Reference Values for the Revised Anti-Müllerian Hormone Generation II Assay: Infertile Population-based Study

Joong Yeup Lee,1 Soyeon Ahn,2 Jung Ryeoil Lee,3 Byung Chul Jee,3 Chung Hyon Kim,1 Soyeon Seo,4 Chang Suk Suh,5 and Seok Hyun Kim6

1Hamchoon Women’s Clinic, Seoul, Korea; 2Medical Research Collaborating Center, Seoul National University Bundang Hospital, Seongnam, Korea; 3Department of Obstetrics and Gynecology, Seoul National University Bundang Hospital, Seongnam, Korea; 4Samkwang Medical Laboratories, Seoul, Korea; 5Department of Obstetrics and Gynecology, Seoul National University College of Medicine, Seoul, Korea

Received: 19 September 2016 Accepted: 30 January 2017

Address for Correspondence: Seok Hyun Kim, MD, PhD Department of Obstetrics and Gynecology, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 03080, Korea E-mail: seokhyun@snu.ac.kr

Funding: This research was supported by the Korea Health Care Technology R & D Project, Ministry of Health and Welfare, Korea (grant No. A120043).

INTRODUCTION

The percentage of women delaying childbearing continues to increase, and demand for an accurate assessment of ovarian reserve has grown. Reproductive aging varies markedly among women. Anti-Müllerian hormone (AMH) is recognized as a reliable predictor of ovarian response in assisted reproduction and is a good indicator of ovarian reserve in other clinical settings (1). To demonstrate the value of AMH as a biomarker, a reliable and commercially available assay is essential (2). The original second generation (Gen II) AMH assay was developed by modification of the Immunotec (IOT) and Diagnostic Systems Laboratory (DSL) assay. The Gen II assay uses the same antibody from the DSL system, but adopted calibration against the standard from the IOT system (3). Recent publications reported a discrepancy in AMH values obtained using the original Gen II assay compared to those with the IOT assay (4), as well as a significant decrease in values compared to those with the DSL assay (5-7). Beckman-Coulter released a field safety notice stating that all kit lots (≤ 326,119) were affected by complement interference, which caused a time-dependent 70% reduction in the AMH value, and they revised their assay methodology (i.e., premixing the sample with a buffer before plating) (8). This technical issue is of critical importance, and it may have hampered the clinical use of the AMH.

Reference values do not have a meaning per se unless they are referred to in a particular context, which is usually a physiological one; thus, those values represent a specific population (9). Reference values and centile charts for AMH could serve as a tool for clinicians and patients when assessing ovarian reserve at a given age, which would assist in decision-making and selecting treatment options. The samples for reference value studies must be collected under conditions representative of those used in clinical practice (10). We previously reported on age-specific AMH values measured using the IOT assay in a cohort of 16,972 infertility women, and we confirmed that the quadratic model is the most appropriate for describing the AMH-age relationship (11). Since the revised Gen II assay is a substitution for the IOT and original Gen II assay, new age-specific AMH values are urgently needed.

The aim of this study was to establish age-specific reference values and centile chart for the revised Gen II assay. To achieve this objective, we studied mass anonymized AMH data, which were obtained in the same laboratory during routine clinical use.
MATERIALS AND METHODS

Study cohort
We recruited AMH data obtained from unselected infertile women aged 25 to 45 years old between June 2013 and June 2014. Serum AMH measurements using the revised Gen II assay were performed at Samkwang Medical Laboratories. AMH samples were collected from many infertility centers in Korea for ovarian reserve estimation. Only data on the patients’ age, hospital, race/ethnicity, and AMH levels were available from the laboratory’s registry. To collect data specifically for the purpose of developing the reference values for infertile women, we included women aged 25–45 years for infertility work-up in the analysis to reduce selection bias, and we excluded non-Korean women to rule out the influence of ethnicity (11). If there were multiple AMH values from the same patient, each woman was included only once, and the first AMH value was chosen for analysis. The study cohort was randomly divided into training or validation data groups to establish and validate the optimal AMH-age model.

