Ovarian reserve tests in the prediction of the fertile lifespan and current fertility

Martine Depmann

Ovarian reserve tests in the prediction of the fertile lifespan and current fertility

PhD thesis, Utrecht University; with a summary in Dutch

Contact: m.depmann@gmail.com
ISBN: 978-90-393-6604-2
Cover photo: Karin Vermeer
Printing: Printsupport4you, Meppel

Ovarian reserve tests in the prediction of the fertile lifespan and current fertility

Het gebruik van ovariële reserve testen in het voorspellen
van de vruchtbare levensfase en de huidige vruchtbaarheid
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag
van de rector magnificus prof. dr. G.J. van der Zwaan, ingevolg het besluit
van het college voor promoties
in het openbaar te verdedigen op
donderdag 8 september 2016
des ochtends te 10.30 uur

door
Martine Depmann
geboren op 27 september 1984
te Tilburg

Promotoren: Prof. dr. F.J.M. Broekmans

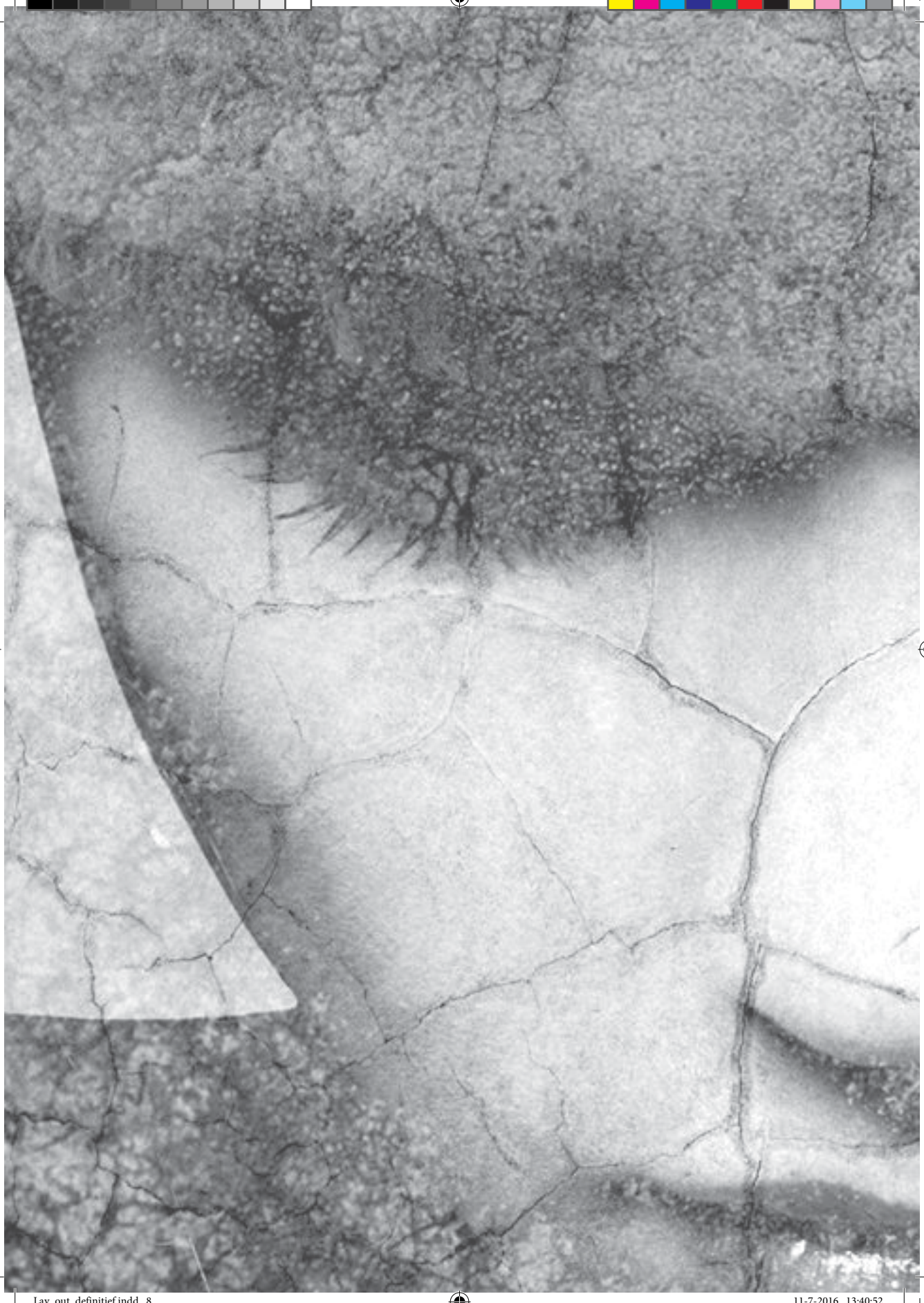
Prof. dr. B.W.J. Mol

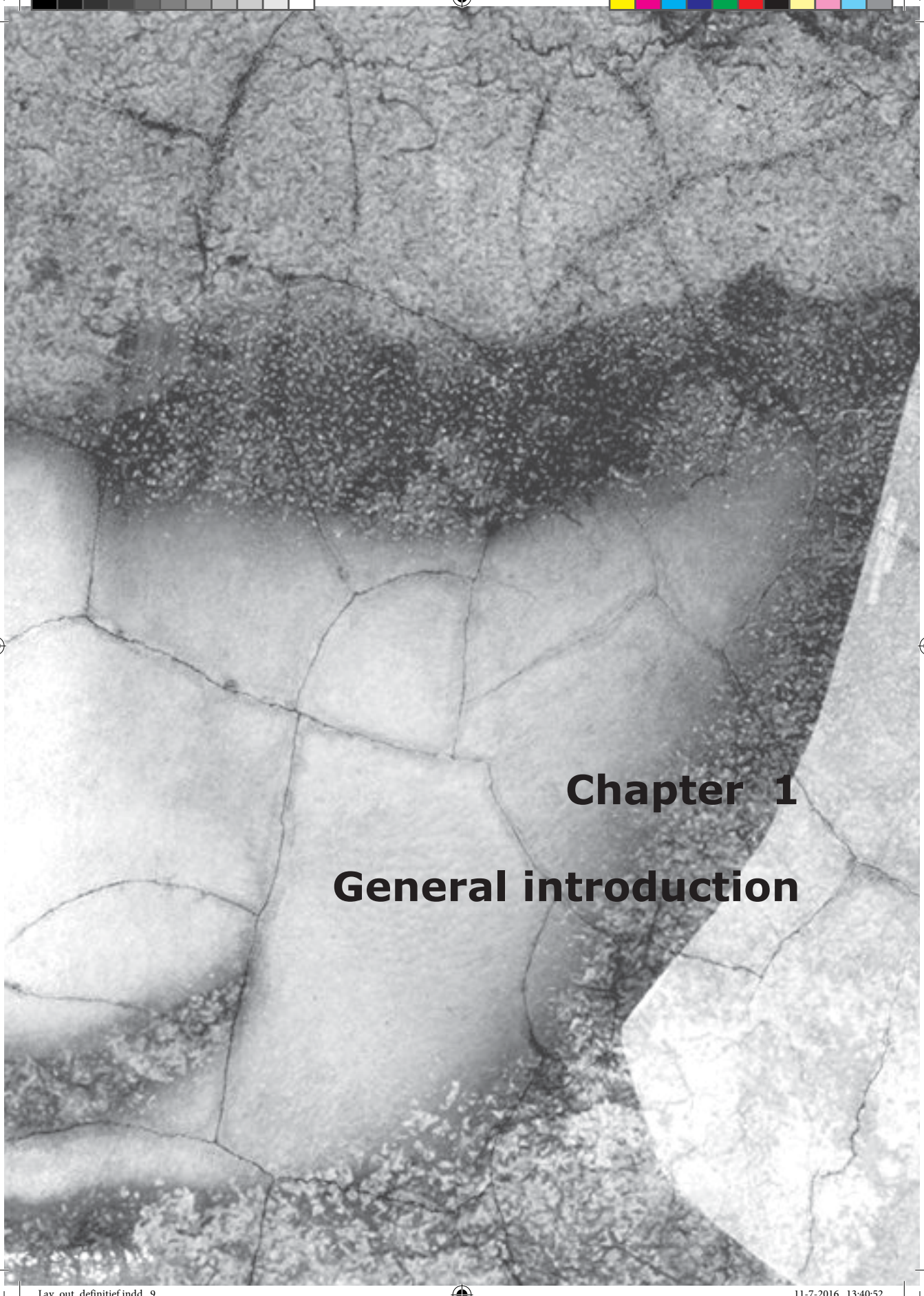
Copromotor: Dr. S.L. Broer

Publication of this thesis was financially supported by:
Chipsoft, Ferring Pharmaceuticals, Gedeon Richter, Goodlife Pharma, Memidis Pharma, BMA Mosos, Olympos, Origio, Roche and Will Pharma.

