Anti-Müllerian hormone: clairvoyance or crystal clear?

Scott M. Nelson1,*, Richard A. Anderson2, Frank J. Broekmans3, Nick Raine-Fenning4, Richard Fleming1,5, and Antonio La Marca6

1School of Medicine, University of Glasgow, Glasgow, UK 2MRC Centre for Reproductive Health, Queens Medical Research Institute, University of Edinburgh, Edinburgh, UK 3Department of Reproductive Medicine and Gynecology, Division of Woman and Baby, University Medical Center Utrecht, GA Utrecht, The Netherlands 4Division of Human Development, School of Clinical Sciences, University of Nottingham, Nottingham, UK 5Glasgow Centre for Reproductive Medicine Limited, Glasgow, UK 6Mother-Infant Department, Institute of Obstetrics and Gynecology, University of Modena and Reggio Emilia, 41100 Modena, Italy

*Correspondence address. Tel: +44-141-211-2329; Fax: +44-141-552-0873; E-mail: scott.nelson@glasgow.ac.uk

Submitted on October 9, 2011; resubmitted on November 9, 2011; accepted on November 16, 2011

ABSTRACT: The clinical use of anti-Müllerian hormone (AMH) has increased exponentially due to its unique relationship with the ovarian reserve and ability to predict ovarian response, facilitate pretreatment counselling and individualize treatment strategies to minimize the risk of ovarian hyperstimulation syndrome. There is now a rapidly increasing literature examining additional possibilities for AMH, all of which suggest that its reach extends far beyond the assisted conception clinic. The recognition that it is a significantly more accurate and reliable measure of ovarian reserve than the antral follicle count or FSH has led to its adoption by physicians to counsel women on their reproductive lifespan, the impact of gonadotoxic chemotherapy, radiotherapy and surgery on the ovarian reserve and allow polycystic ovarian syndrome to be diagnosed by primary care physicians. We propose that there is an adequate literature base to embrace this technology while continuing to develop and refine how AMH can optimize patient care.

Key words: Mülllerian inhibitory substance / ovarian reserve / ultrasound

Introduction

The debate article by Loh and Maheshwari presents a perspective on the potential utility of anti-Müllerian hormone (AMH) in women’s health (Loh and Maheshwari, 2011). While acknowledging the ability of AMH as an accurate predictor of the ovarian response to exogenous gonadotrophins, the authors argue that AMH should not be adopted by reproductive endocrinologists, while highlighting their own preference for the antral follicle count (AFC). They also briefly examine the potential alternative uses for measurement of circulating AMH concentrations, and conclude that at present there is insufficient evidence for more widespread use of AMH. We do not consider that Loh and Maheshwari provide a thorough analysis of the use (and limitations) of AMH in contemporary reproductive medicine and propose a contrary position in this debate article.

AMH—crystal clear clarity today

So can AMH justify its inclusion in the clinician’s armamentarium? Introduction of a novel hormone/marker assay into routine clinical practice requires more than scientific validation in the research context. Accredited clinical biochemistry laboratories will already have the prerequisite infrastructure required to provide accurate reporting of blood tests. For AMH, the independent laboratories of Nelson and Fleming report an inter- and intra-assay coefficient of variation <3%, which is an improvement on most other endocrine measures. AMH is now part of the Europe-wide National External Quality Assessment Service scheme which promotes inter-laboratory standardization. Pre-existing issues regarding different AMH assays have now been resolved with Beckman Coulter purchasing the patents for all previous versions and establishing the AMH Generation II assay, although this generates a monopoly (Nelson and La Marca, 2011). This assay retains the cross-species specificity of the DSL (Diagnostic Systems Laboratories) assay and is calibrated to the Immunotech standards. Consequently, the values generated are similar to the original Immunotech assay and 40% higher than the previous DSL version (Wallace et al., 2011). Although licensing agreements for the AMH assay have been established with other automated platform manufacturers, they are contractually required to have similar standardization to allow relative conformity in the reporting of results. Consequently, the days of confusion over interpretation of AMH results are coming to an end (Nelson and La Marca, 2011).

It is with the ongoing resolution of these issues that we have seen an exponential growth in the use of AMH by a range of doctors including primary care physicians, general gynaecologists and reproductive endocrinologists. Blood tests clearly have marked advantages over...
ultrasound for primary care physicians but many in reproductive medi-
cine work in an environment where a transvaginal ultrasound examin-
ation is a routine part of patient assessment; thus, are there still advantages there?

