

# Anti-Müllerian hormone: poor assay reproducibility in a large cohort of subjects suggests sample instability

Oybek Rustamov<sup>1,\*</sup>, Alexander Smith<sup>2</sup>, Stephen A. Roberts<sup>3</sup>, Allen P. Yates<sup>2</sup>, Cheryl Fitzgerald<sup>1</sup>, Monica Krishnan<sup>4</sup>, Luciano G. Nardo<sup>5</sup>, and Philip W. Pemberton<sup>2</sup>

<sup>1</sup>Department of Reproductive Medicine, St Mary's Hospital, Central Manchester University Hospitals Foundation Trust, Manchester M13 0JH, UK <sup>2</sup>Department of Clinical Biochemistry, Central Manchester Foundation Trust, Manchester M13 9WL, UK <sup>3</sup>Health Sciences – Methodology, Manchester Academic Health Science Centre (MAHSC), University of Manchester, Manchester M13 9PL, UK <sup>4</sup>School of Medicine, University of Manchester, Manchester M13 9WL, UK <sup>5</sup>Reproductive Medicine and Gynaecology Unit, GyneHealth, Manchester M3 4DN, UK

\*Correspondence address. E-mail: oybek.rustamov@cmft.nhs.uk, oybek\_rustamov@yahoo.co.uk

Submitted on March 1, 2012; resubmitted on June 8, 2012; accepted on June 14, 2012

**STUDY QUESTION:** What is the variability of anti-Müllerian hormone (AMH) concentration in repeat samples from the same individual when using the Gen II assay and how do values compare to Gen I [Diagnostic Systems Ltd (DSL)] assay results?

**SUMMARY ANSWER:** The Gen II AMH assay displayed appreciable variability, which can be explained by sample instability.

**WHAT IS KNOWN ALREADY:** AMH is the primary predictor of ovarian performance and is used to tailor gonadotrophin dosage in cycles of IVF/ICSI and in other routine clinical settings. Thus, a robust, reproducible and sensitive method for AMH analysis is of paramount importance. The Beckman Coulter Gen II ELISA for AMH was introduced to replace earlier DSL and Immunotech assays. The performance of the Gen II assay has not previously been studied in a clinical setting.

**STUDY DESIGN, SIZE AND DURATION:** We studied an unselected group of 5007 women referred for fertility problems between 1 September 2008 and 25 October 2011; AMH was measured initially using the DSL AMH ELISA and subsequently using the Gen II assay. AMH values in the two assays were compared using a regression model in log(AMH) with a quadratic adjustment for age. Additionally, women ( $n = 330$ ) in whom AMH had been determined in different samples using both the DSL and Gen II assays (paired samples) identified and the difference in AMH levels between the DSL and Gen II assays was estimated using the age-adjusted regression analysis. A subset of 313 women had repeated AMH determinations ( $n = 646$  samples) using the DSL assay and 87 women had repeated AMH determinations using the Gen II assay ( $n = 177$  samples) were identified. A mixed effects model in log(AMH) was utilized to estimate the sample-to-sample (within-subject) coefficients of variation of AMH, adjusting for age. Laboratory experiments including sample stability at room temperature, linearity of dilution and storage conditions used anonymized samples.

**MAIN RESULTS AND THE ROLE OF CHANCE:** In clinical practice, Gen II AMH values were ~20% lower than those generated using the DSL assay instead of the 40% increase predicted by the kit manufacturer. Both assays displayed high within-subject variability (Gen II assay CV = 59%, DSL assay CV = 32%). In the laboratory, AMH levels in serum from 48 subjects incubated at RT for up to 7 days increased progressively in the majority of samples (58% increase overall). Pre-dilution of serum prior to assay, gave AMH levels up to twice that found in the corresponding neat sample. Pre-mixing of serum with assay buffer prior to addition to the microtitre plate gave higher readings (72% overall) compared with sequential addition. Storage at  $-20^{\circ}\text{C}$  for 5 days increased AMH levels by 23% compared with fresh samples. The statistical significance of results was assessed where appropriate.

**LIMITATIONS, REASONS FOR CAUTION:** The analysis of AMH levels is a retrospective study and therefore we cannot entirely rule out the existence of differences in referral practices or changes in the two populations.

**WIDER IMPLICATIONS OF THE FINDINGS:** Our data suggests that AMH may not be stable under some storage or assay conditions and this may be more pronounced with the Gen II assay. The published conversion factors between the Gen II and DSL assays appear

to be inappropriate for routine clinical practice. Further studies are urgently required to confirm our observations and to determine the cause of the apparent instability. In the meantime, caution should be exercised in the interpretation of AMH levels in the clinical setting.

