

Just Accepted by Climacteric

Predicting menopausal age with anti-Müllerian hormone: a cross-validation study of two existing models

F. Ramezani Tehrani, M. Dólleman, J. van Disseldorp, S. L. Broer, F. Azizi, M. Solaymani-Dodaran, B. C. Fauser, J. S. E. Laven, M. J. C. Eijkemans and F. Broekmans

doi:10.3109/13697137.2014.898264

ABSTRACT

Objective: This study aimed to cross validate two comparable Weibull models *of prediction of age at* natural menopause (ANM) from two cohorts (SRV and TLGS cohorts). It summarizes advantages and disadvantages of the models and underlines the need for achieving correct time dependency in dynamic variables like AMH.

Methods: Models were fitted in the original datasets and then applied to the cross validation datasets. *The discriminatory capacity of model* was assessed by calculating c-statistics for the models in their own data and in the cross validation data. Calibration of the models on the cross validation data was assessed by measuring the slope, intercept and Weibull shape.

Results: C-statistic for the SRV model on the SRV data was 0.7 (0.7-0.8) and on the TLGS data *was* 0.8 (0.8-0.9). For the TLGS model on the TLGS data it was 0.9 (0.8-0.9) and in the SRV data it was 0.7 (0.6-0.8). After calibration of the SRV model on the TLGS data the slope was 1, the intercept -0.3 and the shape 1.1. The TLGS model on the SRV data had a slope of 0.3, an intercept of 12.7 and a shape of 0.6.

Conclusions: Both models discriminate well between women that enter menopause early or late during follow-up. While the SRV model showed good agreement between the predicted risk of entering menopause and the observed proportion of women who entered menopause during follow-up (calibration) in the cross validation dataset, the TLGS model showed poor calibration.

© 2014 The International Menopause Society. This provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

DISCLAIMER: The ideas and opinions expressed in the journal's *Just Accepted* articles do not necessarily reflect those of Informa Healthcare (the Publisher), the Editors or the journal. The Publisher does not assume any responsibility for any injury and/or damage to persons or property arising from or related to any use of the material contained in these articles. The reader is advised to check the appropriate medical literature and the product information currently provided by the manufacturer of each drug to be administered to verify the dosages, the method and duration of administration, and contraindications. It is the responsibility of the treating physician or other health care professional, relying on his or her independent experience and knowledge of the patient, to determine drug dosages and the best treatment for the patient. *Just Accepted* articles have undergone full scientific review but none of the additional editorial preparation, such as copyediting, typesetting, and proofreading, as have articles published in the traditional manner. There may, therefore, be errors in *Just Accepted* articles that will be corrected in the final print and final online version of the article. Any use of the *Just Accepted* articles is subject to the express understanding that the papers have not yet gone through the full quality control process prior to publication.

Predicting menopausal age with anti-Müllerian hormone: a cross-validation study of two existing models

F. Ramezani Tehrani¹, M. Dölleman^{2,3}, J. van Disseldorp², S. L. Broer², F. Azizi¹, M. Solaymani-Dodaran¹, B. C. Fauser², J. S. E. Laven⁴, M. J. C. Eijkemans³ and F. Broekmans²

¹Reproductive Endocrinology Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran, ²Department of Reproductive Medicine and Gynaecology, University Medical Center Utrecht, ³Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands, ⁴Division of Reproductive Medicine, Department of Obstetrics and Gynecology, Erasmus Medical Center, University Medical Center, Rotterdam

Correspondence: F. Ramezani Tehrani, Reproductive Endocrinology Research Center, Research Institute for Endocrine Sciences, 24 Parvaneh, Yaman Street, Velenjak, P.O. Box 19395-4763, Tehran, I.R. Iran. Postal Code 1985717413; E-mail: ramezani@endocrine.ac.ir

JUST ACCEPTED

ABSTRACT

Objective: This study aimed to cross validate two comparable Weibull models *of prediction of age at natural menopause (ANM) from two cohorts (SRV and TLGS cohorts)*. It summarizes advantages and disadvantages of the models and underlines the need for achieving correct time dependency in dynamic variables like AMH.

Methods: Models were fitted in the original datasets and then applied to the cross validation datasets. *The discriminatory capacity of model* was assessed by calculating c-statistics for the models in their own data and in the cross validation data. Calibration of the models on the cross validation data was assessed by measuring the slope, intercept and Weibull shape.

