RBMO

ELSEVIER

ARTICLE

A multicentre evaluation of the Elecsys® anti-Müllerian hormone immunoassay for prediction of antral follicle count

BIOGRAPHY

Michael H. Jacobs graduated from Case Western Reserve University before completing residency in Obstetrics and Gynecology at Columbia Presbyterian Medical Center, and fellowship in Reproductive Endocrinology/Infertility at the Hospital of the University of Pennsylvania. In 1991, Dr Jacobs founded the Fertility and IVF Center of Miami, and remains its Medical Director.

Michael H. Jacobs[1,*], Laura M. Reuter[2], Valerie L. Baker[3], LaTasha B. Craig[4], Denny Sakkas[5], Eric Surrey[6], Kevin J. Doody[7], Emily S. Jungheim[8], Aykut B. Bayrak[9], Martin Hund[10], Wilma D.J. Verhagen-Kamerbeek[10], Deborah Pardue[11], Katharina Buck[12], Barbara Timm[13]

KEY MESSAGE

Based on a 1.77 ng/ml cut-off, the Elecsys® anti-Müllerian hormone immunoassay identifies women with an antral follicle count >15 with high specificity and sensitivity. This assay provides a reliable means to determine ovarian reserve, and could facilitate informed clinical decision-making for women receiving counselling on assistive reproductive therapy.

ABSTRACT

Research question: What concentration of anti-Müllerian hormone (AMH) corresponds to an antral follicle count (AFC) >15 for determination of ovarian reserve?
Design: A prospective study conducted at 13 US fertility clinics in women aged 21–44 years who presented for AFC evaluation by transvaginal ultrasound. Serum samples were collected at the time of AFC evaluation (menstrual cycle day 2–4). AMH concentrations were measured by the Elecsys® AMH immunoassay; oestradiol and follicle-stimulating hormone (FSH) concentrations were also measured. The serum AMH cut-off able to detect AFC >15 with high sensitivity was determined (derivation cohort). Clinical performance of the AMH assay at the derived cut-off was evaluated (validation cohort). Receiver operating characteristic (ROC) analyses were also performed.

Results: In the derivation cohort (n = 306), an optimal serum AMH cut-off value of 1.77 ng/ml was determined to correspond to AFC >15 with 89.63% sensitivity and 69.01% specificity, using the Elecsys AMH assay. In the validation cohort (n = 856), this 1.77 ng/ml cut-off could identify women with an AFC >15 with a sensitivity of 88.34% and a specificity of 68.29%; corresponding positive predictive and negative predictive values were 75.19% and 84.34%, respectively. ROC analyses demonstrated that AMH performed better than oestradiol or FSH in predicting AFC, with area under the curves of 85.7%, 57.1% and 69.7%, respectively, in the validation cohort.

Conclusion: The Elecsys AMH immunoassay provides a robust and fully automated method to measure serum AMH levels. Women with AMH values below the cut-off of 1.77 ng/ml are unlikely to have AFC >15.

[1] Fertility and IVF Center of Miami, 8950 North Kendall Drive, Suite 103, Miami FL 33176, USA
[2] Midwest Fertility Specialists, 12188-A North Meridian Street, Suite 250, Carmel IN 46032, USA
[3] Stanford Fertility Specialists, 1195 West Fremont Avenue, Sunnyvale CA 94087, USA
[4] University of Oklahoma Health Sciences Center, 1100 North Lindsay Avenue, Oklahoma City OK 73104, USA
[5] Boston IVF, 130 Second Avenue, Waltham MA 02451, USA
[6] Colorado Center for Reproductive Medicine, 10290 RidgeGate Circle, Lone Tree CO 80124, USA
[7] Center for Assisted Reproduction, 1701 Park Pl Avenue, Bedford TX 76022, USA
[8] Washington University Fertility and Reproductive Medicine Center, 4444 Forest Park Avenue No. 3100, St. Louis MO 63110, USA
[9] LA IVF, 2080 Century Park, Los Angeles CA 90067, USA
[10] Roche Diagnostics International Ltd, Forrenstrasse 2, Rotkreuz 6343, Switzerland
[11] Roche Diagnostics Operations Inc., 9115 Hague Road, Indianapolis IN 46250-0457, USA
[12] Roche Diagnostics GmbH, Nonnenwald 2, Penzberg 82377, Germany
[13] Heartland Center for Reproductive Medicine, 7308 South 142nd Street, Omaha NE 68138, USA

*Corresponding author. E-mail address: fivf@aol.com (M. H. Jacobs).
https://doi.org/10.1016/j.rbmo.2018.12.041
1472-6483/
Declaration: MH Jacobs, L Reuter, VL Baker, LB Craig, D Sakkas, K Doody, AB Bayrak and B Timm have no conflicts of interest to disclose. M Hund, WDJ Verhagen-Kamerbeek, D Pardue and K Buck are employees of Roche Diagnostics. ES Jungheim has participated in speaker's panels for Roche Diagnostics. E Surrey has participated in medical advisory boards and speaker's panels for AbbVie Laboratories, and speaker's panels for Ferring.
An abstract was submitted and a poster was presented at the 65th Annual Meeting of the Society for Reproductive Investigation (March 2018).

KEYWORDS

Anti-Müllerian hormone
Antral follicle count
Ommunoassay
Ovarian reserve

INTRODUCTION

Determination of ovarian reserve in women presenting to fertility clinics for counselling on assisted reproductive therapy is part of clinical decision-making in ovarian stimulation. Diminished ovarian reserve is associated with poor response to ovarian stimulation and may reflect poorer IVF cycle outcomes, although this is not consistently shown (*ASRM Committee Opinion, 2012; La Marca et al., 2010*). The antral follicle count (AFC) on day 2–4 of the menstrual cycle, assessed by transvaginal ultrasound (TVUS), is commonly used to determine ovarian reserve. Women with an AFC >15 are identified as having high ovarian reserve (*Anderson et al., 2015*), but are also at increased risk of hyper-response to ovarian stimulation (*Broer et al., 2011*). Biomarkers of ovarian reserve, such as anti-Müllerian hormone (AMH), follicle-stimulating hormone (FSH) and oestradiol are often evaluated (*Broer et al., 2010; Verhagen et al., 2008*); however, there is no consensus about the most accurate method.