**CONFLICT OF INTEREST/STUDY FUNDING:** S. Roberts is supported by the NIHR Manchester Biomedical Research Centre.

**Key words:** Anti-Müllerian hormone / Müllerian inhibitory substance / AMH / AMH Gen II ELISA / DSL Active MIS / AMH ELISA / sample stability

## Introduction

AMH in women is secreted by the granulosa cells of pre-antral and small antral follicles (Vigier *et al.*, 1984; Themmen, 2005) and circulating levels reflect the ovarian pool from which follicles can be recruited (Durlinger *et al.*, 2002; Loh and Maheshwari, 2011). The measurement of AMH has become of paramount significance in clinical practice in IVF units to assign candidates to the most suitable controlled ovarian hyperstimulation protocol and its level is used to predict poor or excessive ovarian response (Nelson *et al.*, 2007; Nardo *et al.*, 2009; Yates *et al.*, 2011). It is also of increasing importance in (i) the prediction of live birth rate in IVF cycles (La Marca *et al.*, 2011); (ii) the screening/diagnosis of polycystic ovarian syndrome (Cook *et al.*, 2002); (iii) the follow-up of women with a history of granulosa cell tumours (Lane *et al.*, 1999); (iv) the prediction of the age of onset of infertility due to the menopause (van Disseldorp *et al.*, 2008; Broer *et al.*, 2011) and finally (v) the assessment of the long-term effect of chemotherapy on fertility (Anderson, 2011).

Following the development of the first laboratory AMH assay in 1990 (Hudson *et al.*, 1990; Lee *et al.*, 1996) first generation, commercially available immunoassays were introduced by Diagnostic Systems Ltd (DSL) and Immunotech Ltd (IOT). These assays used different antibodies and standards (Nelson and La Marca, 2011) and the resulting AMH concentrations obtained using the IOT assay were found to be higher than those produced using the DSL assay by most, but not all, authors (Freour *et al.*, 2007; Taieb *et al.*, 2008; Lee *et al.*, 2011). The AMH Gen II Assay (Beckman Coulter Ltd) replaced both of these assays, using the DSL Gen I antibody with the IOT standards. AMH values obtained using this kit were predicted to correlate with, but be higher than, those using the old DSL kit (Kumar *et al.*, 2010; Nelson and La Marca, 2011). This was confirmed (Wallace *et al.*, 2011) with the AMH Gen II assay giving values ~40% higher than the DSL assay. The recommended conversion factor of 1.4 (AMH Gen II = DSL × 1.4) was also applied to the DSL reference ranges but this recommendation does not appear to have been independently validated.

It is generally accepted that serum AMH concentrations are highly reproducible within and across several menstrual cycles and therefore a single blood sampling for AMH measurement has been accepted as routine practice (Hehenkamp *et al.*, 2006; La Marca *et al.*, 2006; Tsepelidis *et al.*, 2007). However, we recently challenged this view and reported significant sample-to-sample variation in AMH levels using the DSL assay in women who had repeated measurements; 28% difference between samples taken from the same patient with a median time between sampling of 2.6 months and taking no account of menstrual cycle (Rustamov *et al.*, 2011). Although we could not explain the cause of this variability, we speculated that it might be due to true biological variation in secretion of AMH or due to post-sampling and pre-analytical instability of the specimen.

Given the widespread adoption of AMH in Clinical Units, it is critical that the sources of variability in any AMH assay are understood and quantified. This paper presents the results of clinical and laboratory studies on routine clinical samples using the new AMH Gen II assay, specifically comparing assay values with the older DSL assay, assessing between sample variability and investigating analytical and pre-analytical factors affecting AMH measurement.

## Methods

### Study population

Samples were obtained from women of 20–46 years of age attending for investigation of infertility requiring AMH assessment at the secondary (Gynaecology Department) and tertiary (Reproductive Medicine Department) care divisions of St Mary's Hospital, Manchester from 1 September 2008 to 25 October 2011. Samples which were lipaemic or haemolysed and samples not frozen within 2 h of venepuncture were excluded from the study. Anonymized samples from this pool of patients were used for stability studies after routine AMH measurements had been completed. The full data set comprised AMH results on 5868 samples from 5007 women meeting the inclusion criteria. Additionally, we identified 330 women in whom AMH had been determined using the DSL assay in one sample and the Gen II assay in another.

### Sample processing

Collection and handling of all AMH samples were conducted according to the standards set out by the manufacturers and did not vary between the different assays. Serum samples were transported immediately to the Department of Clinical Biochemistry, based in the same hospital, and separated within 2 h of venepuncture using the Modular Pre-Analytics Evo (Roche Diagnostics, Burgess Hill, West Sussex, UK). Samples were frozen in aliquots at  $-20^{\circ}\text{C}$  until analysis, normally within 1 week of receipt. The laboratory participates in the pilot National external quality assessment scheme (UKNEQAS) for AMH in Edinburgh and performance has been satisfactory.

### AMH analysis

All AMH assays were carried out strictly according to the protocols provided by the manufacturer and sample collection and storage also conformed to these recommendations. All AMH samples were analysed in duplicate and the mean of the two replicates was reported as the final result.

