human reproduction open

ORIGINAL ARTICLE

The correlation between AMH assays differs depending on actual AMH levels

Å. Magnusson [ID] [1,2,*], G. Oleröd[3], A. Thurin-Kjellberg[1,2], and C. Bergh[1,2]

[1]Department of Obstetrics, Gynaecology and Reproductive Medicine, Sahlgrenska University Hospital, SE-413 45 Gothenburg, Sweden
[2]Institute of Clinical Sciences, Sahlgrenska Academy, Gothenburg University [3]Department of Clinical Chemistry, Sahlgrenska University Hospital, SE-413 45 Gothenburg, Sweden

*Correspondence address. E-mail: asa.magnusson@vgregion.se [ID] orcid.org/0000-0003-2548-1114

STUDY QUESTION: What is the correlation of serum anti-Müllerian hormone (AMH) levels between two frequently used laboratory assays?

SUMMARY ANSWER: A considerable difference was found in serum AMH levels measured with the two different assays, particularly for low AMH values.

WHAT IS KNOWN ALREADY: AMH is regarded as being a robust, highly sensitive and specific biomarker for ovarian response and has become widely used as the basis for fertility treatment decisions. However, several available assays with different reference values, in addition to inter-laboratory variations and issues of sample stability, make interpretation of the AMH values and their clinical implications complicated.

STUDY DESIGN, SIZE, DURATION: An observational study was performed including 269 serum samples from infertile women, originating from a RCT conducted in 2013-2016 (www.clinicaltrials.gov NCT02013973).

PARTICIPANTS/MATERIALS, SETTING, METHOD: Serum AMH levels analysed with the Modified Beckman Coulter Gen II ELISA assay (Premix method) were compared to AMH levels analysed with the Beckman Coulter Gen II ELISA original assay (Gen II original). All samples were handled identically and analysed with the two assays in a parallel setting.

MAIN RESULTS AND THE ROLE OF CHANCE: The slope of the regression line showed a mean of 18% higher values with the Premix method compared to the Gen II original assay, and more than 40% higher values for AMH levels in the lower range.

LIMITATIONS, REASONS FOR CAUTION: The Gen II original assay is no longer in clinical use as it has been replaced by the Premix method, which, in turn, recently has been further developed into an automated method.

WIDER IMPLICATIONS OF THE FINDINGS: The finding of differences in AMH levels between assays is clinically important and may imply an incorrect classification in the assessment of ovarian reserve. The robustness of serum AMH as a marker for ovarian reserve and as a tool for fertility counselling has to be investigated further. There is an urgent need for international standards on interpretation of AMH values for different assays.

STUDY FUNDING/COMPETING INTERESTS: Financial support was received through Sahlgrenska University Hospital (ALFGBG-70940) and the Hjalmar Svensson Research Foundation. None of the authors declares any conflict of interest.

Key words: anti-Müllerian hormone / assay / comparison

> # WHAT DOES THIS MEAN FOR PATIENTS?
>
> This study looked at the results of Anti-Müllerian hormone (AMH) tests, which are often used in fertility assessments and for treatment decisions, to see whether the results of different types of tests produced different results. AMH tests are used in IVF to assess how a woman may respond to the drugs and more generally to give a prediction of ovarian reserve (the remaining stock of eggs that the ovaries can produce).
>
> There have been previous concerns that the results from AMH tests may vary depending on the laboratory doing the analysis, and also some questions about whether AMH levels can change if analysed on different days in the menstrual cycle. There are also concerns about the different tests themselves giving different results. This particular study looked at two of these tests. The researchers noted that one of them is no longer used in clinics, but even so found some considerable differences in the results, particularly for women with lower AMH results.
>
> The researchers say that they are concerned that there is a real risk that women who have the test as part of a fertility check may receive incorrect information and have made an urgent call for international standards for the different AMH tests to ensure that results from different tests are interpreted correctly.

Introduction

Anti-Müllerian hormone (AMH) is produced in the granulosa cells of pre-antral and small antral follicles and has become widely used for assessment of ovarian reserve and prediction of ovarian response to hormone stimulation for IVF.

As the number of follicles diminishes with increasing age, there is consequently a physiological decrease in serum AMH levels, and serum AMH levels have been suggested to be a predictive factor for the time to onset of menopause (Broer *et al.*, 2011; Freeman *et al.*, 2012; Dólleman *et al.*, 2014; Depmann *et al.*, 2016). Furthermore, serum AMH has been suggested as a screening biomarker for fertility counselling (Broer *et al.*, 2014), although several studies have shown a lack of correlation between serum AMH levels and natural fecundity (Hagen *et al.*, 2012; Zarek *et al.*, 2015; Depmann *et al.*, 2017; Hvidman *et al.*, 2017).