AMH levels show good correlation with AFC, and are relatively stable throughout the menstrual cycle (*Hehenkamp et al., 2006; La Marca et al., 2007; van Disseldorp et al., 2010; van Rooij et al., 2002*). AMH measurement may therefore offer several advantages over AFC (*Weenen et al., 2004*) and facilitate more reliable assessment of ovarian reserve (*La Marca et al., 2014*). The Elecsys® AMH assay is an electrochemiluminescence immunoassay for quantitative determination of serum AMH, and the first automated AMH assay to receive FDA approval. The assay is standardized against the Beckman Coulter AMH Gen II ELISA and demonstrates high sensitivity and specificity for assessment of ovarian reserve, as well as excellent precision (*Anckaert et al., 2016; Anderson et al., 2015*).

The aim of this study was to determine and validate a cut-off value for the Elecsys AMH assay corresponding to an AFC >15 in women presenting to fertility clinics.

MATERIALS AND METHODS

Study design and participants

This was a prospective, non-interventional study performed according to the study protocol at 13 fertility clinics in the USA between February 2014 and July 2015. A list of participating centres is provided in the Supplementary Material. The study involved two cohorts: the derivation cohort was recruited from six sites, and the validation cohort was recruited from 13 sites (six sites used for the original derivation cohort as well as an additional seven sites).

Women aged 21–44 years who presented at fertility clinics for evaluation of AFC by TVUS (menstrual cycle day 2–4) were enrolled. Exclusion criteria were major ovarian abnormalities (including the presence of only one ovary, or cysts and solid masses >2 cm detected by TVUS); a diagnosis of polycystic ovary syndrome using the Rotterdam criteria (*Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004*); a documented positive pregnancy test at time of presentation; AMH determination in the preceding 3 months known by the study sonographer; body mass index (BMI) ≥40 kg/m^2; endocrine or metabolic abnormalities (diabetes, or pituitary, adrenal, pancreas, liver or kidney disease); ovarian surgery in the past 6 months or hormonal contraceptives in the preceding 3 months; any hormonal medication in the past 21 days (thyroid hormones permitted); and currently undergoing treatment for malignancy or positive serum human chorionic gonadotrophin (HCG) levels.

The study was conducted according to the ICH Guideline for Good Clinical Practice, the Declaration of Helsinki and the Convention of the Council of Europe. All women provided written informed consent. The study protocol was approved prior to study initiation by relevant institutional review boards (see Supplementary Material). Study sites entered patient clinical data into Medrio, a validated database for the electronic capture of patient data; accuracy of the database was validated by Roche Diagnostics.

Study procedures

AFC was determined by two- or three-dimensional TVUS (AFC was defined as the total number of antral follicles with a size of 2–10 mm in both ovaries). Due to the potential for inter-sonographer variability, sonographers at all sites were provided with training and were asked to follow published practical recommendations for accurate TVUS (*Broekmans et al., 2010*); each sonographer used standard equipment available at their site (2D or 3D TVUS equipment [Supplementary **TABLE 1**]).

Patient blood samples were collected during the same visit that TVUS was performed. Samples were processed within 4 h, according to the assay manufacturer's instructions. Samples were centrifuged at 2000*g* for 10 ± 5 min and serum was pipetted into 0.5 ml aliquots, which were frozen at –15°C or colder until shipment on dry ice to Roche Diagnostics (Indianapolis, IN, USA). Samples were subsequently stored at below –70°C until shipment on dry ice to the three US testing laboratories (Mayo Validation Support Services [MVSS], Rochester, Minnesota; Core Laboratory for Clinical Studies at Washington University Medical Centre, St. Louis, Missouri; Nationwide Laboratory Services, Fort Lauderdale, Florida, which moved to Plantation, Florida) where samples were stored at below –15°C and thawed on the day of testing.

Clinical performance

Blood samples were randomly distributed to one of the three US laboratories (listed above) for measurement of AMH using the Elecsys AMH immunoassay on a cobas e 411 analyzer (Roche Diagnostics). Oestradiol, FSH and HCG were also measured using the Elecsys oestradiol II, Elecsys oestradiol III (change in reagent made due to conversion to a new, comparable assay; no significant difference seen in results), Elecsys FSH and Elecsys HCG+β (intact HCG + the β subunit) immunoassays (Roche Diagnostics GmbH, Mannheim, Germany). Assays were performed according to the manufacturer's instructions. HCG was measured to detect possible pregnancies. At each laboratory, the analytical performance of the AMH assay was assessed prior to each run using PreciControl AMH control samples (QC run). Study samples were only analysed if the QC run met predefined criteria, i.e. measured AMH levels were within the target range on the package insert.

Analytical performance

At each site, the analytical performance (repeatability and reproducibility) of the Elecsys AMH assay was evaluated according to CLSI-EP15-A2 guidelines (*Clinical and Laboratory Standards Institute, 2005*). For these experiments, human serum pools

(HSP) containing known quantities of AMH were obtained from BIOMEX GmbH. Fetal bovine serum containing a high level of AMH was added to the HSP samples to derive the required AMH concentrations (measured using the cobas e 411 analyzer). HSP were aliquoted and frozen at below –70°C until use. Preparation and validation of these samples was conducted by Roche Diagnostics Research and Development (Penzberg, Germany) prior to use for this trial.

Statistical analysis

Coefficients of variation and 95% confidence intervals (CI) were determined for reproducibility and repeatability using a variance component analysis approach. Required sample size estimations are reported in the Supplementary Material.