Aan mijn ouders

Contents





Chapter 1

General introduction

The limits of female fertility

From a biological point of view, the optimal timeframe for women to have children lies between 18 and 30 years of age [1]. After the age of 30, female fertility starts to decline, conception rates per cycle fall and sterility eventually occurs. In the present time it is difficult to assess female fertility since the widespread use of contraceptives, allowing for a restriction in childbearing, makes it nearly impossible to estimate true biological potential. Looking at historical cohorts though, where contraceptives were not yet invented and postponing of childbearing in favor of a female career was not relevant, one can assess the true limits of the female fertile lifespan. In a recent study [2], female fertility was assessed in such a historical cohort using age at birth of the last child as a proxy variable for sterility. It was demonstrated that age-related fertility starts to decline at the average age of 30 years, decreasing slowly up till the age 35-40 years, after which it is lost at an accelerating pace until finally sterility occurs. Moreover, this state of natural sterility is present in 50% of women at age 41, compared to 90% and 100% in women aged 45 and 50 years respectively. Translation of the subject of age- related subfertility to the present time, where fertility treatment is an option, was performed by another recent study [1]. In this study, it was calculated that in order to have a 90% chance of reaching the desired family size, women who wish to have 2 children should start attempting to conceive no later than age 31, see Figure 1.1. If, however, In Vitro Fertilization (IVF) is an option, one can start attempting to conceive at age 35.

Nonetheless, there is always a group of women for whom even fertility treatment is simply commenced too late. Due to the large biological variation in age at onset of declining fertility, it is difficult to differentiate which women will lose reproductive potential at an early age and which women can conceive even up till a very high age. For women who will eventually show never to be able to reproduce, or for women who will be faced with tremendous efforts to reproduce without fulfilling their desired family size, together estimated to comprise 15% of all women, an outline of their individual fertile life span could prevent involuntary childlessness.

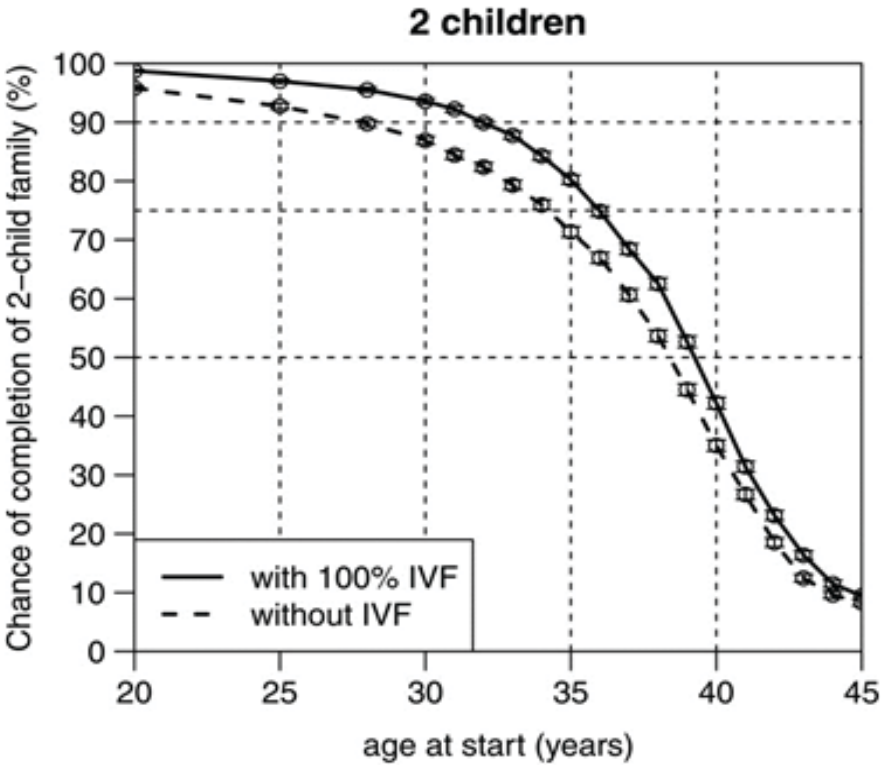


Figure 1.1: Relationship between the female age at which couples start building a family and the chance of realizing a family with one child, with and without use of IVF. The short lines above and below each point in the graph indicate the 95% confidence intervals (reproduced with permission [1]).

The relation between the end of natural fertility and menopause

Menopause, the cessation of menstrual cycles, marks the end of a woman's reproductive life. In Western countries, the mean age at menopause is 51 years. However, and as depicted in Figure 1.2, a large age interval for the occurrence of menopause ranging from 40-60 years accompanies this mean age [3]. Population based studies have demonstrated the end of natural fertility to take place on average ten years prior to menopause, preceded by a period of gradually declining natural fertility. In these studies it was observed that the end of natural fertility occurs with an age interval length similar to the length of the age interval of menopause [4,5].

