
<td>4.87</td>
<td>4.94</td>
<td>3.7</td>
<td>6.82</td>
<td>5.16</td>
<td>5.31</td>
<td>4.99</td>
</tr>
<tr>
<td>Min–Max</td>
<td>18.00–44.00</td>
<td>25.00–44.00</td>
<td>21.00–42.00</td>
<td>26.00–40.00</td>
<td>18.00–44.00</td>
<td>22.00–42.00</td>
<td>21.00–44.00</td>
<td>23.00–43.00</td>
</tr>
<tr>
<td><strong>Age group (y)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–29</td>
<td>123 (27.27)</td>
<td>7 (18.92)</td>
<td>13 (23.64)</td>
<td>9 (15.25)</td>
<td>51 (45.95)</td>
<td>25 (43.86)</td>
<td>16 (13.91)</td>
<td>2 (11.76)</td>
</tr>
<tr>
<td>30–34</td>
<td>148 (32.82)</td>
<td>13 (35.14)</td>
<td>22 (40.00)</td>
<td>32 (54.24)</td>
<td>28 (25.23)</td>
<td>16 (28.07)</td>
<td>34 (29.57)</td>
<td>3 (17.65)</td>
</tr>
<tr>
<td>35–39</td>
<td>120 (26.61)</td>
<td>11 (29.73)</td>
<td>13 (23.64)</td>
<td>15 (25.42)</td>
<td>22 (19.82)</td>
<td>13 (22.81)</td>
<td>38 (33.04)</td>
<td>8 (47.06)</td>
</tr>
<tr>
<td>40–45</td>
<td>60 (13.30)</td>
<td>6 (16.22)</td>
<td>7 (12.73)</td>
<td>3 (5.08)</td>
<td>10 (9.01)</td>
<td>3 (5.26)</td>
<td>27 (23.48)</td>
<td>4 (23.53)</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>414 (91.80)</td>
<td>32 (86.49)</td>
<td>46 (83.64)</td>
<td>55 (93.22)</td>
<td>109 (98.20)</td>
<td>47 (82.46)</td>
<td>110 (95.65)</td>
<td>15 (88.24)</td>
</tr>
<tr>
<td>Black</td>
<td>15 (3.33)</td>
<td>0 (0.00)</td>
<td>5 (9.09)</td>
<td>1 (1.69)</td>
<td>0 (0.00)</td>
<td>9 (15.79)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>Asian</td>
<td>16 (3.55)</td>
<td>5 (13.51)</td>
<td>0 (0.00)</td>
<td>3 (5.08)</td>
<td>0 (0.00)</td>
<td>1 (1.75)</td>
<td>5 (4.35)</td>
<td>2 (11.76)</td>
</tr>
<tr>
<td>Other</td>
<td>6 (1.33)</td>
<td>0 (0.00)</td>
<td>4 (7.27)</td>
<td>0 (0.00)</td>
<td>2 (1.80)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td><strong>Smoking habit</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>99 (21.95)</td>
<td>3 (8.11)</td>
<td>8 (14.55)</td>
<td>8 (13.56)</td>
<td>52 (46.85)</td>
<td>16 (28.07)</td>
<td>11 (9.57)</td>
<td>1 (5.88)</td>
</tr>
<tr>
<td>Past</td>
<td>80 (17.74)</td>
<td>6 (16.22)</td>
<td>13 (23.64)</td>
<td>2 (3.39)</td>
<td>8 (7.21)</td>
<td>11 (19.30)</td>
<td>34 (29.57)</td>
<td>6 (35.29)</td>
</tr>
<tr>
<td>Never</td>
<td>272 (60.31)</td>
<td>28 (61.82)</td>
<td>34 (61.82)</td>
<td>49 (83.05)</td>
<td>51 (45.95)</td>
<td>30 (52.63)</td>
<td>70 (60.87)</td>
<td>10 (58.82)</td>
</tr>
<tr>
<td><strong>Female infertility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>309 (68.51)</td>
<td>20 (54.05)</td>
<td>28 (50.91)</td>
<td>27 (45.76)</td>
<td>90 (81.08)</td>
<td>30 (52.63)</td>
<td>109 (94.78)</td>
<td>5 (29.41)</td>
</tr>
<tr>
<td>Yes</td>
<td>142 (31.49)</td>
<td>17 (45.95)</td>
<td>27 (49.09)</td>
<td>32 (54.24)</td>
<td>21 (18.92)</td>
<td>27 (47.37)</td>
<td>6 (5.22)</td>
<td>12 (70.59)</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>4.77</td>
<td>5.91</td>
<td>5.27</td>
<td>3.49</td>
<td>3.28</td>
<td>5.82</td>
<td>5.02</td>
<td>3.82</td>
</tr>
<tr>
<td>Min–Max</td>
<td>16.70–52.70</td>
<td>16.70–40.40</td>
<td>17.70–52.70</td>
<td>18.20–34.10</td>
<td>17.80–33.50</td>
<td>17.70–41.00</td>
<td>17.80–47.30</td>
<td>18.20–33.50</td>
</tr>
</tbody>
</table>