Demographic variables, baseline characteristics and biomarker levels were summarized as *n*, total range, mean, standard deviation, median, and 25th and 75th quantiles (Q1, Q3) for each study cohort. Differences between the derivation cohort and validation cohort were tested using the Mann–Whitney and Fisher's exact tests for continuous and categorical variables, respectively. The AMH cut-off in the derivation cohort was established as the AMH concentration corresponding to an AFC >15 with 90% sensitivity. In the validation cohort, the cut-off value established in the derivation cohort was validated by evaluation of clinical performance in terms of assay specificity and sensitivity. Negative (NPV) and positive predictive values (PPV) with non-parametric 95% CI were also calculated for the quantiles of the Elecsys AMH immunoassay. Using the same methods, additional exploratory analyses were performed to determine the AMH cut-off corresponding to an AFC >20.

In addition, receiver operating characteristic (ROC) curves were used to show the classification potential of the biomarkers to identify high ovarian reserve based on AFC >15; area under the ROC curves (AUC) were calculated for AMH, FSH, oestradiol and age. Coefficients of variation per site were calculated using a variance component analysis (multivariable). Logistic regression analysis (multivariable) was used to assess marker combinations (AMH with FSH, and additionally with age) compared with AMH or age alone as predictors of AFC >15 (*ASRM Committee Opinion, 2012*). For the AUC calculations, continuous AMH values were used, whereas for sensitivity and specificity, AMH was dichotomized in these analyses. ROC curves were calculated from the predicted values and AUC compared by using DeLong's test for AUC (*DeLong et al., 1988*). All analyses were performed using R version 3.2.2 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Patient disposition and baseline demographics

Patient disposition for the derivation (n = 306) and validation (n = 856) cohorts is presented in FIGURE 1. Baseline demographics and clinical characteristics were comparable for the derivation cohort and validation cohort, with no statistically significant differences in age, BMI or race (TABLE 1). In the derivation and validation cohorts, respectively, mean age was 34.6 and 34.2 years, mean BMI was 25.6 kg/m^2 and 25.3 kg/m^2, the majority of women were white/Caucasian (82.4% and 77.5%), and the majority had never smoked (73.5% and 80.8%).

Analytical performance

The Elecsys AMH assay demonstrated excellent repeatability and reproducibility. The coefficients of variation for repeatability were <2.0% at all sites and coefficients of variation for total reproducibility were <5.5% with upper 95% CI of <11.0% (Supplementary TABLE 2).

Clinical performance

AFC and concentrations of AMH, FSH and oestradiol are presented for the derivation cohort and validation cohort in TABLE 2. For AMH, no statistically significant difference was observed between the derivation cohort and validation cohort (TABLE 2), whereas for AFC, a statistically significant difference was observed (P = 0.024; TABLE 2; Supplementary FIGURE 1A). In the overall cohort, between-site variability was smaller for AMH than for AFC (coefficients of variation: 16.2% versus 33.0%, respectively; Supplementary FIGURE 2). There was a good correlation between AMH and AFC in both cohorts (Spearman's r coefficient = 0.8 and 0.7 in the derivation cohort and validation cohort, respectively; Supplementary FIGURE 1B). Spearman's r correlations of AMH and AFC per recruitment site ranged from 0.7 to 0.9.

Of the 306 women in the derivation cohort, 121 (39.54%) with an AFC >15 also had an AMH concentration >1.77 ng/ml, while 118 (38.56%) with an AFC of 0–15 also had an AMH ≤1.77 ng/ml (TABLE 3). An optimal serum AMH cut-off value of 1.77 ng/ml (95% CI 1.44–2.06 ng/ml) was determined to correspond to AFC >15 with 89.63% (95% CI 83.21–94.21) sensitivity and 69.01% (95% CI 61.49–75.84) specificity, using the Elecsys AMH assay. AMH concentrations in the derivation cohort are presented by AFC classification ≤15 or >15 in FIGURE 2.

Of the 856 women in the validation cohort, 394 (46.03%) with an AFC >15 also had an AMH concentration >1.77 ng/ml, while 280 (32.71%) with an AFC of 0–15 also had an AMH ≤1.77 ng/ml (TABLE 3). When evaluated in the validation cohort, the derived 1.77 ng/ml cut-off could identify women with an AFC >15 with a sensitivity of 88.34% (95% CI 84.99–91.17) and a specificity of 68.29% (95% CI 63.55–72.77), using the Elecsys AMH assay. PPV and NPV values at 1.77 ng/ml cut-off were 75.19% (95% CI 71.26–78.83) and 84.34% (95% CI 79.97–88.08), respectively. AMH concentrations in the validation cohort are presented by AFC class (≤15 or >15) in FIGURE 2.

ROC analyses showed that AMH performed markedly better for the determination of ovarian reserve than oestradiol or FSH, with a larger AUC for AMH (90.5%) compared with oestradiol (50.5%) and FSH (69.1%) in the derivation cohort (FIGURE 3). AMH and oestradiol were not correlated (Spearman's r correlation = –0.06) and there was a moderate negative correlation between AMH and FSH (Spearman's r correlation = –0.5; P < 0.01). In the validation cohort, the derived AMH cut-off of 1.77 ng/ml showed good discrimination of AFC >15. AUC were markedly higher for AMH (85.7%) than for oestradiol (57.1%) or FSH (69.7%) (FIGURE 3). AMH and oestradiol were not correlated (Spearman's r correlation = –0.2) and there was a moderate negative correlation between AMH and FSH (Spearman's r correlation = –0.5; P < 0.01). A multivariable model combining AMH