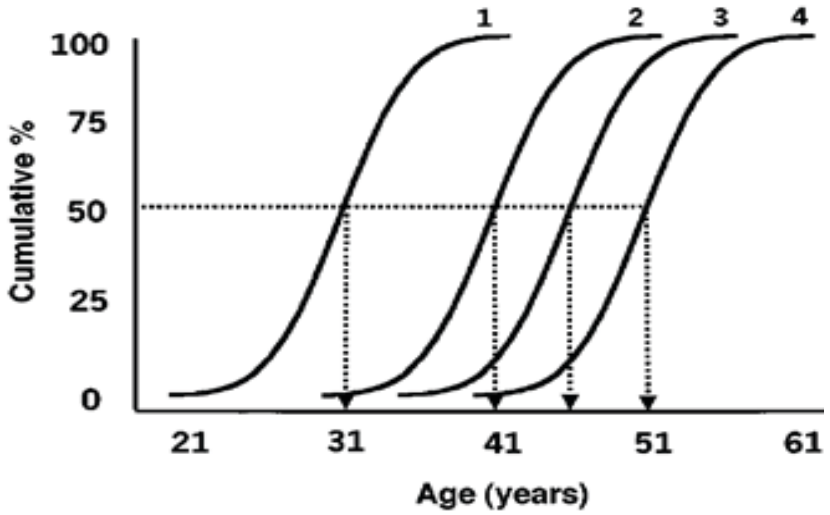


Figure 1.2: Age variations of the various stages of reproductive ageing.

Curve 1: variation in age at the beginning of subfertility (mean age 31 years); curve 2: variation in age at the occurrence of natural sterility (mean age 41 years); curve 3: variation in age at the transition from cycle regularity to irregularity (mean age 46 years); curve 4: variation in age at menopause (mean age 51 years); redrawn after te Velde and Pearson [4]

Moreover, the similarity between the length of the age interval and the shape of the distribution curve for age at natural menopause (ANM) on the one hand, and the end of natural fertility on the other hand, has led to the suggestion that a fixed temporal relation (of ~ ten years) is present between these two reproductive events.

Still, not many prospective follow up studies are available supporting this hypothesis. However, in the study by de Boer et al. [6], a relatively early age at menopause in women with a poor response in IVF was demonstrated, indicative of a low oocyte quantity and an early decline in fertility, an observation that is suggestive of such a relationship. Today's interference with natural fertility by the use of contraception and delay of motherhood jeopardizes any effort to further elaborate on this hypothesis. In a recent study by Daan et al. [7], it became apparent that in women with premature ovarian insufficiency, the interval between the final cessation of menses on the one hand and the loss of natural fertility and regular cycles on the other hand may be more compressed, thereby shedding some doubt as to the stability of the proposed concept. Nevertheless, the idea of a certain systematic time frame between reproductive events could allow for management of individual reproductive behavior based on forecasts of the only noticeable event, i.e. menopause.

Reproductive aging: quantity versus quality

When considering the loss of natural fertility and the definite end of the reproductive life (i.e. menopause) one should make a clear distinction between oocyte quality versus follicle quantity.
The reduction in natural fertility and presumably also the occurrence of subfertility (defined as not achieving an ongoing pregnancy within a one year period of exposure without any obvious explanation) stem from the deterioration of the quality of oocytes.
This is also reflected by an increase in aneuploidy rates resulting in higher rates of spontanous pregnancy loss [8]. The decline in follicle quantity on the other hand causes cycles to become increasingly irregular in the menopausal transition, ultimately leading to the final menstrual period. From conception to menopause, a gradual decline in the number of primordial follicles, also known as the true ovarian reserve, is demonstrated to be present (Figure 1.3) [9]. Menopause is believed to occur when the number of primordial follicles in the human ovary has fallen below a critical threshold [9]. At this threshold, the ovary becomes insufficiently capable of maintaining the production of mature oocytes within the framework of a menstrual cycle, and subsequently menopause occurs. Consequently, the endowed follicle pool at birth is considered to be the main determining factor in the timing of menopause. This theory was further supported in the study by Bjelland et al. [10]. In this paper, a time to menopause analysis was performed using data from 23580 women of whom 1055 underwent surgical oophorectomy for various reasons. Women that had undergone unilateral oophorectomy experienced menopause earlier when compared to women without surgery.
Although the effect was modest, menopause occurred one year earlier in the surgery group, it was statistically significant. These findings demonstrate that the remaining pool size somehow is of influence as a determining factor in the timing of menopause.
However, the limited effect of losing half of the ovarian reserve at some stage of premenopausal life indicates that other regulatory mechanisms must also be in play.

The role for menopause prediction in determining the fertile lifespan

Menopause is a relatively clearly marked event for both women and clinicians, even though it is defined in hindsight one year after the final menstrual period.

It might be suggested that the self-reporting of age at menopause makes this event susceptible to recall bias. It was however demonstrated by den Tonkelaar et al. [11] and by Colditz et al. [12] that both the validity and reproducibility of self-reported age at menopause are quite good.