Note: Percentage of total in parentheses. CI = confidence interval; SD = standard deviation.

clinical teams. Although live birth is the ultimate positive outcome, prediction of ovarian response is also of great value for identifying women who are likely to respond poorly, thus allowing appropriate counseling in advance, and for identifying women likely to show an excessive response, thus predicting a potential for OHSS. Thus, in both groups management strategies can be tailored to optimize outcome and minimize risk (4).

Both AFC and AMH have specific characteristics and thus advantages and disadvantages as biomarkers. Because AFC is widely available, it has the advantage of immediacy, but standardization is a difficulty. Ultrasound equipment varies among centers, and there have been progressive increases in resolution and thus image quality over recent years. This has led to substantial changes in what is regarded as a “normal” AFC, as well as in what might be used as a diagnostic criterion in, for example, polycystic ovarian syndrome (5). There is also well-recognized, substantial interobserver variation (17), compounding the variations among women and across the menstrual cycle (18); as a result, assessment in the early follicular phase is required, as performed in this study. Although AMH may show less intraindividual and interindividual variation than AFC (19–21), the lack of standardization of the calibrators among the different manufacturers is also a significant issue, as are other methodologic problems that have affected reproducibility (12). Previous assays have also all been manual plate–based formats, with inherent susceptibility to variation within and between laboratories and in lot-to-lot variation. The technical characteristics of the fully automated AMH assay used in this study have been recently described, indicating that it may address many of the issues with the measurement of this hormone (13, 22).

The present hormone analyses were all performed in a single academic laboratory, with a very similar technical performance for AMH to that previously described. There was no significant variation in AMH distribution between centers, although there was significant variation in AFC. Likewise, although both AMH and AFC showed the expected inverse relationship with age, there was significant variation in that relationship among centers for AFC but not for AMH. These analyses may indicate that AMH measured by this new assay is a more reliable indicator of the ovarian reserve than AFC. The full automation of the AMH assay removes operator-related variation; the centralization of hormone analysis

### TABLE 2

<table>
<thead>
<tr>
<th>AMH group</th>
<th>AFC 0–7</th>
<th>AFC 8–15</th>
<th>AFC &gt;15</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMH ≤ 0.681</td>
<td>43 (63.2%)</td>
<td>22 (32.4%)</td>
<td>3 (4.4%)</td>
<td>68</td>
</tr>
<tr>
<td>0.681 &lt; AMH ≤ 2.27</td>
<td>20 (12.0%)</td>
<td>95 (56.9%)</td>
<td>52 (31.1%)</td>
<td>167</td>
</tr>
<tr>
<td>AMH &gt; 2.27</td>
<td>3 (1.4%)</td>
<td>52 (24.1%)</td>
<td>161 (74.5%)</td>
<td>216</td>
</tr>
<tr>
<td>N</td>
<td>66</td>
<td>169</td>
<td>216</td>
<td>451</td>
</tr>
</tbody>
</table>

Note: Percentages refer to AMH group numbers (AMH values in ng/mL).
removes a further source of variability, and thus this may not exactly reflect the situation where individual sites perform hormone analysis. The possibility of centralization is an inherent advantage of hormone analysis over ultrasound-based analysis: recording images and their offline and centralized analysis are possible in principle (17), but they are very time consuming, so one of the key advantages of AFC, its immediacy, is lost.

Although AMH and AFC are sometimes regarded as interchangeable, there are emerging data suggesting that AMH is a more reproducible measure of the ovarian response to stimulation. In an analysis of potential markers to predict ovarian response in a randomized controlled trial, AFC was not found to be predictive (23). In contrast, AMH was predictive both of number of retrieved oocytes and of poor versus excessive response. Similarly, in another randomized controlled trial, AMH but not AFC predicted oocyte yield after ovarian stimulation (24). These findings have led to discussion of the limitations of AFC for the prediction of ovarian response, particularly in multicenter trials (25). Although the present data do not include assessment of ovarian response, the finding of significant intercenter variability in the relationship between AFC and age but not between AMH and age is consistent with this growing recognition of difficulties in standardizing AFC assessment among centers.