- (i) The DSL AMH assay: The enzymatically amplified two-site immunoassay (DSL, Active MIS/AMH ELISA; Diagnostic Systems Laboratories, Webster, TX, USA) was used for the measurement of AMH prior to 17 November 2010. The working range of the assay was up to 100 pmol/l with a minimum detection limit of 0.63 pmol/l. The intra-assay coefficient of variation (CV) ( $n = 16$ ) was 3.9% (at

10 pmol/l) and 2.9% (at 56 pmol/l). The inter-assay CV ( $n = 60$ ) was 4.7% (at 10 pmol/l) and 4.9% (at 56 pmol/l).

- (ii) The Beckman Coulter Gen II assay: After 17 November 2010, AMH was measured using the enzymatically amplified two-site immunoassay (AMH Gen II ELISA, Beckman Coulter, Inc., Brea, CA, USA). The working range of the assay is up to 150 pmol/l with a minimum detection limit of 0.57 pmol/l. The intra-assay CV ( $n = 16$ ) is 2.92% (at 18 pmol/l) and 2.03% (at 60 pmol/l). The inter-assay CV ( $n = 28$ ) is 3.57% (at 18 pmol/l) and 3.64% (at 60 pmol/l).

### Sample stability studies

- (i) Stability of AMH in serum at room temperature (RT): serum samples ( $n = 48$ ) were allowed to thaw and then left at RT for 1 week. At 0, 1, 2, 4 and 7 days, 100  $\mu$ l aliquots were removed and immediately stored at  $-80^{\circ}\text{C}$  in 2 ml screw-capped polypropylene tubes (Alpha Laboratories, Eastleigh, UK). Two freeze/thaw cycles had no effect on AMH concentration (results not shown). Samples from individual subjects were analysed for AMH on the same Gen II microtitre plate to eliminate inter-assay variability. Results were expressed as a percentage of the Day 0 value.
- (ii) Linearity of dilution: 100  $\mu$ l fresh serum ( $n = 9$ ) was added to 100  $\mu$ l AMH Gen II sample diluent, incubated for 30 min at RT and the mixture analysed using the standard Gen II assay procedure.
- (iii) Comparison between the standard assay method and an equivalent procedure: in the standard Gen II ELISA assay method, the first steps involve the addition of calibrators, controls or serum samples to microtitration wells coated with anti-AMH antibody; assay buffer is then added to each well. As a comparison, serum and assay buffer were mixed in a separate tube, incubated for 10 min at RT and then added in exactly the same volume and proportions to the microtitre plate. Thereafter, the assay was performed using the standard protocol.
- (iv) Stability of AMH during storage: fresh serum samples ( $n = 8$ ) analysed on the day of reception were compared with aliquots from the same samples that had been frozen for 5 days either in polystyrene tubes at  $-20^{\circ}\text{C}$  or polypropylene tubes at  $-80^{\circ}\text{C}$ .

### Statistical analysis

Data analysis was performed using the Stata 12 statistics package (Stata-Corp, TX, USA). Data management and analysis of clinical data were conducted by one of the researchers (O.R.) and verified independently by another member of the research team (S.R.) using different statistical software (R statistical environment). Approval for the use of the data was obtained from the Local Research Ethics Committee (UK-NHS 10/H1015/22). The age-related relationship of the DSL and Gen II assays to AMH was visualized using scatter plots and quadratic fit on a logarithmic scale (Nelson *et al.*, 2011). The age-adjusted regression analysis of paired samples was used to estimate the difference in AMH levels between the DSL and Gen II assays. A mixed effects model in  $\log(\text{AMH})$  was utilized to estimate the sample-to-sample (within-subject) coefficients of variation

of AMH levels in women who had repeated measurements, within a 1-year period from the patient's first AMH sample, adjusting for age as above. In the sample stability studies, percentage changes are expressed as mean  $\pm$  SEM. In the stability of AMH in serum at RT study, a paired  $t$ -test determined the level of significance between baseline and subsequent days.

## Results

### Population studies and variability

#### AMH concentration

Table 1 summarizes the results of AMH determinations in our population of women attending the IVF Clinic prior to the 17 November 2010 (using the DSL assay) and after that date (using the Gen II assay). A second analysis compares AMH levels in women who had AMH measured using the DSL assay for one sample and the Gen II for another. Results were consistent with lower serum levels of AMH observed when samples were analysed using the Gen II assay compared with that using the DSL assay. Figure 1 shows the correlation of AMH with age for the unselected groups. After adjustment for age, the total cohorts showed Gen II giving AMH values of 34% lower than those for DSL. Analysis restricted to patients with AMH determinations using both assays gave an age-adjusted difference of 21%.