In contrast, there is no demarcation for the end of natural fertility, or declining fertility, and therefore, this event occurs unnoticed. Effective contraceptives and the fact that more women have joined the workforce, has resulted in postponement of childbearing in Western countries in recent years [13]. Considering the theory of the fixed interval between menopause and the loss of natural fertility and the fact that some women may reach menopause already at age forty, one can postulate that some women may begin to lose natural fertility as early as age twenty. It is therefore imaginable that postponement of childbearing has subsequently resulted in an increasing number of women facing age related infertility at the time they hope to conceive, which is often in their early thirties. Such women are then indicated to have a short fertile lifespan. Fertile lifespan could be defined as the life period in which reproduction principally is possible, normally ranging from age 14 until age 41, which would comprise a 27 year range. This range could however greatly vary, from 17 to 37 years, assuming that the variation in menarche is clearly less wide than for menopause. It is needless to say that, based on the concepts laid down in Figure 1.2, optimal fertile lifespan could then even be much shorter.

If a proper outline of the individual fertile lifespan was to be available, women at risk for (relatively) early loss of natural fertility could be identified. They, in turn, could then prevent involuntary childlessness through timely family planning or through other recently developed technologies, such as cryopreservation of oocytes. Considering the above mentioned broad age interval during which the end of natural fertility occurs, female age alone does not provide sufficient information for adequate family planning.

Currently however, and as previously stated, there is no marker available predicting the end of natural fertility or narrowing this end of fertility window. Nonetheless, increasing evidence has become available stating that ovarian reserve tests (ORTs) might be capable of predicting menopause [14-24]. Consequently, in the prediction of the end of natural fertility, use has been made of the previously explained fixed temporal relation between menopause and fertility loss. Extrapolating the individual prediction of age at menopause could therefore provide a more precise estimation of the individual fertile life span. This information in turn could then be used in family planning in order to prevent involuntary childlessness.

The true ovarian reserve, natural fertility and the occurrence of menopause

As previously mentioned, from conception to menopause a gradual decline in the number of primordial follicles is hypothesized to be present (Figure 1.3) [9]. Based on this theory of gradual decline, the number of primordial follicles present at a certain moment could be used to predict age at natural menopause. Furthermore, and based on the presumed fixed interval between menopause and the end of natural fertility, the true ovarian reserve could be used to predict the end of natural fertility. Since the definite loss of natural fertility is preceded by a decline in fertility, use of the true ovarian reserve could also be made to predict current or actual female fertility.

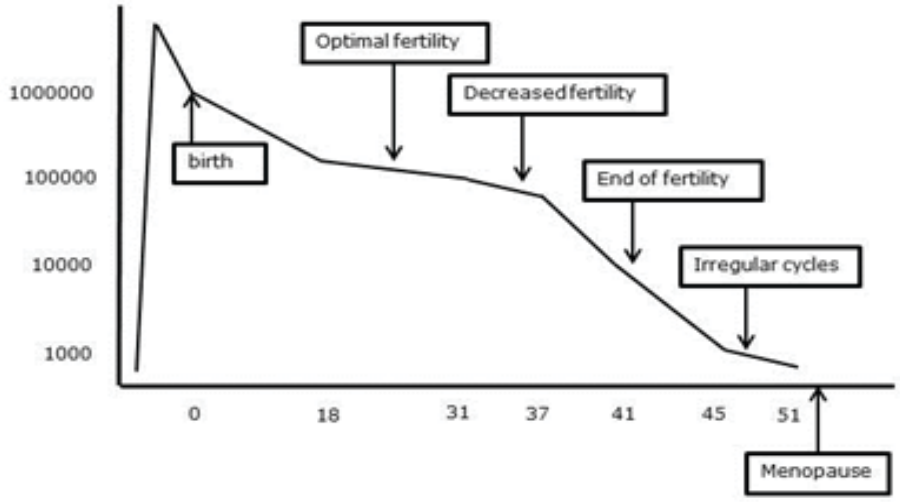


Figure 1.3: The relation between the true ovarian reserve and reproductive events. The Y-axis represents the number of primordial follicles, the X-axis represents age. Redrawn after Faddy et. al [25]

The true ovarian reserve and current fertility

As previously explained, the decline in the number of primordial follicles is accompanied by a decline in the quality of the oocytes held within the follicles. Currently, there is no marker available assessing the quality of the oocyte, which is exposed to possible fertilization on a monthly basis. Therefore, and comparable to the use that has been made of the marker "menopause" in the end of natural fertility, ovarian reserve tests are used to predict current fertility. This is an understandable concept, after all, if the decline in the number of

primordial follicles and the decline in oocyte quality go hand in hand, ORTs reflective of the number of primordial follicles could potentially reflect the oocyte quality.