Results: C-statistic for the SRV model on the SRV data was 0.7 (0.7-0.8) and on the TLGS data was 0.8 (0.8-0.9). For the TLGS model on the TLGS data it was 0.9 (0.8-0.9) and in the SRV data it was 0.7 (0.6-0.8). After calibration of the SRV model on the TLGS data the slope was 1, the intercept -0.3 and the shape 1.1. The TLGS model on the SRV data had a slope of 0.3, an intercept of 12.7 and a shape of 0.6.

Conclusions: Both models discriminate well between women that enter menopause early or late during follow-up. While the SRV model showed good agreement between the predicted risk of entering menopause and the observed proportion of women who entered menopause during follow-up (calibration) in the cross validation dataset, the TLGS model showed poor calibration.

JUST ACCEPTED

INTRODUCTION

Anti-müllerian hormone (AMH) has emerged as an important biomarker of ovarian ageing. As menopause is the final event that marks depletion of the ovarian reserve, AMH has been suggested to predict the age at which a woman will enter menopause¹⁻⁶. Although the mean age at menopause occurs at around 51 years, a considerable range exists from 40-60 years with approximately 10% of women in the general population becoming menopausal by the age of 45 years⁷. Preceding reproductive events such as cycle irregularity in the perimenopause and subfertility are suggested to happen in a fixed temporal fashion with the end of natural fertility occurring approximately 10 years before the final menstrual period. Identifying women with early menopause and thus early subfertility could be used for the primary prevention of infertility by counseling those women to conceive early or to cryopreserve oocytes. Furthermore accurate estimation of time of menopause could identify those at higher risk of cardiovascular disease, osteoporosis, and breast or endometrial cancers due to early or late menopause.

Four models for predicting age at menopause with AMH have been presented in recent literature^{1; 3-5}. This study aims to solidify existing evidence that AMH-based predictions of age at natural menopause (ANM) is possible by cross-validating two existing models of menopause prediction from the two most comparable studies^{1; 3; 6}. We further attempt to make suggestions towards an optimal predictive model for individualized fertility forecasting based on *the* serum AMH concentration.

MATERIALS AND METHODS

Patients

Data came from two different studies namely the SRV (Scheffer, van Rooij, de Vet) cohort and the TLGS (Tehran Lipid and Glucose Study) cohort.

The SRV cohort study

The SRV cohort is a combined population from three highly comparable prospective longitudinal studies on ovarian function from the Netherlands that were used for the prediction of menopause in a previous publication¹. It consists of 257 normo-ovulatory women aged between 21 and 46 years. Women were selected if they had regular and predictable cycles, no history of infertility or endocrine disorders, if they had not used contraceptives for at least 3 months, and if they had no history of ovarian or uterine surgery. In these women, baseline (T1) AMH was determined. At follow up (T2), which was approximately 11 years later, *cycle's* status was reassessed by questionnaires in 158 women. From these, 48 women were post-menopausal. The study was approved by the medical ethical review committee and all participants gave written informed consent. For a more detailed description the reader is referred to Broer et al, 2011¹.

The TLGS cohort study

The TLGS is an ongoing prospective longitudinal study in Tehran that was set up to assess prevalence and risk factors of non-communicable diseases. From the original cohort of 2,412 women, 1,265 women met the eligibility criteria and from these women 266 study participants were randomly selected to model age at menopause using AMH. Eligible women were aged between 20 and 50 years, they had regular and predictable menstrual cycles, they had not taken any form of hormonal contraceptives for at least three months, they were proven fertile defined as having at least one-term pregnancy within 1 year after stopping contraception and they had no history of ovarian or uterine surgery. AMH was measured at baseline (T1) and reproductive status was assessed by questionnaires on average 6 years later (T2). From the 266 original participants, 63 women reached menopause. The study was approved

by the institutional review board and all participants gave written informed consent. For a more detailed description the reader is referred to Tehrani et al 2011³.

Definitions and Endpoints

In the SRV cohort a regular and predictable cycle was defined as a cycle of 19-35 days where consecutive menses were predictable within 7 days. In the TLGS cohort study a regular cycle was defined as a predictable cycle with 21-35 day intervals. The endpoint of this cross validation was menopause. Both studies defined menopause as a period of amenorrhea *for* at least 12 consecutive months, for which no other pathologic or physiologic cause was present. Menopause was assessed at follow up, on average 11 years after inclusion in the SRV cohort and on average 6 years after inclusion in the TLGS cohort.