Several studies have shown that the serum AMH level is a good predictive biomarker of ovarian response to hormone stimulation during IVF treatment (Nelson *et al.*, 2007; Brodin *et al.*, 2015; Nelson *et al.*, 2015; Iliodromiti *et al.*, 2015) and in recent years it has become widely used as the basis for treatment decisions, especially regarding the hormone dosage used for ovarian stimulation (Nelson *et al.*, 2009; Magnusson *et al.*, 2017; Nyboe Andersen *et al.*, 2017). Serum AMH levels have also been found to be highly correlated to the antral follicle count (AFC), another biomarker with similar sensitivity for ovarian response (Brodin *et al.*, 2015; Nyboe Andersen *et al.*, 2017). AMH, being a quantitative laboratory parameter, is expected to lack the inter-observer variation that is inevitably associated with the AFC assessment performed by sonography. Nevertheless, considerable variations were recently described in a study comparing AMH values from 10 different laboratories analysing the same serum samples using the same assay (Gen II original) (Zuvela *et al.*, 2013). Serum AMH levels have been considered to be cycle independent, making it a suitable parameter for the assessment of ovarian reserve, as a blood sample can be taken on any day of the cycle (Van Disseldorp *et al.*, 2010). However, in recent years several studies have reported considerable intra-cycle variation (Wunder *et al.*, 2008; Overbeek *et al.*, 2012). A recently published study showed that assessment of multiple blood samples from the same cycle resulted in alteration of AMH class according to ovarian response in almost 30% of the women (Hadlow *et al.*, 2016) implicating a risk of misclassification and an incorrect treatment decision depending on the cycle day of blood sampling.

There have also been concerns regarding the different assays used for the analysis of serum AMH. In 2010, the first generation assays by Diagnostics Systems Laboratory (DSL) and Immunotech (IOT) were replaced by the Gen II original assay. The correlation between these assays was investigated and considered to be good, although the AMH values measured with the Gen II original assay were found to be approximately 40% higher as compared to the DSL assay (Wallace *et al.*, 2011). However, in 2013 Beckman Coulter reported instability in their assay owing to complement interference and analyses of serum samples after storage at room temperature resulted in AMH values 20–40% lower with the Gen II original assay compared to the DSL assay (Rustamov *et al.*, 2012). The Gen II original assay was modified with a pre-diluting step before analysis to eliminate complement interference and the Modified AMH assay (Premix method) was released. A study, including a mixed population of 28 non-pregnant/early pregnant and 42 infertile women, compared AMH values analysed with Gen II original assay to values analysed when premixing serum with the assay buffer. This study suggested an increase of up to 2-fold for AMH values in the infertile group after the pre-diluting procedure (Han *et al.*, 2014). In addition, the same study found a significant increase in AMH values when serum samples were stored in room temperature for more than 8 h.

A recent study has compared three new AMH assays—the Ultra-Sensitive AMH/MIS ELISA kit (Ansh Labs), the automated Access AMH assay (Beckman-Coulter), and the Elecsys® AMH Immunoassay (Roche)—to the Gen II original in a parallel setting. The study showed good correlations between the assays, but significantly different AMH values when the four assays were compared on the same serum samples (Li *et al.*, 2016).

The aim of this study was to investigate, in a parallel setting, the relationship between AMH values analysed simultaneously with the Gen II original assay and the Premix method assay in a secondary analysis of a RCT of well characterized infertile women.

Material and Methods

A recently published RCT compared two algorithms used for deciding the starting dose of gonadotrophin for follicular stimulation before IVF

(Magnusson et al., 2017). A serum sample was taken at the first visit to the clinic and immediately stored at −70°C. Randomization was performed after down-regulation and before the start of stimulation. If randomized to the algorithm including AMH, the patient´s serum sample was thawed, AMH was analysed and the result was used to decide the gonadotrophin starting dose.

After the completion of the study, the serum samples from the patients randomized to the non-AMH algorithm were thawed and all samples were analysed simultaneously with the Premix method and the Gen II original. In total, 269 serum samples were analysed and compared.