However and as extensively demonstrated in studies assessing ORT based prediction of embryo or oocyte quality in IVF cycles, the relation between ORTs en oocyte quality is not easy to grasp [26-28]. Specifically, the occurrence of a live birth after IVF can be hardly predicted by any prior information.

In contrast, female age does have the capacity to inform on the chances to produce a live birth from IVF, and ORTS do add some information to such prognostication, emphasizing the complex interrelation between quantity and quality in the ovarian ageing process. To date, the possible role for any ORT, but specifically anti-Müllerian hormone (AMH), to allow for information on a woman's current fertility is under quite some debate, as fertility clinics have here and there started to offer AMH as a test for informing on fertility, both current and future, to young women without an actual wish to start a family. The question here should be posted, whether such practice is scientifically well based, and as such will be addressed in this thesis.

Derivatives of the true ovarian reserve: Ovarian Reserve Tests

Since one cannot remove an ovary to quantify the primordial follicle pool, it is not possible to perform an in vivo assessment of the relationship between the true ovarian reserve and age at natural menopause. Research did however demonstrate the number of antral follicles present in the ovary at a given time to be reflective of the primordial follicle pool [29]. Therefore, markers reflective of the number of antral follicles, the so-called ovarian reserve tests (ORTs), have been extensively reviewed as to their capacity to predict menopause, and thus (the end of) natural fertility. The relation between the true ovarian reserve and ORTs is depicted in Figure 1.4.

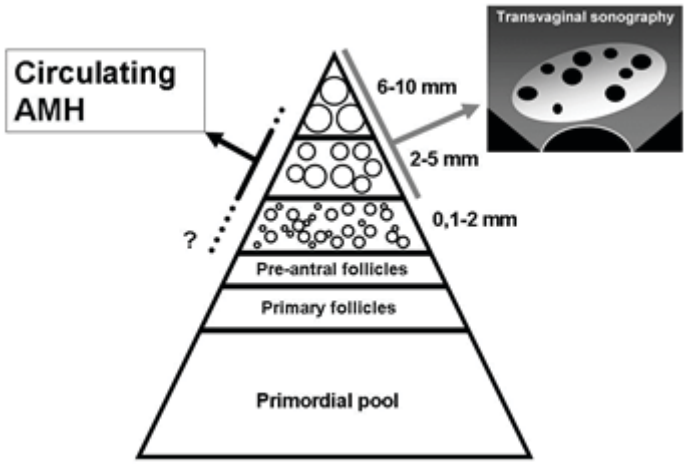


Figure 1.4 the relation between the primordial follicle pool and ovarian reserve tests reflective of this true ovarian reserve.

Antral Follicle Count

The Antral Follicle Count (AFC), as conducted via transvaginal ultrasound, is a measurement of all ovarian follicles ranging 2-10 (or 2-8) mm in size. The sum of the counts in the left and right ovary constitutes the AFC. The number of ovarian follicles has been demonstrated to represent the true ovarian reserve well in histologically counted ovarian specimens [30].

Moreover, it is mainly the number of relatively small antral follicles (also mainly responsible for the production of AMH, see below) that declines with increasing age [31]. The documented inter-observer reproducibility of the antral follicle count was demonstrated to be quite good, indicating that potentially this marker may offer promising clinical applications. Still, rigorous quality control on the execution of this test may seem more troublesome compared to most laboratory tests [32,33].

Anti-Müllerian hormone

Serum AMH levels stem from granulosa cells of ovarian antral follicles.

Whilst in the embryonic phase the absence of AMH results in female differentiation, the presence of AMH later in life (first detected around birth) results in inhibition of growth of resting ovarian follicles [28]. This mechanism in concert with other local, paracrine regulation factors, regulates the tempo of primordial follicles entering the pool of growing follicles, and thus is likely to determine the exhaustion of the primordial follicle pool and thus the occurrence of menopause, together with the size of the pool at the time of endowment of

the follicles into the ovaries in development [28].

The main source for circulating anti- Müllerian hormone is believed to be the pool of ovarian follicles ranging 2-8 mm in size [34,35]. Moreover, in another study, a positive and steady correlation was observed between levels of circulating AMH and small antral follicles and no correlation was observed between AMH and follicles greater than 12mm in size [36].

The role of AMH has been extensively researched as to its capacity to predict ovarian response in assisted reproduction technology. Moreover, data is becoming available assessing the capacity of AMH to predict menopause and current fertility.

Follicle Stimulating Hormone

Levels of Follicle Stimulating Hormone (FSH), one of the pituitary hormones, rise with the depletion of the ovary. FSH has been demonstrated to accurately reflect the current reproductive state, and is therefore still used in the classification of the current phases of reproductive ageing [37].

Other ovarian reserve tests

Several other ovarian reserve tests have been researched as to their capacity to reflect the number of antral follicles present in the ovary.