Predictors

The variables included in the predictive model were age at AMH measurement and serum AMH level. The measurement techniques for AMH were different as described below.

AMH assay

In the SRV cohort, all AMH values were measured in the same laboratory. In one of the three studies AMH was measured using the Active MIS/AMH ELISA from Diagnostics Systems Laboratories (Webster, TX). This system had an interassay and intraassay coefficient of variation of less than 5% at the level of 3.0 ng/ml and less than 11% at the level of 13.0 ng/ml. The limit of detection (LOD) of the assay was 0.026 ng/ml. In the other two studies, contributing to the SRV cohort, AMH was measured with the ultrasensitive immunoenzymometric assay (Immunotech-Coulter, Marseille, France). The limit of detection was 0.05 ng/ml and intra- and interassay coefficients of variation were less than 5% and less than 8%, in the two studies, respectively. A correction factor was required to justify comparing and pooling of AMH samples from these three studies. As previously described and successfully applied, serum AMH concentration measured in the Immunotech-Coulter assay was corrected by a factor of 0.5 to be comparable to the DSL values²¹.

In the TLGS cohort AMH was measured using the Active MIS/AMH ELISA from Diagnostics Systems Laboratories (DSL-10-14400, Webster, TX). The LOD was 0.006 ng/mL the intra and interassay coefficients of variation were 5.2 and 9.1, respectively. To compare measures of AMH in the SRV cohort to the TLGS cohort, both AMH values were converted into values representative of the new AMH Generation II assay (AMH Gen II assay) from Immunotech-Coulter by applying a conversion factor. For the SRV cohort a lab-specific conversion factor was determined which amounted to 'DLS x 1.564=Gen II'. For the TLGS cohort we applied a conversion factor of 'DSL x 1.4 – 0.0868= Gen II' as previously described in the literature^{8,9}.

Predictive models

In the SRV cohort, the predictive model was based on female age and AMH. A Weibull survival model was built with female age on the time axis, with delayed entry at T1 and age-specific percentiles of AMH as a single covariate. Age-specific AMH levels were calculated per participant as follows: first all AMH values were log-transformed, then a flexible spline function was applied to the scatter plot of log(AMH) with age at T1 and with the assumption that there was a normal distribution of residuals around this fitted curve. In a tabular presentation of the model, AMH measurement values, and their corresponding percentiles were shown per age category, as well as the 5th, 25th, 50th and 75th percentile values of the predicted ANM¹.

The TLGS prediction of ANM was also based on a Weibull survival model with female age on the time axis, and delayed entry at T1. This model contained both Age at T1 and AMH as covariates.

The following predictive equation was obtained, where 0.5 is the constant to determine the median ANM and age T1 is the age at inclusion:

$$\text{Age at menopause} = (-\ln(0.5))^{-0.037} \exp(0.10 \text{ AMH} + 0.016 \text{ ageT1} + 3.2) \quad ^6$$

Statistical Analysis

Data from both studies were converted to contain the same format with regard to variables and coding. Baseline measures were calculated for both studies and compared to the originally published articles. Baseline characteristics were described as means with standard deviation for continuous variables and numbers with percentages for categorical variables. Baseline measures were compared between cohorts using a Mann-Whitney-U test for continuous variables and a chi-squared test for binomial variables. Kaplan Meier curves were created with female age on the time axis and delayed entry at T1, where women were either censored at the end of their follow-up, or continued until their age at menopause.

The basic principle in this analysis was to fit the SRV regression model on the SRV data and to then apply this model to the TLGS data. The performance of the prediction rule was studied in terms of discrimination and calibration. Discrimination is quantified by C-statistics and it reflects the capacity of the models to discriminate between women who will enter menopause early or late during follow-up. Calibration reflects the agreement between the predicted risk of entering menopause and the observed proportion of women who entered menopause during follow-up. For a Weibull model, calibration may be assessed through fitting a Weibull model on the validation data using a transformed time axis according to the baseline of the prediction model that is to be validated, and containing the linear predictor of that model, evaluated in the validation data, as a single covariate. A perfect calibration has an intercept of zero, a slope of 1 and the Weibull shape will be equal to 1. Because the models are fitted to their own data (which means that the model will always be perfectly calibrated in its original dataset) the calibration was only done for the model in the cross validation data. All analyses were then repeated, but with reversed roles, where the TLGS cohort was used to fit the model and then the performance of that predictive rule was assessed in the SRV data, as described above.