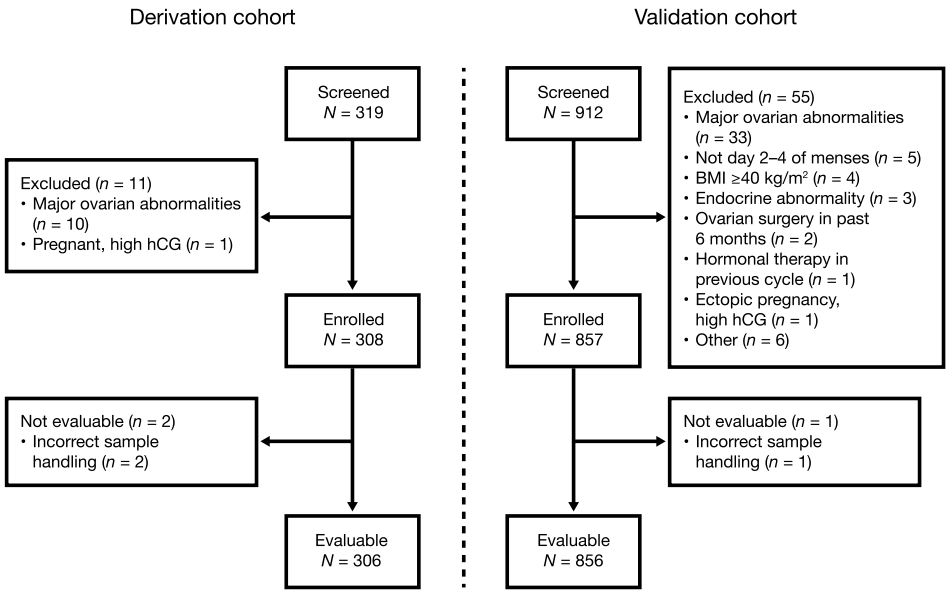

FIGURE 1 Patient flow diagram. BMI = body mass index; HCG = human chorionic gonadotrophin.

with FSH or age did not improve the predictive value over that of AMH alone whereas AMH alone or in combination with age was better than age alone in predicting AFC >15 (Supplementary **TABLE 3**).

Exploratory analyses

Of the 306 women in the derivation cohort, 87 had an AFC >20, of whom 78 (25.49% of the derivation population) also had an AMH concentration >2.64 ng/ml. Of the 219 women with an AFC of 0–20, 179 (58.50% of the derivation population) also had an AMH ≤2.64 ng/ml (Supplementary **TABLE 4**). A serum AMH cut-off value of 2.64 ng/ml (95% CI 1.77–2.81) was determined to

TABLE 1 SUMMARY OF DEMOGRAPHICS FOR THE DERIVATION AND VALIDATION COHORTS

Characteristic[a]	Parameter	Derivation cohort (*n* = 306; 6 sites)	Validation cohort (*n* = 856; 13 sites)
Age, years	Mean (±SD)	34.59 (±4.71)	34.20 (±4.89)
	Median (Q1–Q3)	35.00 (31.00–38.00)	34.00 (31.00–38.00)
	Min–max	23.00–44.00	21.00–44.00
Age group, years, *n* (%)	21–29	48 (15.69)	159 (18.57)
	30–34	104 (33.99)	277 (32.36)
	35–39	99 (32.35)	290 (33.88)
	40–44	55 (17.97)	130 (15.19)
Race, *n* (%)	White/Caucasian	252 (82.35)	663 (77.45)
	Asian	28 (9.15)	87 (10.16)
	Black/African American	18 (5.88)	56 (6.54)
	Indian/Alaskan native	0	1 (0.12)
	Other	6 (1.96)	21 (2.45)
	Multiple	2 (0.65)	28 (3.27)
BMI, kg/m²	Mean (±SD)	25.55 (±5.12)	25.30 (±4.95)
	Median (Q1–Q3)	24.21 (22.13–28.24)	23.99 (21.73–28.22)
	Min–max	17.54–39.99	14.76–39.86
Smoking, *n* (%)	Ex-smoker	57 (18.63)	122 (14.25)
	Never smoked	225 (73.53)	692 (80.84)
	Smoker	24 (7.84)	41 (4.79)
	No information	0	1 (0.12%)

[a] *P*-values for comparison between derivation and validation cohorts are based on the Mann–Whitney test for continuous values and on Fisher's exact test for categorical variables. Of all characteristics analysed, only smoking had a significant *P*-value (0.027). BMI = body mass index; SD = standard deviation.

TABLE 2 SUMMARY OF AFC AND AMH, FSH AND OESTRADIOL LEVELS IN THE DERIVATION AND VALIDATION COHORTS

Biomarker	Parameter	Derivation cohort (*n* = 306)	Validation cohort (*n* = 856)	*P*-value[a]
AFC, *n* (%)	0–15	171 (55.88)	410 (47.90)	0.02
	>15	135 (44.12)	446 (52.10)	
AFC (*n*)	Mean (±SD)	16.81 (±10.34)	18.87 (±12.43)	0.024
	Median (Q1–Q3)	14.00 (10.00–22.00)	16.00 (10.00–24.00)	
	Min–max	1.00–60.00	1.00–99.00	
AMH (ng/ml)	Mean (±SD)	2.75 (±2.69)	2.77 (±2.19)	NS
	Median (Q1–Q3)	2.04 (1.13–3.53)	2.23 (1.19–3.75)	
	Min–max	0.03–22.11	0.01–18.54	
FSH (IU/l)	Mean (±SD)	7.94 (±3.37)	7.99 (±3.28)	NS
	Median (Q1–Q3)	7.19 (6.01–8.78)	7.29 (6.10–8.80)	
	Min–max	0.16–31.97	1.86–31.62	
Oestradiol (pg/ml)	Mean (±SD)	43.51 (±28.14)	44.21 (±34.08)	NS
	Median (Q1–Q3)	37.26 (28.40–54.03)	39.25 (28.27–53.18)	
	Min–max	5.00–311.00	5.00–724.50	

[a] *P*-values for comparison between derivation and validation cohorts are based on the Mann–Whitney test for continuous values and on Fisher's exact test for categorical variables. AFC = antral follicle count; AMH = anti-Müllerian hormone; FSH = follicle-stimulating hormone; IU = international units; NS = not statistically significant; SD = standard deviation.

correspond to AFC >20 with 89.66% (95% CI 81.27–95.16) sensitivity and 81.74% (95% CI 75.97–86.62) specificity, using the Elecsys AMH assay.