Further characterization of the relationship between AFC and AMH showed good performance in an agreement table and in ROC analyses. These results are similar to those previously reported using other AMH assays (4, 6) and support the value of AMH across the dynamic range found in women of ages across the reproductive life span.

In conclusion, these data demonstrate the clinical performance of the new fully automated Elecsys AMH assay in the analysis of the ovarian reserve in women of reproductive age. The expected relationships with age and AFC were observed, with evidence of much lower variability in the determination of AMH compared with AFC. This supports the use of the automated AMH assay in a range of contexts in reproductive medicine such as in physiologic, therapeutic, and potentially pathologic investigations.

REFERENCES
SUPPLEMENTAL FIGURE 1

Distribution of AMH (ng/mL) in different AFC groups for all sites. The P values in the graph refer to t tests of AMH means between the AFC 0–7 and 8–15 group and between the 8–15 and >15 group.

**SUPPLEMENTAL TABLE 1**

Characteristics of ultrasonography equipment used for transvaginal sonography at study sites.

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Study site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturer</td>
<td>1</td>
</tr>
<tr>
<td>Device type</td>
<td>GE LOGIQ P3</td>
</tr>
<tr>
<td>Probe frequency</td>
<td>8 MHz</td>
</tr>
<tr>
<td>Probe resolution</td>
<td>2 mm</td>
</tr>
<tr>
<td>Yearly maintenance</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### SUPPLEMENTAL TABLE 2

Antral follicle count (AFC), antimüllerian hormone (AMH), follicle-stimulating hormone (FSH), and estradiol (E₂) data in study participants, overall and by study site.