In order to also graphically illustrate differences between the two models, we plotted their Weibull curves for predicted age at menopause for four given ages: 25, 30, 35 and 40 years. Within each age category, three AMH values were chosen to represent a low age-specific AMH, an average age-specific AMH and a high age-specific AMH. For a 25-year old, the selected AMH values were 0.8ng/ml, 3ng/ml and 7 ng/ml. For a 30 year old the selected AMH values were 0.8ng/ml, 2.5ng/ml and 5ng/ml. For a 35 year old the selected AMH values were 0.3 ng/ml, 1.5 ng/ml and 3 ng/ml. These values were 0.2 ng/ml, 1 ng/ml and 2 ng/ml for a 40 year old. Data was analyzed with SPSS 15.0 (SPSS Inc., Chicago, IL) and R version 2.9.0. (<http://www.r-project.org/>).

RESULTS

Baseline characteristics of both cohorts are displayed in table 1. The mean age at inclusion in both cohorts was approximately the same, but the variation in age at inclusion was higher in the TLGS cohort (35.3 years +/-6SD and 37.6 years +/-10 SD for the SRV and TLGS cohorts, respectively; p=0.57). This explains how approximately the same proportion of women entered menopause during follow-up (48 women (26%) in SRV versus 63 women (24%) in TLGS) despite a shorter time to follow-up in the TLGS cohort. The Kaplan Meier curves in Figure 1 show the similarity between the two cohorts with regard to the occurrence of menopause during follow up.

The C-statistic for the SRV model in the SRV data was 0.73 (95%CI 0.65-0.82) and for the SRV model in the TLGS data was 0.82 (95%CI 0.75-0.88). When the TLGS model was applied to the TLGS

data the c-statistics was 0.88 (95% CI 0.83-0.94), in the SRV data it was 0.72 (95% CI 0.63-0.81) (Table 2).

Results of the *calibration are* displayed in table 3. When the SRV model was calibrated to the TLGS dataset, the slope was 0.99, the intercept was -0.26 and the shape 1.13. The slope, which is close to a value of 1, indicates that the predicted risk of entering menopause and the observed proportion of women who entered menopause during follow up are in good agreement. The negative intercept *indicates the systematically overestimate of menopausal age when SRV model calibrated in* the TLGS data set. Conversely, when the TLGS model was calibrated on the SRV data, the slope was 0.30, the intercept was 12.66 and the shape was 0.62. The slope is not close to a value of 1, indicating poor agreement between the predicted risk of entering menopause and the observed proportion of women who entered menopause during follow-up. The intercept is higher than 0, indicating an under estimation of ANM, and the shape is also not close to a value of 1.

Figures 2a-2d compare the Weibull curves for the prediction of menopause based on age and AMH in the two predictive models. The solid curves show the predicted ages at menopause for women with a low age-specific AMH, the dotted curves for women with an average age-specific AMH and the dashed curves for women with a high age-specific AMH. The apex of the curves corresponds to the most likely predicted age at menopause with the left and right tails of the curve representing the total range of possible ages at which menopause could occur at a given AMH concentration. The upright Weibull curves represent the predicted age at menopause by the SRV model at the specified AMH values and the upside-down Weibull curves represent predicted age at menopause by the TLGS model at those same specified AMH values. Both models give a Weibull distribution curve that is skewed with a longer left tail. It is evident from these graphs that the predictions according to the SRV model lie closer together and have wide distributions that overlap. Regardless of the age at inclusion, (the age at which AMH was measured), a woman with a low age-specific AMH has a mean ANM of approximately 49 years, a woman with an average age-specific AMH has a mean ANM of approximately 51 years and a woman with a high age-specific AMH has a predicted mean ANM of approximately 54 years.

The Weibull curves for the TLGS model result in predicted ANMs that are spread further apart for women of the same age with different AMH values. The corresponding distributions are narrower than those in the SRV cohort. While a 25-year old with a low age-specific AMH is predicted to reach menopause at 38, a 25 year old with a high age-specific AMH is predicted to reach menopause at 53. Conversely, for a 30 year old the predicted age at menopause is 41 for woman with a low AMH and 49 for a woman with a high AMH. This trend continues so that a 35 year old with a low or a high AMH is predicted to reach menopause at 44 or 50 years respectively and a 40 year old woman with a low or high age-specific AMH has a predicted ANM of 47 and 54 years, respectively (figures 2a-2d).