However, these tests (such as Inhibin B, the Gonadotrophin releasing hormone Agonist Stimulation Test, Exogenous FSH ORT, basal estradiol, or the Clomiphene Citrate Challenge Test) have all been proven to be inferior when compared to AMH, FSH or the AFC [38].

Aims and outline of the thesis

The studies presented in this thesis evaluate some basic concepts of ovarian reserve tests by assessing the role of the primordial follicle pool in menopause prediction and assessing the origin of fluctuation of AMH. Moreover, the use of ORTs in the prediction of menopause and of spontaneous ongoing pregnancy was researched.

Since AMH and the AFC have been proven to be the most promising ORTs, emphasis on these two tests has been placed in this thesis.

The objectives of this thesis can be listed as follows:

1. To investigate some basic concepts of ovarian reserve testing.
2. To determine the role for ovarian reserve tests in the prediction of the reproductive lifespan.
3. To examine the role for ovarian reserve tests in the assessment of current fertility.

Aim 1

Chapter 2: Studies the accuracy of the primordial follicle pool in the prediction of age at menopause.

Chapter 3: Studies the origin of fluctuations in circulating levels of AMH throughout the menstrual cycle and relates these fluctuations to fluctuations in the antral follicle count.

Aim 2

Chapter 4: Provides a systematic review of literature of existing evidence regarding the prediction of age at menopause (or time to menopause), using AMH, the antral follicle count, or mother's age at natural menopause.

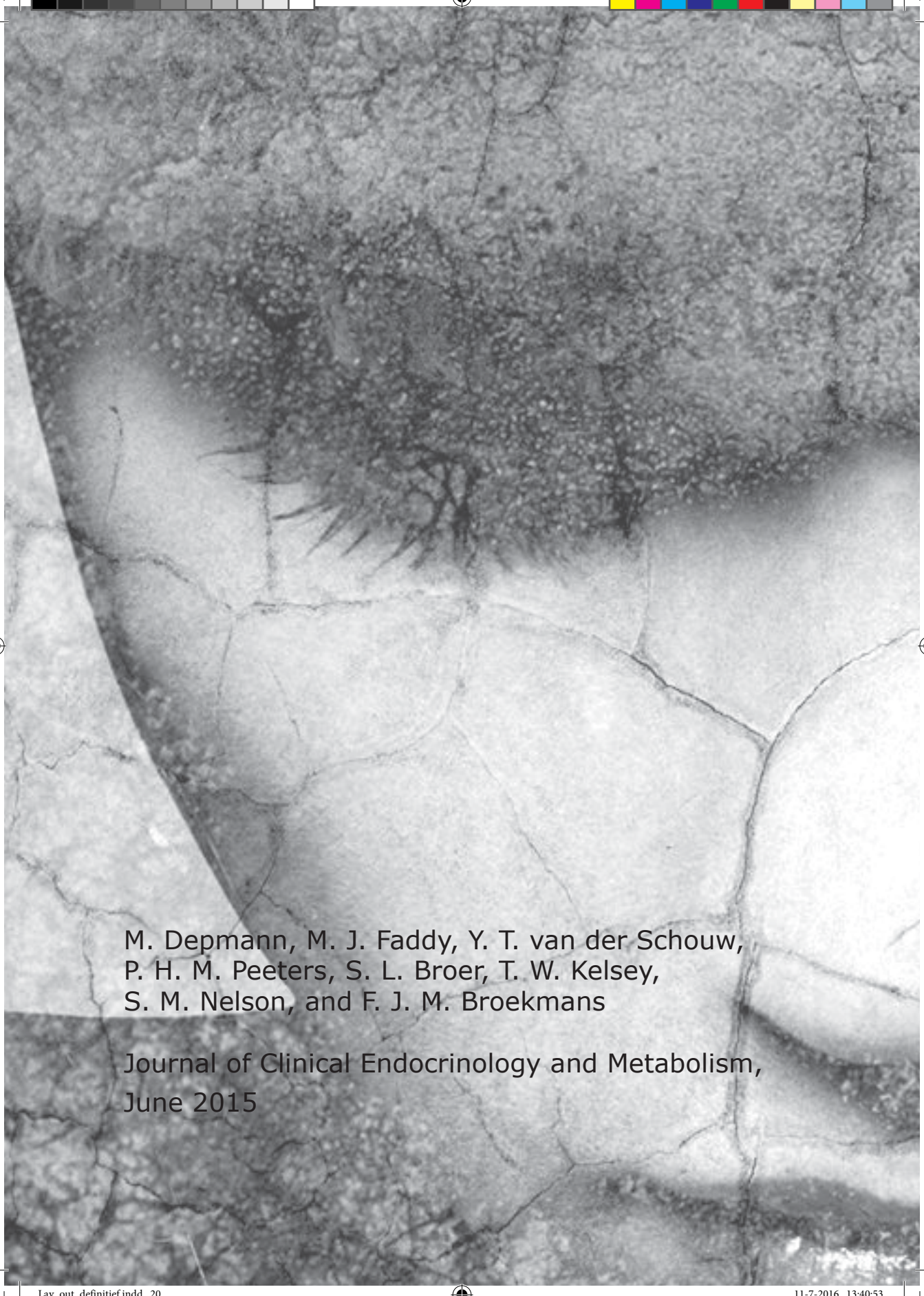
Chapter 5: Describes the results of a long- term follow up study conducted in healthy women researching the predictive capacity of ovarian reserve tests in menopause prediction.