#### AMH variability

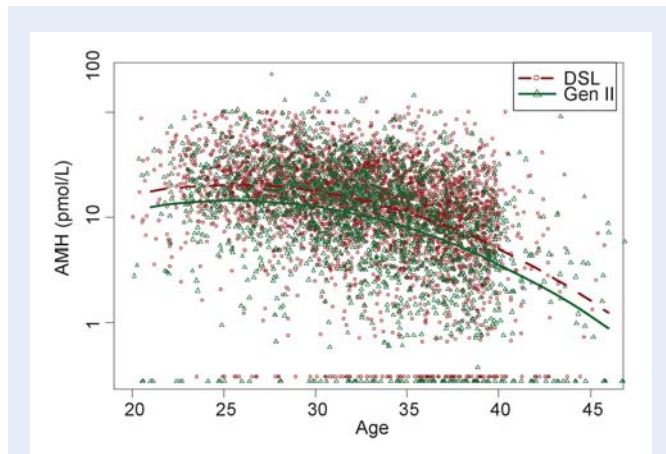
During the study period, 313 women had repeated AMH determinations ( $n = 646$  samples) using the DSL assay with 295 patients having two samples, 17 three samples and 1 five samples. The median time between samples was 5.1 months. Eighty-seven women had repeated AMH determinations using the Gen II assay ( $n = 177$  samples) with 84 women having two samples and 3 having three samples. The median interval between repeat samples was 3.2 months. Both assays exhibit high sample-to-sample variability (CV); this was 32% in the DSL assay group (our previous finding, Rustamov *et al.*, 2011, in a smaller group was 28%); variability in the Gen II assay group was much higher (59%).

### Sample stability studies

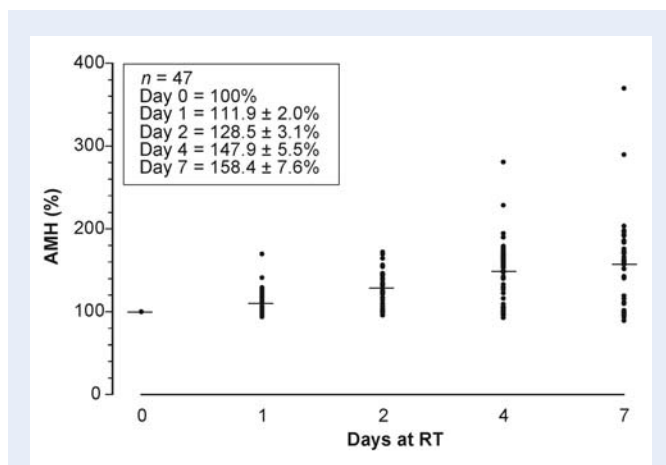
- (i) Stability of AMH in serum at RT: AMH levels in 11 of the 48 individuals remained relatively unchanged giving values within  $\pm 10\%$  of the original activity over the period of a week and one patient had an undetectable AMH at all time points. The remaining 36 serum samples had AMH values that increased progressively with time. In the 47 samples with detectable AMH, levels increased significantly ( $P < 0.001$ ) for each time interval

**Table 1** Median and inter-quartile range for the two assays in the different data sets, along with the mean difference from an age-adjusted regression model expressed as a percentage.

	DSL			Gen II			Difference (%)
	<i>n</i>	Age	AMH (pmol/l)	<i>n</i>	Age	AMH (pmol/l)	
All data	3934	33 (29, 36)	14.7 (7.8, 25.0)	1934	33 (29, 36)	11.2 (4.5, 21.6)	-33.5 (-39.5 to -27.5)
Paired samples	330	32 (29, 36)	14.9 (7.4, 24.7)	330	34 (30, 37)	11.0 (5.6, 20.9)	-21.4 (-36.2 to -6.4)



**Figure 1** Unselected AMH values from DSL (circles) and Gen II (triangles) assays as a function of age. Lines show the regression fits of  $\log(\text{AMH})$  against a quadratic function of age; solid lines Gen II, broken lines DSL.



**Figure 2** Stability of AMH in serum at RT. Results at each time interval are expressed as a percentage of the patient's AMH concentration at Day 0. Means  $\pm$  SEM are indicated.

compared with baseline, the increase at Day 7 being  $158.4 \pm 7.6\%$  (Fig. 2).

- (ii) Linearity of dilution: in a group of nine anonymized samples, proportionality with 2-fold sample dilution did not hold and, on average, there was a  $57.4 \pm 12.3\%$  increase in the apparent AMH concentration on dilution, compared with neat sample (see Table II). Two samples which gave the highest increases were diluted further. It was apparent that, after the anomalous doubling of AMH concentration on initial two-fold dilution, subsequent dilutions gave a much more proportional result (see Table III). Linearity of dilution was maintained only in samples that showed no initial increase on two-fold dilution.
- (iii) Comparison between the standard assay method and an equivalent procedure: serum samples that had been pre-mixed with buffer prior to addition gave, on average,  $71.8 \pm 4.8\%$  higher

**Table II** Proportionality with two-fold dilution of serum.

Sample no.	AMH (pmol/l)		Recovery %
	Neat serum	$\times 2$ dilution	
1	110.5	229.4	207.6
2	49.41	99.00	200.4
3	4.15	4.83	116.4
4	9.23	11.22	121.6
5	28.01	30.66	109.1
6	3.62	6.28	173.4
7	27.39	39.62	144.7
8	5.53	10.34	187.0
9	18.49	28.92	156.4

**Table III** Linearity with multiple dilution of serum.

Sample no.	Dilution	AMH (pmol/l)		Recovery (%)
		Measured	Expected	
1	$\times 1$	110.5	110.5	100
	$\times 2$	114.7	55.25	207.6
	$\times 4$	55.32	27.63	200.2
	$\times 7$	30.72	15.79	194.6
	$\times 10$	21.45	11.05	194.1
2	$\times 1$	49.41	49.41	100
	$\times 2$	49.50	24.71	200.3
	$\times 4$	22.86	12.35	185.1
	$\times 7$	12.28	7.06	173.9
	$\times 10$	8.57	4.94	173.5

readings than those added sequentially using the standard procedure (see Table IV).