DISCUSSION

This study aimed to cross validate and *compares the two* models of menopausal *age's* prediction, using age and AMH. With this study, we have provided further evidence that prediction of age at menopause

in the cohort then it is easier to discriminate those who will and will not become *menopause* during follow up than if all women are of late reproductive age). The TLGS cohort had a larger variation in both the age at which AMH was measured and a larger variation in AMH values (see table1).

We have described that in general, the SRV model slightly overestimates age at menopause when applied to the TLGS cohort, and conversely that the TLGS model underestimates age at menopause when applied to the SRV cohort. There are several factors that must be addressed when trying to understand the incongruence of the models on the external data. The median age at menopause, as measured by the Kaplan Meier curves, is approximately the same in the SRV cohort and the TLGS cohort (51.9 is SRV versus 51.5 in TLGS). Therefore, this cannot explain why the one model overestimates ANM and the other underestimates *it*. Furthermore, the Kaplan Meier curves overlap one another reasonably.

There are some discrepancies in the data, however. First of all, from the table with baseline characteristics it is evident that despite the shorter follow up time, and lower age at follow up in the TLGS study, approximately the same proportion of women has become *menopause* during the study (26% in TLGS versus 24% in SRV). One would expect this proportion to be smaller. One explanation is that the range of ages at inclusion is much larger in the TLGS cohort with 107 out of the 266 women with an age of 40 or above. Perhaps this discrepancy explains why the SRV model overestimates and the TLGS underestimates age at menopause, the TLGS model expects women to become menopausal sooner during follow up while the SRV model expects women to become menopausal later during follow up. It may also be argued that the relative risk for menopause occurring at a certain age differs between different ethnicities. However, the median age at menopause discerned from the TLGS cohort was 49.7 years which is comparable to the median age at menopause from large Caucasian cohort studies (49.2 years).

Another issue that remains to be discussed is that the AMH values in the TLGS cohort are higher than in the SRV cohort despite the age at AMH measurement being slightly older in the TLGS cohort. This could be due to differences in lifestyle factors in the cohort, for example smoking has been shown to be an important determinant of AMH¹⁰. In the SRV cohort 16.7% of women smoked, in the TLGS this was 1.2%. Another plausible solution may be that some incongruity in AMH assay system has remained despite application of suitable conversion factors. However, this is probably of more interest to the TLGS model than the SRV model as the SRV model uses percentiles of AMH and not absolute AMH values. A woman's percentile, or rank in a population, does not change after application of a conversion factor, while one's absolute AMH value may change considerably. Nevertheless, if it were only a problem associated with a conversion factor problem one would expect a systematic incongruence between the observed and predicted ages at menopause, and not a poor performance of all the calibration parameters. The difference in model performance can thus not merely be due to the value of AMH after conversion to the GenII assay, but may also be due to the way in which AMH is incorporated into the predictive model. Both models use age from birth onwards as their time axis with delayed entry of *a woman* at her age at T1. The challenge for AMH is that it is age, and thus follow-up time, dependent. In the SRV model AMH is assigned per woman as a percentile for her age category,

incorporated is such that within a certain age group a higher AMH is associated with a higher age at menopause leading to a positive coefficient for age. In the SRV model, age indeed is positive while AMH is negative. The effect of the different incorporation of age is evident in figures 2a-2d. The SRV model is barely influenced by age but only by age-specific AMH. Regardless of whether a woman is 30 or 40 years old, if she remains in the lower percentiles of AMH, her ANM will be lower (around 49) in compare to a woman who remains in the higher percentiles of AMH who will enter menopause at around 54. The resulting predictions of age at menopause widely overlap due to a wide range of predictions. The TLGS model, in contrast, is highly affected by both age and AMH. This leads to mean predictions that are spread wide apart and with narrower ranges. Although the discriminatory capacity of the TLGS model is appealing and it has a higher c-statistic, the question remains whether it is also clinically realistic. In the general population there are very few women that enter menopause before the age of 40 or 45. The TLGS model seems to suggest otherwise, when looking at their predictions of young reproductive women who would enter menopause at around 43 years with an age-specific AMH level in the average range. With increasing age the models start performing more similarly as shown in figure 2d. However, the target group, in which predicting age at menopause is most valuable, *is* women that are younger than forty. Therefore, the predicted ages at menopause must also be precise for this group of women. Perhaps the TLGS model extrapolates too extremely for young women. Considering these results it could be suggested that the use of age-specific AMH percentiles yields more realistic prediction of age at menopause than absolute AMH values that are age-corrected in the model where age is already on the time axis. However, it may be more limited in its ability to predict both extreme ends of the spectrum for menopausal age.