The revised Gen II assay and analysis
All the AMH measurements were performed according to the manufacturers’ protocols by the same experienced staff. The Gen II kit (A79765; Beckman-Coulter, Brea, CA, USA) was used with the revised instructions (IFU-REF A92268D), which include a predilution step of the specimen with assay buffer. In the original instructions, calibrator, control, or serum and assay buffer were added sequentially to the microplate. In the revised instructions, calibrator, control, or serum and assay buffer are premixed, and then added to the microplate. The assay has a measuring range from 0.08–22.50 ng/mL. AMH values ≤ 0.08 were censored at a value of 0.08 ng/mL. The intra- and inter-assay coefficient of variation (CV) were 5.4% and 5.6%, respectively.

Statistical analysis
The total samples from the study cohort were pooled and randomly divided into a training or validation set. A set of 5 tentative regression models were applied to the training set, where the logarithm of AMH was used as a response variable in each model, and the age in years was a predictor (12). The same regression models were fitted in the validation set to ascertain the model performance. To evaluate goodness of fit, summary statistics for the sum of the absolute residuals (differences between the log [AMH] values and those predicted), $R^2$ statistic (proportion of the total variation of outcomes explained by the model), and Akaike information criterion (AIC) (a relative estimate of the information lost with a given model) were considered (7,11). The 95% confidence intervals (CIs) were produced from 1,000 bootstrap residual data sets, which were generated from random sampling with replacement from the observed residual data set. A likelihood ratio test was used when comparing the nested models. The partial F test was used to decide whether adding an additional explanatory term to the regression models would be statistically significant.

Based on the optimal AMH-age model, centile chart of the serum AMH was plotted in the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles of the predicted AMH values against age.

All analyses were performed using R for Windows, version 2.11.1 (R Foundation for Statistical Computing, Vienna, Austria).

Table 1. Characteristics of the study cohort

| Cohort | Age, yr | AMH, ng/mL |
|--------|---------|------------|
| Total No. of subjects | 15,801 |
| Mean | 34.6 | 4.40 |
| SD | 4.1 | 3.97 |
| Median | 34.1 | 3.40 |
| Upper quartile | 37.3 | 1.62 |
| Lower quartile | 31.8 | 5.97 |

AMH = anti-Müllerian hormone, SD = standard deviation.

Table 2. AMH-age regression models for the revised Gen II assay

| Model | Parameters | Estimate (95% CI) |
|-------|------------|------------------|
| Linear | log (AMH) = a + b × Age | a (5.824 (5.616, 6.031), b (−0.140 (−0.146, −0.134)) |
| Biphasic | log (AMH) = a + b × (Age − 38) | a (0.738 (0.690, 0.786), b (Age ≤ 38) (0.099 (0.090, 0.108), b (Age > 38) (−0.257 (−0.277, −0.237)) |
| Quadratic | log (AMH) = a + b × Age + c × Age² | a (−3.791 (−5.276, −2.306), b (0.410 (0.326, 0.495), c (−0.008 (−0.009, −0.007)) |
| Cubic | log (AMH) = a + b × Age + c × Age² + d × Age³ | a (3.73054 (−6.10765, 13.56874), b (−0.24374 (−1.09386, 0.60637), c (0.01099 (−0.01329, 0.03527), d (−0.00018 (−0.00041, 0.00005)) |
| Power | log (AMH) = a × Age² + c | a (−0.0014 (−0.0038, −0.0005), b (5.1266 (4.464, 5.8106), c (1.9077 (1.7821, 2.0623)) |

AMH = anti-Müllerian hormone, Gen II = second generation, CI = confidence interval.
Ethics statement
This study was approved by the Institutional Review Board of the Hamchoon Women’s Clinic (IRB No. 79507-201503-EC-001). The requirement for informed consent was waived by the board.

RESULTS
Table 1 shows the baseline demographic characteristics of the study cohort, which was randomly divided into a training or validation set. The $P$ value from the Wilcoxon-Mann-Whitney test showed no significant differences in age and the AMH values between the training and validation set ($P = 0.332$ and $0.421$, respectively). In the training set, 5 regression models were fitted, and the parameters for each model were estimated (Table 2). Table 3 presents summary statistics for the sum of the absolute residuals, $R^2$, and AIC in the validation set. The linear model had the largest sum of absolute residuals and the lowest $R^2$. Despite a small difference in the goodness of fit among the non-linear models, the quadratic, cubic, and power models had the lower sum of absolute residuals and AIC. To determine which model would best fit the data, additional statistical analyses were performed. According to the partial F test (Table 4), a quadratic and cubic model provided a significantly better fit than a linear model. There was no significant difference between the quadratic and cubic models. When 2 models fit the data equally well, a simpler model was preferred because of the ease of interpretation. Thus, we chose the quadratic model over the cubic model. Finally, the quadratic and power models appeared to provide the best fit to our data. For the sake of general usability and accessibility, we chose the quadratic model as the most appropriate clinical model for describing the association between AMH and age.