**AMH stability**

Relative to the AFC, AMH has consistently been shown to have signifi-
cantly lower intra- and inter-cycle variability (Hehenkamp et al., 2006; La Marca et al., 2006; Tsepelidis et al., 2007; van Disseldorp et al., 2009). This unique characteristic among female reproductive hor-
mones allows AMH to be performed at anytime within the menstrual cycle, while still exhibiting strong associations with the outcomes of interest (Nelson et al., 2009). While AMH does show some intra-individual variability, this is within classifications of ovarian re-
response/reserve, and this property alone has major implications for clinical workload planning and patient convenience (Sowers et al., 2010).

**AMH and ovarian ageing**

As biomarkers of ovarian ageing, both AFC and AMH demonstrate strong relationships with the stereological determined estimates of the non-growing follicle (NGF) pool (Hansen et al., 2010). Several investigators have eloquently modelled the age-related decline in the NGF pool (Faddy et al., 1992; Faddy and Gosden, 1996; Faddy, 2000; Hansen et al., 2008), with recent models incorporating foetal data through into late adult life (Wallace and Kelsey, 2010). This cross-sectional approach has now been replicated for AMH from healthy foetuses, children and women, with the decline in adult life described by a quadratic curve (Kelsey et al., 2011). This is in keeping with our previous assessment of AMH in 9601 women and subsequent external validation in a further 15 834 women (Nelson et al., 2011a,b). There is no data set of comparable size for any other reproductive biomarker, including AFC. The robust statistical approaches used in these studies reveal a markedly similar anticipated decline in AMH, which is consist-
ent with all previous smaller AMH nomograms (Mulders et al., 2004; van Rooij et al., 2005; Nardo et al., 2007; La Marca et al., 2010a,b,c; Almog et al., 2011; Shebl et al., 2011). Although Loh and Maheshwari are surprised at the high variability in AMH, we believe that this is to be expected given the wide variability in ovarian reserve and conse-
quently the age at menopause—which has a distribution spanning 20 years (Broekmans et al., 2009; Wallace and Kelsey, 2010). Exam-
ination of rates of follicular recruitment through life reveals a difference between women of nearly 100-fold (Wallace and Kelsey, 2010). AMH distributions thus reflect the true biological variation in ovarian follicle number and rate of growth activation. As ovarian function is not crit-
ical to life in the same way as is, for example, the extracellular concen-
tration of potassium, it is not subject to the same rigorous limitations. Even accounting for this variability, the substantial size of these cross-
sectional cohorts ensures that longitudinal studies will show broadly similar rates of decline at a population level.

Although continuing emphasis has been placed on the necessity for longitudinal cohorts of healthy women before being able to interpret the clinical significance of a given individual AMH, the additional value this will add over large-scale cross-sectional population studies is poten-
tially minimal, at least for women at ages on the AMH down-slope, i.e. over 30. This is in marked contrast to women who have an insult to their ovarian reserve such as chemotherapy or ovarian surgery, where longitudinal assessment is entirely appropriate as their NGF decline may deviate substantially from the normal population (Bath et al., 2003; Anderson et al., 2006; Lie Fong et al., 2008; Chang et al., 2010; Iwase et al., 2010; Tsolakidis et al., 2010; Anderson and Cameron, 2011). Consistent with this a pretreatment, AMH has recently been shown to be a better predictor of post-chemotherapy ovari-
an activity than age in a long-term follow-up study (Anderson and Cameron, 2011). Similarly, uncertainties over changes in AMH in the first two decades of life mean that interpretation of AMH in puberty, adolescence and early adult peak AMH attainment requires considerable additional data (Hagen et al., 2010; Kelsey et al., 2011).

For other fields of medicine such as cardiovascular disease where biomarkers including lipids have a clearly established role, clinicians do not wait for repeated measurements to assess rates of increase or decrease before acting. Instead, they combine the individual result with the clinical history and the potential disease risk. Given our ability to interpret AMH in an age-specific manner, it seems approp-
riate to develop management plans centred on patients’ desire for fertility based on their individual AMH, taking into account their family history of age at menopause, personal circumstances and our ability to potentially prolong reproductive lifespan through fertility preservation techniques (Anderson and Wallace, 2011).