- (iv) Stability of AMH during storage: AMH levels in samples stored at  $-20^\circ\text{C}$  showed an average increase of  $22.5 \pm 11.1\%$  over 5 days compared with fresh values, while those samples stored at  $-80^\circ\text{C}$  showed no change ( $1.8 \pm 3.1\%$ ) (see Table V).

## Discussion

This publication arose from two, initially separate, pieces of work in the Clinical IVF Unit at St Mary's Hospital and in the Specialist Assay Laboratory at Central Manchester Foundation Trust. The IVF Unit had become concerned with their observed increase in variation in AMH values and consequently with the reliability of their AMH-tailored treatment guidance. The Laboratory wished to establish whether the practice of sending samples in the post (which has been adopted by many laboratories, rather than frozen as specified by Beckman) was viable. It soon became clear that these anomalies

**Table IV** Comparison between the standard ELISA method and a modified procedure.

Sample no.	AMH (pmol/l)		B/A (%)
	A	B	
1	14.66	22.84	155.8
2	8.39	16.42	195.7
3	31.51	64.46	204.6
4	12.44	20.14	161.9
5	13.93	22.76	163.4
6	7.01	12.46	177.7
7	7.78	13.58	174.6
8	16.93	32.98	194.8
9	9.55	17.93	187.7
10	28.49	54.37	190.8
11	13.65	20.62	151.1
12	17.73	28.68	161.7
13	14.68	24.29	165.5
14	14.99	21.15	141.1
15	2.49	3.57	143.4
16	12.84	22.89	178.3

A = 20  $\mu$ l of serum added directly to the plate followed by 100  $\mu$ l assay buffer.  
 B = 60  $\mu$ l of serum + 300  $\mu$ l assay buffer mixed and incubated at RT for 10 min;  
 120  $\mu$ l of mixture added to the plate.

**Table V** Stability of AMH in serum on storage for 5 days.

Sample no.	AMH (pmol/l)		
	Fresh	-20°C, PS	-80°C, PP
1	12.41	15.51	13.12
2	42.17	75.42	45.08
3	11.93	17.12	12.39
4	10.42	12.82	12.28
5	9.56	9.05	8.79
6	19.02	26.01	18.84
7	24.02	20.16	23.62
8	1.45	1.37	1.32

PS, polystyrene LP4 tube; PP, polypropylene 2 ml tube.

observed in clinical practice might be explained by a marked degree of sample instability seen in the Laboratory which had not previously been reported and which may, or may not, have been an issue with previous AMH assays.

The data contained in this paper represent the largest retrospective study on the variability of the DSL assay and the first study on the variability of the Gen II assay. Early studies reported insignificant variation between repeated AMH measurements, suggesting that a single AMH measurement may be sufficient in the assessment of ovarian reserve (La Marca *et al.*, 2006; Tsepedis *et al.*, 2007). However, these

recommendations have been challenged by a number of groups (Lahlou *et al.*, 2006; Wunder *et al.*, 2008; Rustamov *et al.*, 2011). The current study in a large cohort of patients has demonstrated substantial sample-to-sample variation in AMH levels using the DSL assay and an even larger variability using the Gen II assay. We suggest that this variability may be due to sample instability related to specimen processing given that (i) AMH is produced non-cyclically and true biological variation is believed to be small (Fanchin *et al.*, 2005; van Disseldorp *et al.*, 2010) and (ii) the intra-and inter-assay variation in our laboratory for both the DSL and Gen II assays is small (<5.0%) suggesting that the observed variation is not due to poor analytical technique.

The population data presented in this paper also suggests that, in routine clinical use, the Gen II assay provides AMH results which are 20–40% lower than those measured using the DSL assay. This is in contrast to validation studies for the Gen II assay which showed that this assay gave AMH values ~40% higher than those found with the DSL assay (Kumar *et al.*, 2010; Preissner *et al.*, 2010; Wallace *et al.*, 2011).