The major strength of this study was that the two studies had a very similar set-up with the same variables in the predictive model and with the same endpoint. A challenge of the studies was making the AMH values comparable. To compare the SRV measures of AMH to the TLGS cohort, both AMH values were converted into values representative of the new AMH Generation II assay (AMH Gen II assay) from Immunotech-Coulter by applying a conversion factor. For the SRV cohort *a* lab-specific conversion factor was determined which amounted to 'DLS x 1.564=Gen II'. For the TLGS cohort the conversion factor of 'DSL x 1.4 - 0.0868= Gen II' as previously described in literature ^{8; 11}. Determination of a specific conversion factors between the two cohorts used in this study was not possible as the original TLGS serum samples were no longer available. For the same reason, a lab-specific Gen II conversion factor could not be made for the TLGS cohort. Nevertheless, after applying these conversion factors the AMH values were more comparable than before applying the Gen II conversion factor. Although the mean AMH values still differ significantly at baseline, this may also be due to differences in the cohort and not to the AMH conversion factor alone.

Two other prospective studies, both performed in the same cohort of women, have also provided evidence for the notion that AMH can be used to make a more *individualized* prediction of age at menopause. The cohort consisted of women of later reproductive age (the mean age was 40 years), with a follow up of 14 years. The first study ⁴ confirmed that AMH is a valuable predictive marker of time to menopause, and the second study elaborates that the rate of change of AMH is a better predictor of

CONCLUSIONS

We have shown that any predictive model must incorporate AMH adequately to make it both age and time dependent, preferably with age-specific AMH values as a time-dependent factor. The current *cross validation study has provided further* evidence that the prediction of age at menopause with AMH has definite potential. Correct prediction of age at menopause may pave the road to individualized prevention of primary age-related infertility and menopause-related conditions, like cardiovascular disease and breast cancer.

ACKNOWLEDGEMENTS

Hereby we thank all participants for volunteering and putting time and effort into this longitudinal study. Thanks to you, research on natural fertility in healthy women is possible.

Disclosure summary: FRT, MD, JvD, SLB, FA, MSD and MJC have nothing to disclose. F.J.B. has received fees and grant support from the following companies (in alphabetic order): Ferring, Gedeon Richter, Merck Serono, MSD and Roche. B.C. Fauser has received fees and grant support from the following companies (in alphabetic order); Andromed, Ardana, Auxogyn, Ferring, Genovum, Gideon-Richter, Merck Serono, MSD, Organon, Pantharei Bioscience, PregLem, Roche, Schering, Schering Plough, Serono, Watson laboratories and Wyeth. J.S.E.L. has received fees and grant support from the following companies (in alphabetic order): Ferring, Genovum, Merck-Serono, Organon, Schering Plough and Serono.