**AMH and live birth prediction**

The mantra that AMH does not predict pregnancy (‘quantity not quality’) has been frequently repeated (Smeenk et al., 2007; Broer et al., 2009). However, the flaws in these analyses are now being elu-
cidated. It is widely accepted that oocyte yield independent of age pre-
dicts live birth in IVF cycles—as demonstrated by an analysis of 400 135 IVF cycles, with a linear relationship between oocyte yield and live birth up to 15 oocytes (Sunkara et al., 2011). AMH is in itself strongly associated with oocyte yield (La Marca et al., 2010a,b,c), and in IVF cycles using protocols designed to achieve maximal follicular recruitment, correlation coefficients >0.7 are frequently reported (Nelson et al., 2007). It would therefore seem counter-intuitive that AMH would not be associated with live birth through its relationship with oocyte yield. Fortunately, a number of studies have now specific-
ally examined this issue. All demonstrate that age and AMH are inde-
pendently associated with live birth (Nelson et al., 2007; Gleicher et al., 2010; La Marca et al., 2010a,b,c; Majumder et al., 2010), with AMH achieving this through its primary relationship with oocyte yield (Nelson et al., 2007; Majumder et al., 2010). This is not to say that AMH can predict oocyte or embryo quality, just that there will potentially be more oocytes and thereby embryos to develop and sub-
sequently select for transfer. We previously developed and subse-
quently externally validated a nomogram for stratifying the probability of live birth based on AMH and age in assisted conception, although confidence intervals (CIs) were wide, in keeping with the small numbers assessed (La Marca et al., 2010a,b,c; A. Khader et al., submitted). Consequently, the reality is that age clearly remains the primary determinant of the probability of a live birth after assisted conception, but for any given age, women with a higher ovarian reserve and thereby a higher circulating AMH will do better than their counterparts with a lower AMH. To put it another
way, as women get older, the requirement for embryo selection becomes more important, and therefore restricted ovarian reserve is a handicap. This is important and straightforward information that can be imparted to our patients. Although clinicians may be tempted to exclude patients with a low AMH from IVF programmes on cost-effective grounds (Yates et al., 2011), this will inevitably deny treatment to patients with a low/moderate chance of success of a live birth (Veghofer et al., 2011).

Moving forwards, the next step should be to assess whether inclusion of AMH improves discrimination and calibration of established prediction models for natural conception and IVF-related live birth (van der Steeg et al., 2007; Leushuis et al., 2009; Nelson and Lawlor, 2011). We anticipate that the combination of a baseline phenotype with biomarkers is likely to offer the best estimates of success for IVF, and it will be this approach which will be used for health economic modelling of provision of state-funded IVF. For natural conception, other weaker biomarkers of the NGF pool including FSH, ovarian volume and length of the menstrual cycle, have all shown associations with fecundity even when adjusted for age (Small et al., 2006; van der Steeg et al., 2007). We anticipate that sizeable cohorts will therefore also be able to confirm initial associations between AMH and the chance of natural conception (Steiner et al., 2011), although the effect size may be weak and the relationship complex, as ovulation is likely to be influenced at both ends of the AMH spectrum.

**AFC—shades of grey in a dark room**

There is no doubt that the AFC, the primary alternative to AMH, provides a direct and immediate index of the ovarian response to stimulation (Broekmans et al., 2006). However, over the last decade it has been clearly established that AFC suffers from significant operator variability (Scheffer et al., 2002; Hansen et al., 2003; Lujan et al., 2009; van Disseldorp et al., 2009; Broekmans et al., 2010). This is apparent even in studies which have used a single operator to assess intra- and inter-cycle variability (Fanchin et al., 2005; van Disseldorp et al., 2009). Recent ultrasonographic technical developments including automated three-dimensional ultrasound measurements have the potential to reduce but not to eliminate this variability (Raine-Fenning et al., 2008, 2009; Deb et al., 2009; Salama et al., 2010). This technology remains restricted to a few select centres, however, with software optimization still required to overcome the necessity for time-consuming off-line data manipulation to ensure an accurate identification of each and every individual follicle and an accurate estimate of ovarian reserve and response, undermining the attractions of speed and immediacy of AFC measurement (Jayaprakasan et al., 2007).

Given that most clinics use real-time 2D ultrasound and typically employ several operators, all of whom will have to work under the various time restraints that clinical practice is inevitably associated with, there is an inherent degree of variability in AFCs. A significant investment in training, ongoing quality control and technology are all required to fully realize the potential for AFC and the lack of these has limited its value as a clinical tool in multi-centre trials.