All samples in this retrospective study were subject to the same handling procedures, and were analyzed by the same laboratory; the two populations were comparable, with the same local referral criteria for investigation of infertility. We are unaware of any other alterations in practice, which might produce such a large effect on AMH, but we cannot rule out the possibility of other changes in the population being assayed that were contemporaneous with the assay change. However, any such change would have had to produce a 50% decrease in observed AMH levels to explain our findings. We did note a weak trend towards decreasing AMH over calendar time; assuming a linear trend in the analysis implies that AMH values might be 12% (2–22%) lower when the Gen II assay was being used compared with the Gen I assay. This suggests that the age-adjusted analysis of repeat samples on individuals, showing a 21% decrease in AMH with the Gen II assay, is currently the best estimate of the assay difference.

This is the first study to compare AMH assays in a routine clinical setting in a large group of subjects, and as such is likely to reflect the true nature of the relationship between AMH measured by two different ELISA kits and avoids some of the issues in other published studies. Previous laboratory studies have compared AMH assays in aliquots from the same sample which only provides data on the within-sample relationship between the two assays (Kumar *et al.*, 2010; Preissner *et al.*, 2010; Wallace *et al.*, 2011). Although it is difficult to give a definitive explanation for the discrepancy between the previously published studies (on within-sample relationships) and this study (on between-sample relationships), we suggest that it may be due to degradation of the specimen in one (or both) of the assays. If AMH in serum is unstable under certain storage and handling conditions, this might result in differing values being generated because of differential sensitivity of the two assays to degradation products. Unfortunately, we cannot suggest which step of sample handling might have caused this discrepancy, since the published studies did not provide detailed information.

The present study used samples which were frozen very soon after phlebotomy and analysed shortly thereafter, hopefully minimizing storage effects. The most striking change followed incubation over a period of 7 days at RT; this showed a substantial increase in AMH

levels, rather than the expected decline. Previously, *Kumar et al. (2010)* had shown that the average variation between fresh serum samples and those stored for 7 days was found to be ~4% at 2–8°C and <1% at –20°C but presented no data on RT stability. *Zhao et al. (2007)* reported that AMH values were likely to differ by <20% in samples incubated at RT for 2 days compared with those frozen immediately.

Several supplementary experiments were performed in order to investigate this observed increase in AMH when samples were incubated at RT. These included (i) addition of the detergent Tween-20 to assay buffer to disclose potential antibody-binding sites on the AMH molecule and (ii) the removal of heterophilic antibodies from serum using PEG precipitation or heterophilic blocking tubes. None of these approaches affected AMH levels significantly (results not shown).

Examination of the data presented here shows that, in some samples, AMH levels tend towards twice those expected, while results greater than that only occur in two outliers found in Fig. 2. The AMH molecule is made up of two identical 72 kDa monomers which are covalently bound (*Wilson et al., 1993; di Clemente et al., 2010*). During cytoplasmic transit, each monomer is cleaved to generate 110-kDa N-terminal and 25-kDa C-terminal homodimers, which remain associated in a non-covalent complex. The C-terminal homodimer binds to the receptor but, in contrast to other TGF- $\beta$  superfamily members, AMH is thought to require the N-terminal domain to potentiate this binding to achieve full bioactivity of the C-terminal domain. After activation of the receptor, the N-terminal homodimer is released (*Wilson et al., 1993*). One possible explanation for our findings is that the N- and C-terminal homodimers dissociate gradually under certain storage conditions and that either the two resulting N- and C-terminal components bind to the ELISA plate or a second binding site on the antigen is exposed by the dissociation, effectively doubling the concentration of AMH. It has been shown (*di Clemente et al., 2010*) that no dissociation occurs once the complex is bound to immobilized AMH antibodies. The observation that, in some of our samples, there was no change after 1 week at RT might be explained by the supposition that in those samples AMH is already fully dissociated. A mixture of dissociated and complex forms in the same sample would, therefore, account for the observed recoveries between 100 and 200% in the experiments presented in this paper. Rapid sample processing and storage of the resulting serum in a different tube type at –80°C might slow down this breakdown process.

The change in ionic strength or pH that occurs on dilution also seems to have the same effect in increasing apparent AMH levels and again may be due to dissociation or exposure of a second binding site. Our results contradict those reported by *Kumar et al. (2010)* who showed that serum samples in the range of 36–93 pmol/l of AMH when diluted in Gen II sample diluent showed linear results across the dynamic range of the assay with average recoveries on dilution close to 100%. This might be explained if Kumar's samples were already dissociated before dilution. Linearity is one of the cornerstones of assay validation and it is essential that a proportional response is obtained on dilution of sample, but our results do not seem to support this.

These findings have significant clinical relevance, given the widespread use of AMH as the primary tool for assessment of ovarian reserve and as a marker for tailoring the dose of gonadotrophins in

cycles of IVF/ICSI. As no guideline studies have been published using the new Gen II assay, some ART centres have adopted modified treatment 'cut-off levels' for ovarian stimulation programmes based on the old DSL assay-based 'cut-off levels' multiplied by a conversion factor of 1.4 (*Nelson et al., 2007; Nelson et al., 2009; Wallace et al., 2011*). The data presented in this paper suggest that this approach could result in patients being allocated to the wrong ovarian reserve group. Poor performance of the Gen II assay in terms of sample-to-sample variability (up to 59%) could also lead to unreliable allocation to treatment protocols. It is a matter of some urgency, therefore, that any possible anomalies in the estimation of AMH using the Gen II assay be thoroughly investigated and that this work should be repeated in other centres.