Of the 856 women in the validation cohort, 221 (25.82% of the validation population) with an AFC >20 also had an AMH concentration >2.64 ng/ml, while 437 (51.05% of the validation population) with an AFC of 0–20 also had an AMH ≤2.64 ng/ml (Supplementary **FIGURE 3**; Supplementary **TABLE 4**). In the validation cohort, the derived 2.64 ng/ml cut-off could identify women with an AFC >20 with a sensitivity of 74.41% (95% CI 69.05–79.28) and a specificity of 78.18% (95% CI 74.52–81.53), using the Elecsys AMH assay. At a cut-off of 2.64 ng/ml, the PPV was 64.43% (95% CI 59.11–69.50) and the NPV was 85.19% (95% CI 81.81–88.15).

DISCUSSION

Determination of ovarian reserve by AFC is commonly used to help clinical decision-making for women receiving counselling on assisted reproductive technology (*ASRM Committee Opinion, 2012*). A high AFC may indicate an increased risk of hyper-response to ovarian stimulation and consequently a potential risk of developing ovarian hyperstimulation syndrome (OHSS) (*Broer et al., 2011*). Although there is currently no consensus, AFC cut-offs between 14 and 16 appear to provide the optimal balance between sensitivity and false-positive rate for prediction of hyper-response (*Broer et al., 2011; La Marca et al., 2014; Oudshoorn et al., 2017*). The present prospective study derived a serum cut-off AMH of 1.77 ng/ml for classification of women with an AFC >15. This cut-off was subsequently validated in a second cohort of 856 patients and demonstrated good clinical performance, identifying women with an AFC >15 with high sensitivity (88.34%) and specificity (68.29%).

ROC curve analyses confirmed that the serum AMH cut-off of 1.77 ng/ml provided an optimal balance between sensitivity and specificity and provided further support for the use of AMH-based ovarian reserve determination over other biomarkers. In both the derivation and validation cohorts, AUC for AMH were substantially higher than for FSH and oestradiol, suggesting that these biomarkers are less adequate for ovarian reserve determination compared with AMH. Furthermore, FSH was only moderately correlated with AMH and no correlation was found between oestradiol and AMH. These results are similar to previous ROC analyses by *Anderson et al. (2015)*, but are more robust due to

TABLE 3 AGREEMENT BETWEEN ELECSYS® AMH ABOVE AND BELOW THE 1.77 NG/ML CUT-OFF AND AFC ABOVE AND BELOW 15 IN THE DERIVATION AND VALIDATION COHORTS

No. of patients (%)		AFC 0–15	AFC >15	Total
Derivation cohort	AMH ≤1.77 ng/ml	118 (38.56)	14 (4.58)	132 (43.14)
	AMH >1.77 ng/ml	53 (17.32)	121 (39.54)	174 (56.86)
	Total	171 (55.88)	135 (44.12)	306 (100)
Validation cohort	AMH ≤1.77 ng/ml	280 (32.71)	52 (6.07)	332 (38.79)
	AMH >1.77 ng/ml	130 (15.19)	394 (46.03)	524 (61.21)
	Total	410 (47.90)	446 (52.10)	856 (100)

AFC = antral follicle count; AMH = anti-Müllerian hormone.

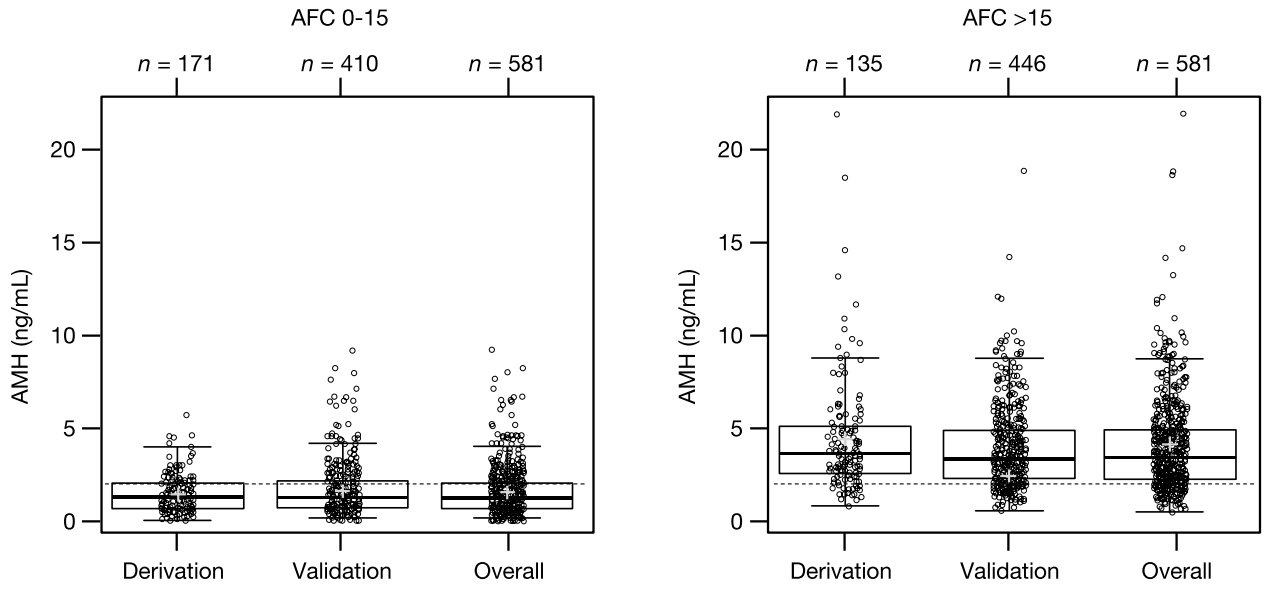

FIGURE 2 Summary of AMH by AFC class (0–15 versus >15) per study cohort (derivation, validation and overall cohort). AFC = antral follicle count; AMH = anti-Müllerian hormone.

the validation step in a second cohort. Using a multivariate modelling approach, it was also shown that the addition of FSH and age to AMH could not improve prediction of AFC class versus AMH alone, as seen in previous studies (*Verhagen et al., 2008*).