The performance of the Premix method was tested in our laboratory and the total coefficient of variance (total CV) for serum AMH was 10.6% at the level of 2.6 ng/ml, 4.9% at the level of 4.9 ng/ml and 10.2% at the level of 8 ng/ml. The lower limit of quantification was 0.2 ng/ml (Magnusson et al., 2017). A previous analysis of the Gen II original assay had shown similar performance.

Microsoft Excel was used for analysis and graphic description of the results.

The study was approved by the regional ethics committee at Gothenburg University (Dnr 219-12).

Results

The characteristics of the study population are summarized in Table I. The correlation coefficient between the two assays was 0.94. When comparing results from Beckman Coulter AMH Gen II ELISA using the Original method and the modified Premix method, the difference varied depending on the concentration of serum AMH. When all samples were analysed including AMH concentrations up to 20 μg/l, the regression line had a slope of 1.176 and an intercept of 0.602 ($r =$ 0.944). The slope of the regression line showed a deviation between the methods, with approximately 18% higher values with the Premix method ($n = 269$). If the results using the Original method were

Table I Characteristics of the study population ($N =$ 269) in a comparison of anti-Müllerian hormone assays.

	Mean	SD	Minimum	Maximum	n/N
BMI (kg/m^2)	23.53	3.67	18.00	35.00	
Age (years)	32.15	3.94	20.00	39.00	
PCO[a]					50/269 (18%)
AMH BCG II O	3.29[a]	3.09	0.10	18.60	
AMH BCG II P	4.43[b]	3.79	0.20	22.00	

AMH, anti-Müllerian hormone.
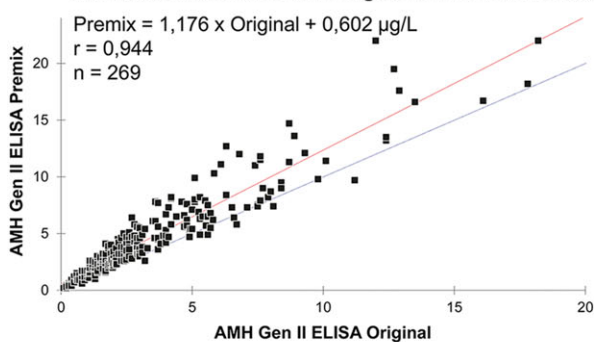[a]Polycystic ovaries: according to Rotterdam criteria.
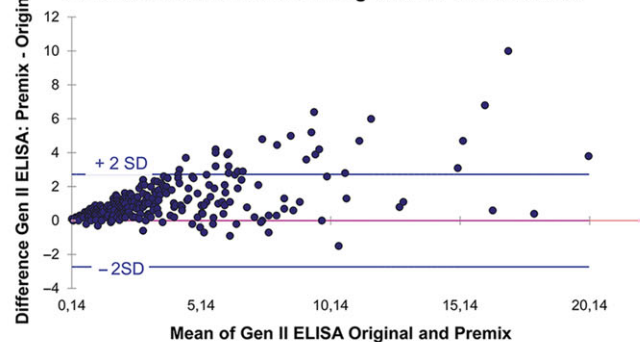[b]Unit for both AMH assays: μg/l.
Beckman Coulter Gen II O: Gen II original method.
Beckman Coulter Gen II P: Gen II Premix method.

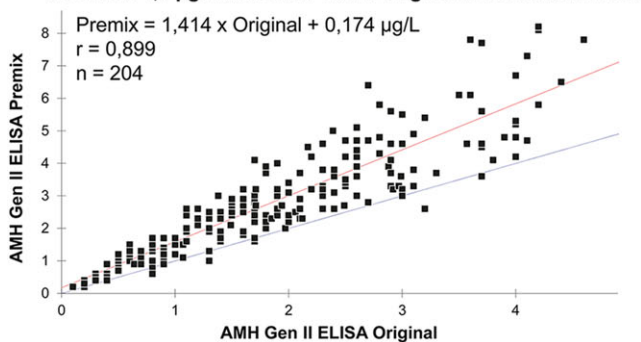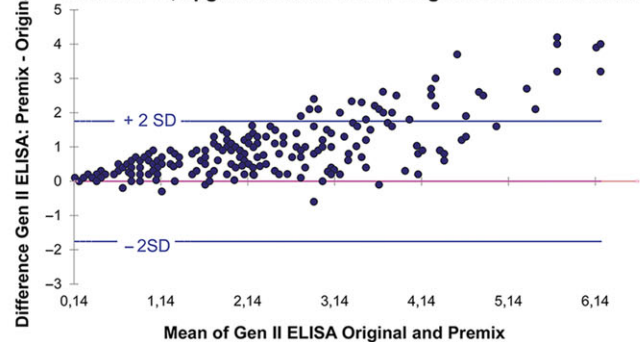**Figure 1** Comparison of the results from the Beckman Coulter AMH Gen II ELISA, using the Original method and the modified Premix method. The upper charts show all anti-Müllerian hormone (AMH) results (up to about 20 μg/l) and the charts below show results where the Original method concentrations are up to 4.5 μg/l (corresponding to an average of about 6.5 μg/l with the Premix method). The left panels represent the linear least squares regression plots and the right panels represent the Bland–Altman plots. r = correlation coefficient; n = number of samples.