## Authors' roles

O.R. led on clinical aspects of this study with responsibility for collation of the clinical database and the analysis of the clinical data. OR prepared the first draft of the clinical work and was involved in preparation of the whole paper and submission of the final manuscript. C.F. and L.G.N. contributed to clinical data analysis, discussion, draft preparation and approval of the final manuscript. M.K. was involved in clinical data collation and approval of the final draft. PWP was the laboratory lead responsible for all of the laboratory-based experiments and for the routine analysis of clinical samples. P.W.P. prepared the first draft of the laboratory work and was involved in the preparation of the whole paper and submission of the final manuscript. A.S. suggested the sample stability studies and was involved in discussion, draft preparation and approval of the final manuscript. A.P.Y. was involved in some of the routine clinical analyses, and discussion and progression of drafts to approval of the final manuscript. S.A.R. was involved in clinical study design, oversaw the statistical analysis and contributed to discussion and progression of drafts through to approval of the final manuscript. O.R. and P.W.P. should be considered as joint first authors.

## Funding

Dr S.R. is supported by the NIHR Manchester Biomedical Research Centre.

## Conflict of interest

None declared.

## References

- Anderson RA. Pretreatment serum anti-Mullerian hormone predicts long-term ovarian function and bone mass after chemotherapy for early breast cancer. *J Clin Endocrinol Metab* 2011;**96**:1336–1343.
- Broer SL, Eijkemans MJ, Scheffer GJ, van Rooij IA, de Vet A, Themmen AP, Laven JS, de Jong FH, Te Velde ER, Fauser BC et al. Anti-mullerian hormone predicts menopause: a long-term follow-up study in normoovulatory women. *J Clin Endocrinol Metab* 2011;**96**:2532–2539.
- Cook CL, Siow Y, Brenner AG, Fallat ME. Relationship between serum Mullerian inhibiting substance and other reproductive hormones in