In exploratory analyses, a serum AMH cut-off of 2.64 ng/ml was derived for classification of women with an AFC >20, and validated in a cohort of 856 patients using the same approach as for AFC >15. This cut-off demonstrated good clinical performance in the validation cohort, identifying women with an AFC >20 with high sensitivity (74.41%) and specificity (78.18%); however, these findings should be interpreted with caution due to limited power, particularly in the derivation arm where the pre-specified minimum required sample size of 98 per AFC class was not achieved.

Based on these findings, ovarian reserve prediction based on serum AMH concentration (as measured by the Elecsys AMH assay) could be used to inform clinical decision-making, enabling well-informed management of patients. The exact role of AMH classification in clinical decision-making

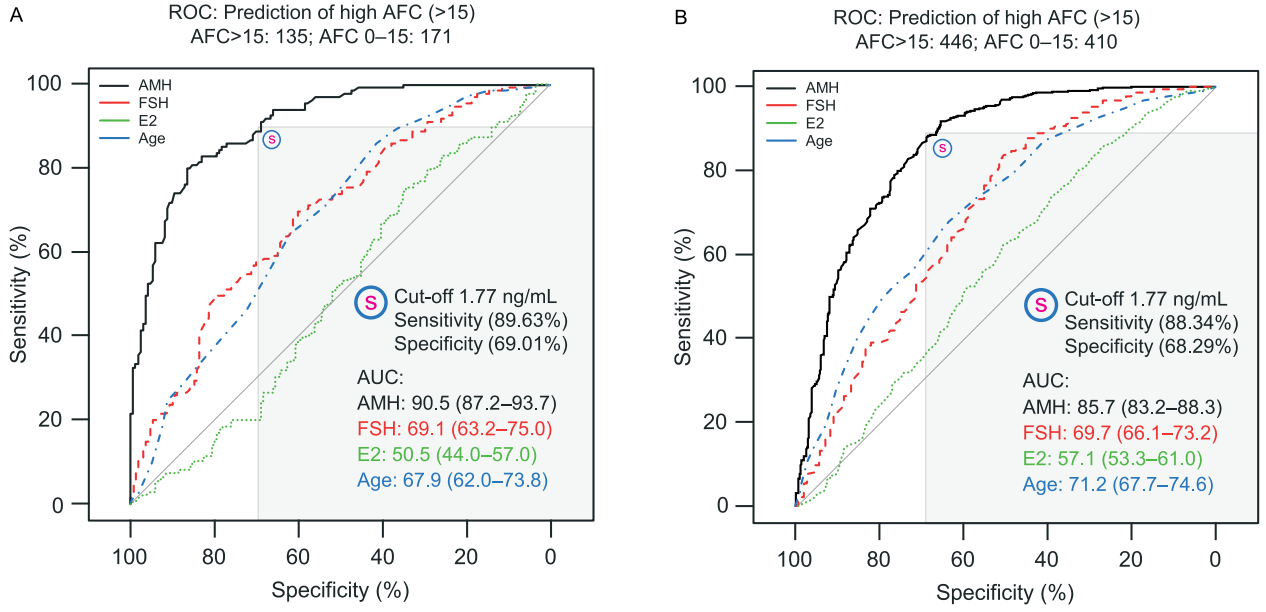

FIGURE 3 ROC curve for prediction of AFC >15 by AMH, FSH and oestradiol in (A) the derivation cohort (*n* = 306) and (B) the validation cohort (*n* = 856). AFC = antral follicle count; AMH = anti-Müllerian hormone; AUC = area under curve; E2 = estradiol; FSH = follicle-stimulating hormone; ROC = receiver operating characteristic.

is currently being investigated in clinical trials (e.g. NCT01956110, NCT01956123 and NCT03564509). Accordingly, the Elecsys AMH assay was recently used to individualize FSH dosing for ovarian stimulation in assisted reproductive therapy as part of a Phase III randomized non-inferiority study (ESTHER-1) (*Nyboe Anderson et al., 2016*). In this setting, an AMH measurement performed at screening was used to guide clinicians when determining the optimal FSH dose (follitropin delta) for ovarian stimulation. The overall aim is to use AMH-based classification to reduce the risk of hyper-response and consequently limit the potential for serious adverse events, such as OHSS.

Previous studies have established a clear association between pretreatment AMH concentrations in serum and ovarian reserve following ovarian stimulation (*Dewailly et al., 2014*). However, in the OPTIMIST study, a post-hoc cost-effectiveness analysis of the utility of AMH suggested that AMH is not cost-effective for individualizing ovarian stimulation (*van Tilborg et al., 2017*). In this study, individualized FSH dosing for ovarian stimulation was determined based solely on AFC. *Nelson and Andersen (2018)* questioned the validity of these retrospective analyses compared with real clinical outcomes. However, evidence from a large-scale, multicentre randomized controlled trial showed that individualized dosing of FSH (follitropin delta) based on pretreatment serum AMH values and body weight resulted in similar efficacy and improved safety compared with conventional ovarian stimulation (*Nyboe Andersen et al., 2017*), supporting the relevance of the findings of this study for clinics.

The results of this study confirm the correlation between serum AMH levels measured using the Elecsys AMH immunoassay and AFC measured by TVUS; however, measurement of AMH by immunoassay offers several advantages over current standard of care (TVUS) for determining ovarian reserve. AMH is stable throughout the menstrual cycle (*La Marca et al., 2007*), so the assay can be performed any time during the cycle and provides a reliable assessment of ovarian reserve; in contrast, AFC assessment is optimally performed between days 2 and 4 of the menstrual cycle. The impact of hormonal contraceptives on AFC results has not been uniformly demonstrated, however (*Birch Peterson et al., 2015; Deb et al., 2012; Johnson et al., 2014*).