limited to concentrations up to 4.5 μg/l (corresponding to an average of about 6.5 μg/l with the Premix method), the regression line had a slope of 1.414 and an intercept of 0.174 ($r = 0.899$). Thus, in the lower concentration range, the slope of the regression line showed a greater deviation between the methods, with more than 40% higher values with the Premix method ($n = 204$). In addition, the method comparisons showed a large variance around the regression line, indicating considerable differences in individual results between the two methods. The Scatter plot diagrams with linear least squares regression lines, regression equation, as well as the Bland–Altman plots are shown in Fig. 1.

Discussion

The principal finding of this study, showing that the discrepancy in AMH values between the two assays is larger for low AMH values, is new and provides valuable information for physicians and their patients. The aim of the study was to provide more insight into the relationship between AMH values analysed with two frequently used assays where the second is an improved version of the first, including a pre-dilution step to solve the problem of complement interference. AMH is considered to be a favourable biomarker for clinical use and a large number of publications have described the use of AMH in ART and fertility counselling. The use of AMH in fertility counselling is especially troublesome given that the discrepancy between the original and the modified assay seems to be more pronounced for AMH values in the low range. It is also obvious that there is a certain variation between laboratories and a lack of international standards for interpretation, also for the new automatized assays. Hence there is a considerable risk of incorrect information being provided to women who want to take decisions on advancing or postponing pregnancy or even having oocyte retrieval and vitrification for fertility preservation. A recent study investigating low AMH values in subfertile and fertile women did not show any correlation between low AMH values and subfertility (Somigliana *et al.*, 2015). Low AMH values have also been described in users of the combined contraceptive pill (Bentzen *et al.*, 2012) and low AMH values in women with hypothalamic amenorrhoea have not been found to indicate inability to respond to ovarian stimulation (Billington and Corenblum 2016).

The original classification of patients scheduled for IVF in the categories high-, normo and low responders was based on the DSL assay (Nelson *et al.*, 2009) and as new assays have been released different conversion factors have been suggested for interpretation between the assays (Nelson and La Marca 2011; Wallace *et al.*, 2011). However, there may be considerable variation between laboratories using the same assay (Zuvela *et al.*, 2013) and handling and storage of samples may vary and therefore influence the analysis results (Rustamov *et al.*, 2012; Han *et al.*, 2014) making interpretation and generalization problematic. Another issue of great concern is how to interpret the serum levels between the different assays that are now commercially available. A recent study compared the two most frequently used automatized assays (Elecsys® and Access) and found that 28% of patients having had their starting dose decided based on serum AMH level might have been misclassified if values from the two assays were used with the same reference intervals (Iliodromiti *et al.*, 2017). Previous studies comparing the Gen II Original assay and the Premix method (Han *et al.*, 2014; Bonifacio *et al.*, 2015) have included pregnant women and although comparing assays for AMH in pregnant

women has a certain value it seems more relevant to assess AMH in infertile women, being the population where it is currently used for prediction of fertility and gonadotrophin stimulation regimes worldwide (Iliodromiti *et al.*, 2017).

The large number of well characterized infertile patients in this study, together with the optimal handling of samples, i.e. immediate freezing, thawing just before analysis and simultaneous analysis using both assays, adds useful information about the relationship between the Original Beckman Coulter AMH Gen II ELISA assay and the currently used Modified AMH assay. However, it should be observed that the original Gen II assay is no longer in use, being replaced by the Premix AMH assay and now further developed by Beckman Coulter into the automated Access AMH Immunoassay. The correlation between the manual and the automated method has been investigated and shown to be very good (Demirdjian *et al.*, 2016).