- untreated women in PCOS and normal women. *Fertil Steril* 2002; **77**:141–146.
- di Clemente N, Jamin SP, Lugovskoy A, Carmillo P, Ehrenfels C, Picard JY, Whitty A, Josso N, Pepinsky RB, Cate RL. Processing of anti-mullerian hormone regulates receptor activation by a mechanism distinct from TGF-beta. *Mol Endocrinol* 2010; **24**:2193–2206.
- Durlinger ALL, Visser JA, Themmen APN. Regulation of ovarian function: the role of anti-Mullerian hormone. *Reproduction* 2002; **124**:601–609.
- Fanchin R, Taieb J, Mendez Lozano DH, Ducot B, Frydman R, Bouyer J. High reproducibility of serum anti-Mullerian hormone measurements suggests a multi-staged follicular secretion and strengthens its role in the assessment of ovarian follicular status. *Hum Reprod* 2005; **20**:923–927.
- Freour T, Mirallie S, Bach-Ngohou K, Denis M, Barriere P, Masson D. Measurement of serum anti-Mullerian hormone by Beckman Coulter ELISA and DSL ELISA: comparison and relevance in assisted reproduction technology (ART). *Clin Chim Acta* 2007; **375**:162–164.
- Hehenkamp WJ, Looman CW, Themmen AP, de Jong FH, Te Velde ER, Broekmans FJ. Anti-Mullerian hormone levels in the spontaneous menstrual cycle do not show substantial fluctuation. *J Clin Endocrinol Metab* 2006; **91**:4057–4063.
- Hudson PL, Dugas I, Donahoe PK, Cate RL, Epstein J, Pepinsky RB, MacLaughlin DT. An immunoassay to detect human mullerian inhibiting substance in males and females during normal development. *J Clin Endocrinol Metab* 1990; **70**:16–22.
- Kumar A, Kalra B, Patel A, McDavid L, Roudebush WE. Development of a second generation anti-Müllerian hormone (AMH) ELISA. *J Immunol Methods* 2010; **362**:51–59.
- La Marca A, Stabile G, Artesio AC, Volpe A. Serum anti-Mullerian hormone throughout the human menstrual cycle. *Hum Reprod* 2006; **21**:3103–3107.
- La Marca A, Nelson SM, Sighinolfi G, Manno M, Baraldi E, Roli L, Xella S, Marsella T, Tagliasacchi D, D'Amico R et al. Anti-Mullerian hormone-based prediction model for a live birth in assisted reproduction. *Reprod Biomed Online* 2011; **22**:341–349.
- Lahlou N, Chabbert-Buffet E, Gainer E, Roger M. Diphasic pattern of anti-Mullerian hormone (AMH) in ovulatory cycles as evidenced by means of ultra-sensitive assay: new insights into ovarian function. *Fertil Steril* 2006; **86**:S11.
- Lane AH, Lee MM, Fuller AF Jr, Kehas DJ, Donahoe PK, MacLaughlin DT. Diagnostic utility of Mullerian inhibiting substance determination in patients with primary and recurrent granulosa cell tumors. *Gynecol Oncol* 1999; **73**:51–55.
- Lee MM, Donahoe PK, Hasegawa T, Silverman B, Crist GB, Best S, Hasegawa Y, Noto RA, Schoenfeld D, MacLaughlin DT. Mullerian inhibiting substance in humans: normal levels from infancy to adulthood. *J Clin Endocrinol Metab* 1996; **81**:571–576.
- Lee JR, Kim SH, Jee BC, Suh CS, Kim KC, Moon SY. Antimullerian hormone as a predictor of controlled ovarian hyperstimulation outcome: comparison of two commercial immunoassay kits. *Fertil Steril* 2011; **95**:2602–2604.
- Loh JS, Maheshwari A. Anti-Mullerian hormone—is it a crystal ball for predicting ovarian ageing? *Hum Reprod* 2011; **26**:2925–2932.
- Nardo LG, Gelbaya TA, Wilkinson H, Roberts SA, Yates A, Pemberton P, Laing I. Circulating basal anti-Müllerian hormone levels as predictor of ovarian response in women undergoing ovarian stimulation for *in vitro* fertilization. *Fertil Steril* 2009; **92**:1586–1593.
- 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-Mullerian hormone-based approach to controlled ovarian stimulation for assisted conception. *Hum Reprod* 2009; **24**:867–875.
- Nelson SM, Messow MC, Wallace AM, Fleming R, McConnachie A. Nomogram for the decline in serum antimullerian hormone: a population study of 9,601 infertility patients. *Fertil Steril* 2011; **95**:736–741.
- Preissner CM, Morbeck DE, Gada RP, Grebe SK. Validation of a second generation assay for anti-Mullerian hormone. *Clin Chem* 2010; **56**:A54.
- Rustamov O, Pemberton PV, Roberts SA, Smith A, Yates AP, Patchava SD, Nardo LG. The reproducibility of serum anti-Müllerian hormone in subfertile women: within and between patient variability. *Fertil Steril* 2011; **95**:1185–1187.
- Taieb J, Coussieu C, Guibourdenche J, Picard JY, di Clemente N. Two immunoassays for antimullerian hormone measurement: analytical and clinical performances. *Ann Biol Clin (Paris)* 2008; **66**:537–547.
- Themmen AP. Anti-Mullerian hormone: its role in follicular growth initiation and survival and as an ovarian reserve marker. *J Natl Cancer Inst Monogr* 2005; **(34)**:18–21.
- Tsepelidis S, Devreker F, Demeestere I, Flahaut A, Gervy C, Englert Y. Stable serum levels of anti-Mullerian hormone during the menstrual cycle: a prospective study in normo-ovulatory women. *Hum Reprod* 2007; **22**:1837–1840.
- van Disseldorp J, Faddy MJ, Themmen AP, de Jong FH, Peeters PH, van der Schouw YT, Broekmans FJ. Relationship of serum antimullerian hormone concentration to age at menopause. *J Clin Endocrinol Metab* 2008; **93**:2129–2134.
- van Disseldorp J, Lambalk CB, Kwee J, Looman CWN, Eijkemans MJC, Fauser BC, Broekmans FJ. Comparison of inter- and intra-cycle variability of anti-Mullerian hormone and antral follicle counts. *Hum Reprod* 2010; **25**:221–227.
- Vigier B, Picard JY, Tran D, Legeai L, Josso N. Production of anti-Mullerian hormone: another homology between Sertoli and granulosa cells. *Endocrinology* 1984; **114**:1315–1320.
- Wallace AM, Faye SA, Fleming R, Nelson SM. A multicentre evaluation of the new Beckman Coulter anti-Mullerian hormone immunoassay (AMH Gen II). *Ann Clin Biochem* 2011; **48**:370–373.
- Wilson CA, Di Clemente N, Ehrenfels C, Pepinsky RB, Josso N, Vigier B, Cate RL. Müllerian inhibiting substance requires its N-terminal domain for maintenance of biological activity, a novel finding within the transforming growth-factor-beta superfamily. *Mol Endocrinol* 1993; **7**:247–257.
- Wunder DM, Yared M, Kretschmer R, Birkhauser MH. Statistically significant changes of anti-Mullerian hormone and inhibin levels during the physiologic menstrual cycle in reproductive age women. *Fertil Steril* 2008; **89**:927–933.
- Yates AP, Rustamov O, Roberts SA, Lim HY, Pemberton PV, Smith A, Nardo LG. Anti-Mullerian hormone-tailored stimulation protocols improve outcomes whilst reducing adverse effects and costs of IVF. *Hum Reprod* 2011; **26**:2353–2362.
- Zhao J, Ng SY, Rivnay B, Leader BS. Serum Anti-Mullerian hormone (AMH) measurement inter-assay agreement and temperature stability. *Fertil Steril* 2007; **88**:S1.