Fig. 1 shows centile charts of the predicted AMH values against the age acquired by the IOT assay and by the revised Gen II assay.

Table 3. Summary of the goodness of fit for AMH-age regression models in the validation dataset

| Model      | $R^2$ (95% CI) | Sum of absolute residuals (95% CI) | AIC          |
|------------|----------------|-----------------------------------|--------------|
| Linear     | 0.22 (0.19, 0.26) | 6,375.91 (6,249.67, 6,492.11)     | 23,828       |
| Biphasic   | 0.25 (0.21, 0.28) | 6,297.76 (6,173.46, 6,419.51)     | 23,682       |
| Quadratic  | 0.25 (0.21, 0.28) | 6,300.39 (6,171.43, 6,433.02)     | 23,668       |
| Cubic      | 0.25 (0.21, 0.28) | 6,295.61 (6,170.16, 6,420.72)     | 23,667       |
| Power      | 0.25 (0.21, 0.28) | 6,294.67 (6,173.02, 6,431.25)     | 23,665       |

A 95% CI was determined using bootstrapping methods. AMH = anti-Müllerian hormone, $R^2$ = coefficient of determination, CI = confidence interval, AIC = Akaike information criterion.

Table 4. The partial F test for comparing the model with and without an additional analysis of variance

| Parameters          | Res.Df | RSS    | Df | Sum of squared residuals | F value | $P$ value for F test |
|---------------------|--------|--------|----|--------------------------|---------|----------------------|
| Linear vs. quadratic| 7,898  | 9,156.9| 1  | 164.21                   | < 0.001  |
| Linear vs. cubic    | 7,896  | 8,878.9| 2  | 192.13                   | 1.00     |
| Quadratic vs. cubic | 7,896  | 8,878.9| 1  | 2.6886                   | 0.129    |

Res.Df = residual degree of freedom, RSS = residual sum of square, Df = degree of freedom.

Fig. 1. Centile charts for the predicted AMH values acquired by the IOT and revised Gen II assays. AMH = anti-Müllerian hormone, IOT = Immunotec, Gen II = second generation.
say; both were based on the most suitable quadratic models. There is a little difference between the shapes of 2 centile charts; the centile charts of the revised Gen II assay had a wider separation in the younger age group and a narrower separation in the older age group. The observed, not predicted, age-specific AMH values measured using the revised Gen II assay in 15,801 infertility patients and corresponding percentiles are summarized in Table 5.

**DISCUSSION**

Discrepancy in AMH values measured using different kits is a major concern that hampers their application in the reproductive field. Our study aimed to establish age-specific reference values for the revised Gen II assay with population-based data.

With the introduction of the revised Gen II assay, it is necessary to establish new cut-off values and age-specific reference values of AMH. The manufacturer alleged that the existing reference established with the IOT assay can be applied when interpreting the AMH values with the original Gen II without any problems theoretically. However, adapting clinical cut-off values from the IOT assay to the original Gen II without verification was not recommended because of a different antibody pair. This is actually correct, considering the aforementioned AMH discrepancy. We derived reference values for AMH in a large population-based infertility cohort aged 25–45 years. There are 2 approaches for establishing age-specific AMH reference values. One way is to construct percentile values from observed population data (Table 5), and the other way is using predicted AMH values based on the best model fitted to the population data (Fig. 1). Raw centiles can be calculated from appropriate data, but the perturbations will unlikely reflect the underlying AMH decrease with age at the population level. Therefore, it is reasonable to use statistical methods to derive a series of smoothed curves showing how the centiles of AMH change when plotted against age (13). We previously reported that the quadratic model optimally described AMH-age relation in an analysis of 16,972 women with AMH measurements obtained using the IOT assay (11). The current study also demonstrated that the quadratic model was best fit for describing the AMH-age relation irrespective of the types of AMH assay.

A limitation of this study was that the centile charts were constructed to provide reference values of AMH for infertile women, not for a normal population. Because of our large databases, reference values for the infertile population could be extrapolated to the general population with some assumptions. However, it is necessary to confirm and validate the reference values obtained by using a more scientifically sound a priori selection of the general population (9).