In conclusion, this study showed substantially higher AMH values with the Premix method compared to the Gen II ELISA original assay, and particularly for AMH levels in the lower range. The considerable difference in AMH levels measured with different assays creates concerns both clinically and scientifically. International standards for reference intervals, handling of samples and conversion factors between different assays are urgently needed.

Authors' roles

All the authors contributed to the conception and design of the study and to the analysis and interpretation of data. G.O. performed the AMH analyses and contributed to the technical data on the AMH assays. Å.M. drafted the manuscript. Å.M. and C.B. finalized it and all the authors revised the manuscript and approved the final version.

Funding

Conflict of interest

None declared.

References

Bentzen JG, Forman JL, Pinborg A, Lidegaard Ø, Larsen EC, Friis-Hansen L, Johannsen TH, Nyboe Andersen. Ovarian reserve parameters: a comparison between users and non-users of hormonal contraception. *Reprod Biomed Online* 2012;**25**:612–619.

Billington EO, Corenblum B. Anti-Mullerian hormone levels do not predict response to pulsatile GnRH in women with hypothalamic amenorrhea. *Gynecol Endocrinol* 2016;**32**:728–732.

Bonifacio M, Bradley C, Karia S, Livingstone M, Bowman M, McArthur S. The original Beckman Coulter Generation II assay significantly underestimates AMH levels compared with the revised protocol. *J Assist Reprod Genet* 2015;**32**:1691–1696.

Brodin T, Hadziosmanovic N, Berglund L, Olofsson M, Holte J. Comparing four ovarian reserve markers-associations with ovarian response and live births after assisted reproduction. *Acta Obstet Gynecol Scand* 2015;**94**:1056–1063.

Broer S, Eijkemans M, Scheffer GJ, van Rooij I, de Vet A, Themmen A, Laven JS, de Jong F, Te Velde ER, Fauser B *et al*. Anti-mullerian hormone predicts menopause: a long-term follow-up study in normoovulatory women. *J Clin Endocrinol Metab* 2011;**96**:2532–2539.

Broer S, Broekmans F, Laven J, Fauser B. Anti-Müllerian hormone: ovarian reserve testing and its potential clinical implications. *Hum Reprod Update* 2014;**20**:688–701.

Demirdjian G, Bord S, Lejeune C, Masica R, Rivière D, Nicouleau L, Denizot P, Marquet PY. Performance characteristics of the Access AMH assay for the quantitative determination of anti-Müllerian hormone (AMH) levels on the Access* family of automated immunoassay systems. *Clin Biochem* 2016;**49**:1267–1273.

Depmann M, Eijkemans M, Broer S, Scheffer G, van Rooij I, Laven J, Broekmans F. Does anti-Müllerian hormone predict menopause in the general population? Results of a prospective ongoing cohort study. *Hum Reprod* 2016;**31**:1579–1587.

Depmann M, Broer S, Eijkemans M, van Rooij I, Scheffer G, Heimensem J, Mol B, Broekmans F. Anti-Müllerian hormone does not predict time to pregnancy: results of a prospective cohort study. *Gynecol Endocrinol* 2017;**33**:644–648.

Dólleman M, Depmann M, Eijkemans M, Heimensem J, Broer S, van der Stroom E, Laven J, Van Rooij I, Scheffer G, Peeters P *et al*. Anti-Mullerian hormone is a more accurate predictor of individual time to menopause than mother's age at menopause. *Hum Reprod* 2014;**29**:584–591.

Freeman E, Sammel MD, Lin H, Gracia C. Anti-müllerian hormone as a predictor of time to menopause in late reproductive age women. *J Clin Endocrinol Metab* 2012;**97**:1673–1680.

Hadlow N, Brown S, Habib A, Wardrop R, Joseph J, Gillett M, Maguire R, Conradie J. Quantifying the intraindividual variation of antimüllerian hormone in the ovarian cycle. *Fertil Steril* 2016;**106**:1230–1237.

Hagen C, Vestergaard S, Juul A, Skakkebæk N, Andersson A, Main K, Hjøllund N, Ernst E, Bonde J, Anderson RA *et al*. Low concentration of circulating antimüllerian hormone is not predictive of reduced fecundability in young healthy women: a prospective cohort study. *Fertil Steril* 2012;**98**:1602–1608.

Han X, McShane M, Sahertian R, White C, Ledger W. Pre-mixing serum samples with assay buffer is a prerequisite for reproducible anti-Müllerian hormone measurement using the Beckman Coulter Gen II assay. *Hum Reprod* 2014;**29**:1042–1048.