JUST ACCEPTED

References

- (1) Broer SL, Eijkemans MJ, Scheffer GJ et al. Anti-mullerian hormone predicts menopause: a long-term follow-up study in normoovulatory women. *J Clin Endocrinol Metab* 2011 August;96(8):2532-9.
- (2) van Disseldorp J., Faddy MJ, Themmen AP et al. Relationship of serum antimullerian hormone concentration to age at menopause. *J Clin Endocrinol Metab* 2008 June;93(6):2129-34.
- (3) Tehrani FR, Shakeri N, Solaymani-Dodaran M, Azizi F. Predicting age at menopause from serum antimullerian hormone concentration. *Menopause* 2011 July;18(7):766-70.
- (4) Freeman EW, Sammel MD, Lin H, Gracia CR. Anti-Mullerian Hormone as a Predictor of Time to Menopause in Late Reproductive Age Women. *J Clin Endocrinol Metab* 2012 February 29.
- (5) Freeman EW, Sammel MD, Lin H, Boorman DW, Gracia CR. Contribution of the rate of change of antimullerian hormone in estimating time to menopause for late reproductive-age women. *Fertil Steril* 2012 November;98(5):1254-9.
- (6) Tehrani FR, Solaymani-Dodaran M, Tohidi M, Gohari MR, Azizi F. Modeling age at menopause using serum concentration of anti-mullerian hormone. *J Clin Endocrinol Metab* 2013 February;98(2):729-35.
- (7) te Velde ER, Pearson PL. The variability of female reproductive ageing. *Hum Reprod Update* 2002 March;8(2):141-54.
- (8) Nelson SM, La MA. The journey from the old to the new AMH assay: how to avoid getting lost in the values. *Reprod Biomed Online* 2011 October;23(4):411-20.
- (9) Kelsey TW, Wright P, Nelson SM, Anderson RA, Wallace WH. A validated model of serum anti-mullerian hormone from conception to menopause. *PLoS One* 2011;6(7):e22024.
- (10) Dolleman M, Verschuren WM, Eijkemans MJ et al. Reproductive and Lifestyle Determinants of Anti-Mullerian Hormone in a Large Population-based Study. *J Clin Endocrinol Metab* 2013 March 26.
- (11) 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 July;48(Pt 4):370-3.

TABLE LEGENDS

Table 1

The baseline characteristics

	SRV	TGLS	P-value
Number of women in cohort	185	266	
Age at AMH measurement (T1)	35.5 (5.9)	37.6 (9.6)	0.57
Mean Age at Follow up (T2)	47.5 (6.11)	42.63 (8.09)	<0.001
AMH (ng/ml)	2.8 (2.8)	4.1 (4.1)	0.02
BMI (T1)	24.0 (4.0)	27.7 (5.0)	<0.001
Number of post-menopausal women at follow up	48 (25.9%)	63 (23.7%)	0.561*

Comparison of the baseline characteristics between two cohorts. All p-values are calculated with independent samples t-tests, except those denoted with a * where the p-value is measured with a chi-squared analysis.

Table 2

Accuracy of the models in their original dataset and in the cross validation dataset

	Model applied to:			
	SRV		TLGS	
Model sampled from:	C-statistic	95% CI	C-statistic	95% CI
SRV	0.73	(0.65-0.82)	0.82	(0.75-0.88)
TLGS	0.72	(0.63-0.81)	0.88	(0.83-0.94)

This table shows the c-statistics for the two models in their own data and in the cross validation data. C-statistics represent the discriminatory capacity of the model, i.e. the accuracy with which the model can distinguish between those women that will enter menopause early during the follow up period and those ones that will enter menopause late during the follow up period.

Table 3

Calibration of the models in the cross validation dataset

	Slope	Intercept	Shape
	(95% CI)	(95% CI)	(95% CI)
SRV model on TLGS data	0.99	-0.26	1.13
	(0.69-1.28)	(-0.51- 0.00)	(1.08-1.39)
TLGS model on SRV data	0.30	12.66	0.62
	(0.14-0.46)	(12.15-13.18)	(0.50-0.77)

This table shows the calibration of the models on the cross validation data. Calibration consists of the calculation of the slope, intercept and Weibull shape. A perfect calibration has an intercept of zero, a slope of 1 and a Weibull shape of 1. Calibration reflects the agreement between the predicted risk of entering menopause and the observed proportion of women who entered menopause during follow-up.

JUST ACCEPTED

FIGURE LEGENDS

Figure 1

Kaplan Meier curves depicting the decrease in the proportion of premenopausal women with increasing age at follow-up.