It is clear that the AFC has a relationship with ovarian reserve and response (Broekmans et al., 2006; Jayaprakasan et al., 2010; Broer et al., 2011a,b). However, the clinical utility of the AFC has been limited, in part due to recognition of the considerable variability in clinical definitions and technical methodology. In an attempt to overcome these issues, we recently proposed a standardized AFC measurement (Broekmans et al., 2010). To minimize confounding, combined with the known difficulties of measuring the AFC in women with previous ovarian surgery or ovarian cysts, we suggested excluding all women with irregular menstrual cycles or co-existing pathology and limiting transvaginal examination to Days 2–5 of the menstrual cycle (Broekmans et al., 2010). Although a step towards standardization of the AFC, by excluding women with oligomenorrhoea or ovarian pathology and performing it in the early follicular phase, this creates additional workload while limiting the generalizability of the AFC. Accurate clinical interpretation is also limited due to the lack of large-cale age-specific AFC normograms (La Marca et al., 2011). With respect to prediction of live birth, AFC has recently been shown to be positively associated with live birth, independent of oocyte yield and age, in a large single-centre study ($n = 2092$ patients, 4308 cycles; adjusted odds ratio of 1.23 (95% CI 1.10–1.37) for AFC stratav) (Holte et al., 2011). This striking finding brings into question the overall premise that the AFC, and thereby also potentially AMH, reflects just quantity not quality, and confirmation by others is eagerly awaited.

With respect to polycystic ovary syndrome (PCOS) diagnosis, it is now widely appreciated that the original AFC criterion is unhelpful (Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group, 2004), as with modern ultrasound technology, the threshold recommended identifies a morphological feature that is present in up to 35% of adolescent and adult women (Duijkers and Klipping, 2010; Johnstone et al., 2010; Kristensen et al., 2010; Raine-Fenning, 2011). Recent reassessment, by the group who proposed the original AFC criterion (Jonard et al., 2003), recognises this limitation and suggests that an AFC of 19 may be a better threshold for diagnosis (Dewailly et al., 2011). This requires debate and validation by other groups, but given the human aspects of the AFC reflected in the observer variability, any novel biomarker assessed against this criterion will inevitably perform suboptimally.

**AMH as a long-term predictor**

Loh and Maheshwari briefly discuss the limitations of AMH in other clinical contexts, including the diagnosis of PCOS, in oncology/fertility preservation and in the prediction of the menopause. We agree that there is a gulf between the demonstration that AMH has diagnostic or predictive value in these contexts and evaluation of the true usefulness of AMH in clinical practice. Full evaluation of the current standing of these potential indications is beyond the remit of this article. However, it is clear that there are strong data for the value of AMH in all these indications (van Rooij et al., 2004; Sowers et al., 2008; van Disseldorp et al., 2008; Anderson and Cameron, 2011; Broer et al., 2011a,b; Dewailly et al., 2011), and more besides: the robust tradition of critical evaluation that has emerged in reproductive medicine will ensure that the value and limitations of AMH will be rigorously evaluated. This particularly applies to our inability to interpret the long-term fertility implications of an AMH measurement in young women, and we agree with Loh and Maheshwari’s caution regarding the widespread uptake of AMH measurement without clearly defined indications. It is however difficult to imagine that AMH will
be of less diagnostic value in PCOS than the current biochemical tests to which we subject our patients.

**Conclusions**

We agree with the proposition that AMH does not meet World Health Organization criteria for screening for a disease state. However, its role as a clinical test is clear. AMH should not be assessed against the criteria of disease, but as a test for ovarian ageing. We argue that AMH meets the specific test criterion ‘there should be a suitable and acceptable screening test or examination’. Whether society will decide to place sufficient worth on ovarian ageing, and its well-established natural history, implications for infertility and now-established treatment options remains to be seen—perhaps we need a crystal ball after all.

**Authors’ roles**

S.M.N. constructed the first draft, with all authors contributing and agreeing to the content of the final submitted version and revision.

**Funding**

No external funding was either sought or obtained for this study.

**Conflict of interest**

S.M.N., R.A.A., F.J.B., R.F. and A.L.M. have all received honoraria from Beckman Coulter and Roche for participation in Advisory Boards regarding AMH. N.R.-F. has received honoraria from GE Systems.

**References**


Anti-Müllerian hormone


Loh JS, Maheshwari A. Anti-Müllerian hormone—is it a crystal ball for predicting ovarian ageing? *Hum Reprod* 2011;26:2925–2932.