Although a low ovarian reserve does not mean decreased fertility potential, it is clear that a low AMH value is associated with a worse outcome during assisted conception. Consequently, many clinicians will advise patients with a low AMH value not to delay commencing family planning or to consider a rapid in vitro fertilization conversion as an infertility treatment.

To the best of our knowledge, this is the largest population-
based study to establish age-specific reference values of AMH using the revised Gen II assay. Further studies of larger general population are consistently required to verify the reference AMH values.

ACKNOWLEDGMENT

We wish to acknowledge Doyeong Hwang and Ki Chul Kim from the Hamchoon Women’s Clinic for their valuable comments regarding our manuscript.

DISCLOSURE

The authors have no potential conflicts of interest to disclose.

AUTHOR CONTRIBUTION

Conceptualization: Lee JY, Lee JR. Data curation: Kim CH, Seo S. Formal analysis: Lee JY, Ahn S, Lee JR. Funding acquisition: Kim SH. Investigation: Ahn S, Jee BC. Writing - original draft: Lee JY, Kim SH. Writing - review & editing: Suh CS.

ORCID

Joong Yeup Lee http://orcid.org/0000-0002-1536-3811
Soyeon Ahn http://orcid.org/0000-0003-3440-2027
Jung Ryeol Lee http://orcid.org/0000-0003-3743-2934
Byung Chul Lee http://orcid.org/0000-0003-2289-6090
Chung Hyon Kim http://orcid.org/0000-0001-7643-2975
Soyeon Seo http://orcid.org/0000-0001-8216-6562
Chang Suk Suh http://orcid.org/0000-0003-1835-7350
Seok Hyun Kim http://orcid.org/0000-0003-0649-3224

REFERENCES

1. Dewailly D, Andersen CY, Balen A, Broekmans F, Dilaver N, Fanchin R, Griesinger G, Kelsey TW, La Marca A, Lambalk C, et al. The physiology and clinical utility of anti-Müllerian hormone in women. Hum Reprod Update 2014; 20:370-85.
2. Puntmann VO. How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease. Postgrad Med J 2009; 85: 538-45.
3. 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: 51-9.
4. Li HW, Ng EH, Wong BP, Anderson RA, Ho PC, Yeung WS. Correlation between three assay systems for anti-Müllerian hormone (AMH) determination. J Assist Reprod Genet 2012; 29: 1443-6.
5. Rustamov O, Smith A, Roberts SA, Yates AP, Fitzgerald C, Krishnan M, Nordo LG, Pemberton PW. Anti-Müllerian hormone: poor assay reproducibility in a large cohort of subjects suggests sample instability. Hum Reprod 2012; 27: 3085-91.
6. 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: 1042-8.
7. Nelson SM, Illidromitl S, Fleming R, Anderson R, McConnachie A, Messow CM. Reference range for the antimüllerian hormone Generation II assay: a population study of 10,984 women, with comparison to the established diagnostics systems laboratory nomogram. Fertil Steril 2014; 101: 523-9.
8. Puntmann VO. How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease. Postgrad Med J 2009; 85: 538-45.
9. Ceriotti F, Hinzmänn R, Panteghini M. Reference intervals: the way forward. Ann Clin Biochem 2009; 46: 8-17.
10. Solberg HE, PetitClerc C. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values. Approved recommendation (1988) on the theory of reference values. Part 3. Preparation of individuals and collection of specimens for the production of reference values. J Clin Chem Clin Biochem 1988; 26: 593-8.
11. Lee JY, Lee BC, Lee JR, Kim CH, Park T, Yeon BR, Seo SY, Lee WD, Suh CS, Kim SH. Age-related distributions of anti-Müllerian hormone level and anti-Müllerian hormone models. Acta Obstet Gynecol Scand 2012; 91: 970-5.
12. Silverwood RJ, Cole TJ. Statistical methods for constructing gestational age-related reference intervals and centile charts for fetal size. Ultrasound Obstet Gynecol 2007; 29: 6-13.
13. Bonellie S, Chalmers J, Gray R, Greer J, Jarvis S, Williams C. Centile charts for birthweight for gestational age for Scottish singleton births. BMC Pregnancy Childbirth 2008; 8: 5-15.