The Elecsys AMH assay is fully automated and therefore AMH determination is not susceptible to any inter-observer variability, unlike TVUS which is highly dependent on the individual sonographer. To minimize the effect of such inter-operator variability in this study, participating sonographers were previously familiar with the AFC procedure, received training and followed a standard method. Even so, between-site variability (overall cohort) was greater for AFC than for AMH, which may partially be due to the varying number of sonographers between sites (2–13) reflecting 'real-life' inter-operator variability in AFC determination combined with differences in TVUS specifications between clinics and other unknown factors. The Elecsys AMH assay may also be advantageous in situations where TVUS is unavailable or undesirable (e.g. due to geographical separation or absence of trained sonographers). Potential future applications of AMH include use in paediatric populations to avoid the need for TVUS, and in chemotherapy patients who wish to preserve fertility (*Peigne et al., 2014*).

This study was performed in a selected population presenting to a fertility clinic, representing the target population for ovarian reserve testing, and therefore is not representative of a fertile population. Lack of racial diversity in the patient population, which was primarily Caucasian, may also limit generalizability of findings; however, a cross-sectional study previously found no variation in serum AMH levels between ethnic groups (*Bhide et al., 2015*). Collection of blood samples for AMH measurement on the same day as the AFC procedure allowed accurate representation of serum AMH compared with AFC, as measured in standard care, and provides additional value as it reflects real-life clinical practice. Reproducibility of the results is also demonstrated, as serum analyses were randomized and sent to three separate laboratories for analysis. A key strength of this study was also the use of a separate cohort to validate the strong clinical performance of the derived AMH cut-off. Evaluation of clinical outcomes and relevance of AFC 0–15 was beyond the scope of the current study.

AMH measurement in this study was performed with the Elecsys AMH assay, and the performance of the 1.77 ng/ml cut-off may therefore not be generalizable to other AMH assays. However, the Elecsys AMH immunoassay represents the first assay that is standardized, automated and FDA-approved, and represents an important advance over other commercially available AMH assays, which lack standardization and show substantial between-assay variability (*Su et al., 2014*). Moreover, Su and colleagues showed that conversion of AMH levels derived from different immunoassays using regression equations is potentially highly inaccurate (*Su et al., 2014*).

In conclusion, the Elecsys AMH assay provides a robust, fully automated method to measure serum AMH. Women with AMH values below the cut-off of 1.77 ng/ml are unlikely to have AFC >15. Therefore, the Elecsys AMH assay is a reliable means to determine ovarian reserve and support clinical decision-making for women receiving counselling on infertility treatment.

ACKNOWLEDGEMENTS

We would like to thank the individuals who participated in this study for their time and contribution to this important field of research, and Ying He for assistance with statistical analyses. All authors fulfil the requirements for authorship of the International Committee of Medical Journal Editors. Medical writing assistance, under the guidance of the authors, was provided by David Evans, PhD, and Chloe Fletcher, MSc (Gardiner-Caldwell Communications, Macclesfield, UK), and was funded by Roche Diagnostics International Ltd (Rotkreuz, Switzerland).

This study was funded by Roche Diagnostics Operations, Inc. (Indianapolis, IN, USA). Medical writing support was funded by Roche Diagnostics International Ltd (Rotkreuz, Switzerland).

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rbmo.2018.12.041.

REFERENCES

Anckaert, E., Öktem, M., Thies, A., Cohen-Bacrie, M., Daan, N.M., Schiettecatte, J., Müller, C., Topcu, D., Gröning, A., Ternaux, F., Engel, C., Engelmann, S., Milczynski, C. **Multicentre analytical performance evaluation of a fully automated anti-Müllerian hormone assay and reference interval determination.** Clin. Biochem. 2016; 49: 260–267

Anderson, R.A., Anckaert, E., Bosch, E., Dewailly, D., Dunlop, C.E., Fehr, D., Nardo, L., Smitz, J., Tremellen, K., Denk, B., Geistanger, A., Hund, M. **Prospective study into the value of the automated Elecsys antimullerian hormone assay for the assessment of the ovarian growing follicle pool.** Fertil. Steril. 2015; 103: 1074–1080

ASRM Committee Opinion. Testing and interpreting measures of ovarian reserve: a committee opinion. Fertil. Steril. 2012; 98: 1407–1415

Bhide, P., Gudi, A., Shah, A., Homburg, R. **Serum anti-Mullerian hormone levels across different ethnic groups: a cross-sectional study.** BJOG 2015; 122: 1625–1629

Birch Petersen, K., Hvidman, H.W., Forman, J.L., Pinborg, A., Larsen, E.C., Macklon, K.T., Sylvest, R., Andersen, A.N. **Ovarian reserve assessment in users of oral contraception seeking fertility advice on their reproductive lifespan.** Hum. Reprod. 2015; 30: 2364–2375

Broekmans, F.J., de Ziegler, D., Howles, C.M., Gougeon, A., Trew, G., Olivennes, F. **The antral follicle count: practical recommendations for better standardization.** Fertil. Steril. 2010; 94: 1044–1051

Broer, S.L., Mol, B., Dolleman, M., Fauser, B.C., Broekmans, F.J. **The role of anti-Mullerian hormone assessment in assisted reproductive technology.** Curr. Opin. Obstet. Gynecol. 2010; 22: 193–201

Broer, S.L., Dolleman, M., Opmeer, B.C., Fauser, B.C., Mol, B.W., Broekmans, F.J. **AMH and AFC as predictors of excessive response in controlled ovarian hyperstimulation: a meta-analysis.** Hum. Reprod. Update. 2011; 17: 46–54