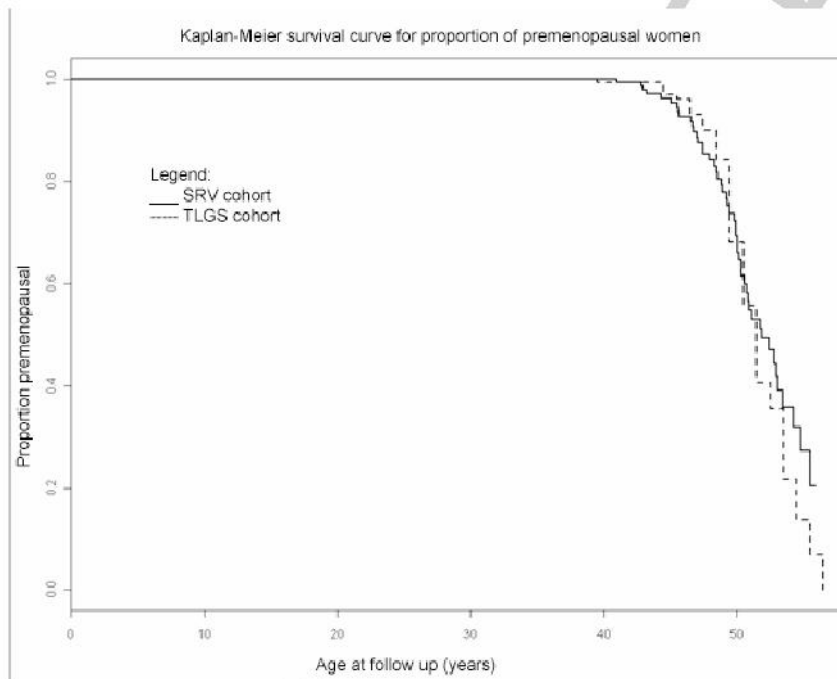
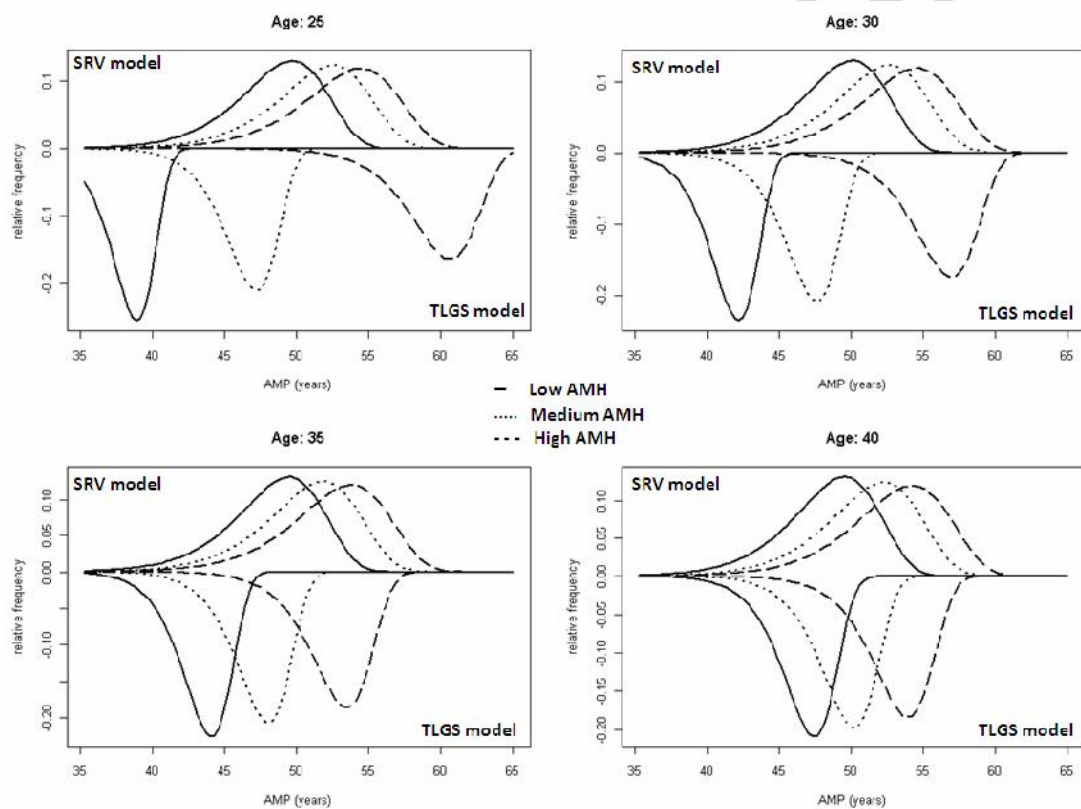


Figure 2

Differential distribution of age at menopause according to the age and AMH in the two different models. This figure shows the distribution of ANM for women in 4 age categories with a low AMH for their age (red curve), an average AMH for their age (green curve) and a high AMH for their age (blue curve). The upright model shows the predicted distributions of ANM in the SRV model and the upside down curve for the predicted distributions according to the same age and same AMH in the TLGS model. The figure displays the increasing conformity of the two models with increasing age.



JUS