<table>
<thead>
<tr>
<th>AFC group</th>
<th>Overall (N = 451)</th>
<th>1 (n = 37)</th>
<th>2 (n = 55)</th>
<th>3 (n = 59)</th>
<th>4 (n = 111)</th>
<th>5 (n = 57)</th>
<th>6 (n = 115)</th>
<th>7 (n = 17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–7</td>
<td>66 (14.63)</td>
<td>5 (13.51)</td>
<td>13 (23.64)</td>
<td>10 (16.95)</td>
<td>20 (18.02)</td>
<td>2 (3.51)</td>
<td>12 (10.43)</td>
<td>4 (23.53)</td>
</tr>
<tr>
<td>8–15</td>
<td>169 (37.47)</td>
<td>7 (18.92)</td>
<td>18 (32.73)</td>
<td>32 (54.24)</td>
<td>34 (30.63)</td>
<td>18 (31.58)</td>
<td>49 (42.61)</td>
<td>11 (64.71)</td>
</tr>
<tr>
<td>&gt;15</td>
<td>216 (47.89)</td>
<td>25 (57.57)</td>
<td>24 (43.64)</td>
<td>17 (28.81)</td>
<td>57 (51.35)</td>
<td>37 (64.91)</td>
<td>54 (46.96)</td>
<td>2 (11.76)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AFC [n]</th>
<th>Mean (95% CI)</th>
<th>SD</th>
<th>Min–Max</th>
<th>Mean (95% CI)</th>
<th>SD</th>
<th>Min–Max</th>
<th>Mean (95% CI)</th>
<th>SD</th>
<th>Min–Max</th>
<th>Mean (95% CI)</th>
<th>SD</th>
<th>Min–Max</th>
<th>Mean (95% CI)</th>
<th>SD</th>
<th>Min–Max</th>
<th>Mean (95% CI)</th>
<th>SD</th>
<th>Min–Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>16.17 (15.37–16.96)</td>
<td>8.58</td>
<td>0–56.00</td>
<td>19.03 (16.06–22.00)</td>
<td>8.91</td>
<td>4.00–38.00</td>
<td>15.31 (12.85–17.77)</td>
<td>9.1</td>
<td>0.00–35.00</td>
<td>13.39 (11.42–15.36)</td>
<td>7.55</td>
<td>2.00–38.00</td>
<td>15.63 (14.20–17.06)</td>
<td>7.62</td>
<td>3.00–56.00</td>
<td>21.30 (18.39–24.21)</td>
<td>10.96</td>
<td>2.00–38.00</td>
</tr>
<tr>
<td>1 (n = 37)</td>
<td>6.66 (5.50–7.82)</td>
<td>1.28</td>
<td>2.00–17.02</td>
<td>7.26 (6.28–8.24)</td>
<td>0.47</td>
<td>3.75–13.75</td>
<td>6.98 (6.49–7.46)</td>
<td>0.56</td>
<td>3.81–23.81</td>
<td>7.43 (6.84–8.03)</td>
<td>0.74</td>
<td>1.88–18.88</td>
<td>7.56 (6.39–8.73)</td>
<td>0.88</td>
<td>3.17–23.17</td>
<td>7.18 (6.65–7.71)</td>
<td>1.00</td>
<td>4.42–24.42</td>
</tr>
<tr>
<td>2 (n = 55)</td>
<td>7.66 (6.50–8.82)</td>
<td>1.38</td>
<td>2.00–17.02</td>
<td>7.51 (6.42–8.60)</td>
<td>0.55</td>
<td>3.75–13.75</td>
<td>7.08 (6.59–7.57)</td>
<td>0.50</td>
<td>3.81–23.81</td>
<td>7.36 (6.77–8.06)</td>
<td>0.73</td>
<td>1.88–18.88</td>
<td>7.56 (6.39–8.73)</td>
<td>0.88</td>
<td>3.17–23.17</td>
<td>7.18 (6.65–7.71)</td>
<td>1.00</td>
<td>4.42–24.42</td>
</tr>
<tr>
<td>3 (n = 59)</td>
<td>8.96 (7.80–10.12)</td>
<td>1.62</td>
<td>2.00–17.02</td>
<td>8.01 (6.97–9.06)</td>
<td>0.68</td>
<td>3.75–13.75</td>
<td>7.96 (7.47–8.46)</td>
<td>0.51</td>
<td>3.81–23.81</td>
<td>7.43 (6.84–8.03)</td>
<td>0.74</td>
<td>1.88–18.88</td>
<td>7.56 (6.39–8.73)</td>
<td>0.88</td>
<td>3.17–23.17</td>
<td>7.18 (6.65–7.71)</td>
<td>1.00</td>
<td>4.42–24.42</td>
</tr>
<tr>
<td>6 (n = 115)</td>
<td>16.20 (15.00–17.40)</td>
<td>2.37</td>
<td>2.00–17.02</td>
<td>15.95 (14.85–17.05)</td>
<td>1.37</td>
<td>3.75–13.75</td>
<td>16.20 (15.00–17.40)</td>
<td>2.37</td>
<td>3.75–13.75</td>
<td>16.20 (15.00–17.40)</td>
<td>2.37</td>
<td>3.75–13.75</td>
<td>16.20 (15.00–17.40)</td>
<td>2.37</td>
<td>3.75–13.75</td>
<td>16.20 (15.00–17.40)</td>
<td>2.37</td>
<td>3.75–13.75</td>
</tr>
</tbody>
</table>

Note: Percentage of total in parentheses. CI = confidence interval; SD = standard deviation.

## Supplemental Table 3

Correlation between AFC and AMH over all sites and for individual sites.

<table>
<thead>
<tr>
<th>Correlation</th>
<th>All</th>
<th>Site 1</th>
<th>Site 2</th>
<th>Site 3</th>
<th>Site 4</th>
<th>Site 5</th>
<th>Site 6</th>
<th>Site 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman</td>
<td>0.68</td>
<td>0.71</td>
<td>0.86</td>
<td>0.49</td>
<td>0.87</td>
<td>0.74</td>
<td>0.62</td>
<td>0.75</td>
</tr>
<tr>
<td>P value (Spearman)</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Intercept</td>
<td>−0.25</td>
<td>−0.51</td>
<td>−0.27</td>
<td>0.22</td>
<td>−0.64</td>
<td>−0.29</td>
<td>−0.76</td>
<td>−1.32</td>
</tr>
<tr>
<td>Slope</td>
<td>0.16</td>
<td>0.14</td>
<td>0.16</td>
<td>0.20</td>
<td>0.20</td>
<td>0.11</td>
<td>0.21</td>
<td>0.36</td>
</tr>
<tr>
<td>N</td>
<td>451</td>
<td>37</td>
<td>55</td>
<td>59</td>
<td>111</td>
<td>57</td>
<td>115</td>
<td>17</td>
</tr>
</tbody>
</table>