Clinical and Laboratory Standards Institute. **User verification of performance for precision and trueness; approved guideline. 2.** 2005 CLSI document EP15-A2 Wayne, PA, USA

Deb, S., Campbell, B.K., Pincott-Allen, C., Clewes, J.S., Cumberpatch, G., Raine-Fenning, N.J. **Quantifying effect of combined oral contraceptive pill on functional ovarian reserve as measured by serum anti-Müllerian hormone and small antral follicle count using three-dimensional ultrasound.** Ultrasound. Obstet. Gynecol. 2012; 39: 574–580

DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L. **Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach.** Biometrics 1988; 44: 837–845

Dewailly, D., Andersen, C.Y., Balen, A., Broekmans, F., Dilaver, N., Fanchin, R., Griesinger, G., Kelsey, T.W., La Marca, A., Lambalk, C., Mason, H., Nelson, S.M., Visser, J.A., Wallace, W.H., Andersen, R.A. **The physiology and clinical utility of anti-Müllerian hormone in women.** Hum. Reprod. Update 2014; 20: 370–385

Hehenkamp, W.J., Looman, C.W., Themmen, A.P., de Jong, F.H., Te Velde, E.R., Broekmans, F.J. **Anti-Mullerian hormone levels in the spontaneous menstrual cycle do not show substantial fluctuation.** J. Clin. Endocrinol. Metab. 2006; 91: 4057–4063

Johnson, L.N.C., Sammel, M.D., Dillon, K.E., Lechtenberg, L., Schanne, A., Gracia, C.R. **Anti-Mullerian Hormone and antral follicle count are lower in female cancer survivors and healthy women taking hormonal contraception.** Fertil. Steril. 2014; 102: 774–781

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

La Marca, A., Sighinolfi, G., Radi, D., Argento, C., Baraldi, E., Artenisio, A.C., Stabile, G., Volpe, A. **Anti-Mullerian hormone (AMH) as a predictive marker in assisted reproductive technology (ART).** Hum. Reprod. Update. 2010; 16: 113–130

La Marca, A., Sunkara, S.K. **Individualization of controlled ovarian stimulation in IVF using ovarian reserve markers: from theory to practice.** Hum. Reprod. Update. 2014; 21: 124–140

Nelson, S.M., Andersen, R.A. **Derailing individualized ovarian stimulation.** Hum. Reprod. 2018; 33: 980–981

Nyboe Andersen, A., Nelson, S.M., Fauser, B.C., Garcia-Velasco, J.A., Klein, B.M., Arce, J.C. ESTHER-1 study group. **Individualized versus conventional ovarian stimulation for *in vitro* fertilization: a multicentre, randomized, controlled, assessor-blinded, phase 3 noninferiority trial.** Fertil. Steril. 2016; 107: 387–396

Oudshoorn, S.C., van Tilborg, T.C., Eijkemans, M.J.C., Oosterhuis, G.J.E., Friederich, J., van Hooff, M.H.A., van Santbrink, E.J.P., Brinkhuis, E.A., Smeenk, J.M.J., Kwee, J., de Koning, C.H., Groen, H., Lambalk, C.B., Mol, B.W.J., Broekmans, F.J.M., Torrance, H.L. OPTIMIST study group. **Individualized versus standard FSH dosing in women starting IVF/ICSI: an RCT. Part 2: The predicted hyper responder.** Hum. Reprod. 2017; 32: 2506–2514

Peigne, M., Decanter, C. **Serum AMH level as a marker of acute and long-term effects of chemotherapy on the ovarian follicular content: a systematic review.** Reprod. Biol. Endocrinol. 2014; 12: 26

Rotterdam ESHRE/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

Su, H.I., Sammel, M.D., Homer, M.V., Bui, K., Haunschild, C., Stanczyk, F.Z. **Comparability of antimullerian hormone levels among commercially available immunoassays.** Fertil. Steril. 2014; 101: 1766–1772

van Disseldorp, J., Lambalk, C.B., Kwee, J., Looman, C.W., Eijkemans, M.J., Fauser, B.C., Broekmans, F.J. **Comparison of inter- and intra-cycle variability of anti-Mullerian hormone and antral follicle counts.** Hum. Reprod. 2010; 25: 221–227

van Rooij, I.A., Broekmans, F.J., te Velde, E.R., Fauser, B.C., Bancsi, L.F., de Jong, F.H., Fauser, B.J., Themmen, A.P. **Serum anti-Mullerian hormone levels: a novel measure of ovarian reserve.** Hum. Reprod. 2002; 17: 3065–3071

van Tilborg, T.C., Oudshroom, S.C., Eijkemans, M.J.C., Mochtar, M.H., van Golde, R.J.T., Hoek, A., Kuchenbecker, W.K.H., Fleischer, K., de Bruin, J.P., Groen, H., van Wely, M., Lambalk, C.B., Laven, J.S.E., Mol, B.W.J., Broekmans, F.J.M., Torrance, H.L. OPTIMIST study group. **Individualised FSH dosing based on ovarian reserve testing in women starting IVF/ICSI: a multicentre trial and cost-effectiveness analysis.** Hum Reprod. 2017; 32: 2485–2495

Verhagen, T.E.M., Hendriks, D.J., Bancsi, L.F.J.M.M., Mol, B.W.J., Broekmans, F.J.M. **The accuracy of multivariate models predicting ovarian reserve and pregnancy after *in vitro* fertilization: a meta-analysis. 2008.** Hum. Reprod. Update. 2008; 14: 95–100

Weenen, C., Laven, J.S., Von Bergh, A.R., Cranfield, M., Groome, N.P., Visser, J.A., Kramer, P., Fauser, B.C., Themmen, A.P. **Anti-Mullerian hormone expression pattern in the human ovary: potential implications for initial and cyclic follicle recruitment.** Mol. Hum. Reprod. 2004; 10: 77–83

Received 9 July 2018; received in revised form 10 December 2018; accepted 12 December 2018.