Hvidman H, Bang A, Priskorn L, Scheike T, Birch Petersen K, Nordkap L, Loft A, Pinborg A, Tabor A, Jørgensen N *et al*. Anti-Müllerian hormone levels and fecundability in women with a natural conception. *Eur J Obstet Gynecol Reprod Biol* 2017;**217**:44–52.

Iliodromiti S, Anderson RA, Nelson SM. Technical and performance characteristics of anti-Müllerian hormone and antral follicle count as biomarkers of ovarian response. *Human Reprod Update* 2015;**21**:698–710.

Iliodromiti S, Salie B, Dewailly D, Fairburn C, Fanchin R, Fleming R, Li HWR, Lukaszuk K, Ng EHY, Pigny P *et al*. Non-equivalence of anti-Müllerian hormone automated assays—clinical implications for use as a companion diagnostic for individualised gonadotrophin dosing. *Hum Reprod* 2017;**32**:1710–1715.

Li HW, Wong BP, Ip WK, Yeung WSB, Ho PC, Ng EHY. Comparative evaluation of three new commercial immunoassays for anti-Müllerian hormone measurement. *Hum Reprod* 2016;**31**:2796–2802.

Magnusson Å, Nilsson L, Oleröd G, Thurin-Kjellberg A, Bergh C. The addition of anti-Müllerian hormone in an algorithm for individualized hormone dosage did not improve the prediction of ovarian response—a randomized, controlled trial. *Hum Reprod* 2017;**32**:811–819.

Nelson SM, La Marca A. The journey from the old to the new AMH assay: how to avoid getting lost in the values. *Reprod Biomed Online* 2011;**23**:411–420.

Nelson SM, Yates RW, Fleming R. Serum anti-Müllerian hormone and FSH: prediction of live birth and extremes of response in stimulated cycles—implications for individualization of therapy. *Hum Reprod* 2007;**22**:2414–2421.

Nelson SM, Yates RW, Lyall H, Jamieson M, Traynor I, Gaudoin M, Mitchell P, Ambrose P, Fleming R. Anti-Müllerian hormone-based approach to controlled ovarian stimulation for assisted conception. *Hum Reprod* 2009;**24**:867–875.

Nelson S, Fleming R, Gaudoin M, Choi B, Santo-Domingo K, Yao M. Antimüllerian hormone levels and antral follicle count as prognostic indicators in a personalized prediction model of live birth. *Fertil Steril* 2015;**104**:325–332.

Nyboe Andersen A, Nelson SM, Fauser BC, Garcia-Velasco JA, Klein BM, Arce JC. Individualized versus conventional ovarian stimulation for in vitro fertilization: a multicenter, randomized, controlled, assessor-blinded, phase 3 noninferiority trial. *Fertil Steril* 2017;**107**:387–396.

Overbeek A, Broekmans F, Hehenkamp W, Wijdeveld M, van Disseldorp J, van Dulmen-de Broeder E, Lambalk C. Intra-cycle fluctuations of anti-Müllerian hormone in normal women with a regular cycle: a re-analysis. *Reprod Biomed Online* 2012;**24**:664–669.

Rustamov O, Smith A, Roberts SA, Yates AP, Fitzgerald C, Krishnan M, Nardo LG, Pemberton PW. Anti-Müllerian hormone: poor assay reproducibility in a large cohort of subjects suggests sample instability. *Hum Reprod* 2012;**27**:3085–3091.

Somigliana E, Lattuada D, Colciaghi B, Filippi F, La Vecchia I, Tirelli A, Baffero G, Paffoni A, Persico N, Bolis G *et al*. Serum anti-Müllerian hormone in subfertile women. Acta Obstet Gynecol Scand 2015;**94**:1307–1312.

Wallace AM, Faye SA, Fleming R, Nelson SM. A multicenter evaluation of the new Beckman Coulter anti-Mullerian hormone immunoassay (AMH Gen II). *Ann Clin Biochem* 2011;**48**:370–373.

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

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

Zarek S, Mitchell E, Sjaarda L, Mumford S, Silver R, Stanford J, Galai N, White MV, Schliep K, DeCherney A *et al*. Is Anti-Müllerian hormone associated with fecundability? Findings from the EAGeR trial. *J Clin Endocrinol Metab* 2015;**100**:4215–4221.

Zuvela E, Walls M, Matson P. Within-laboratory and between-laboratory variability in the measurement of anti-müllerian hormone determined within an external quality assurance scheme. *Reprod Biol* 2013;**13**:255–257.