Chapter 6: Describes the results of an Individual Patient Data Meta- Analysis performed researching AMH based menopause prediction.

Aim 3

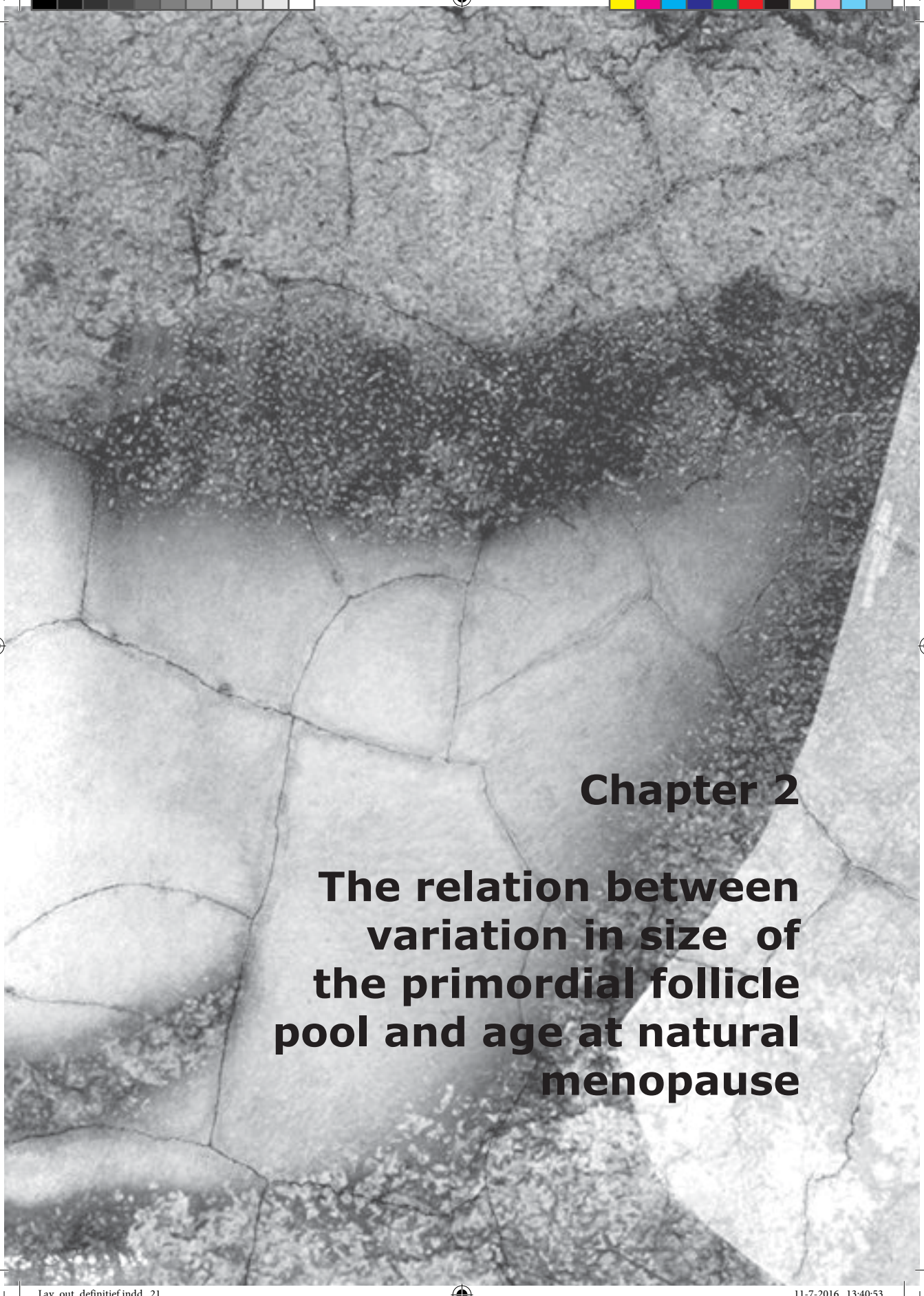
Chapter 7: Provides the results of a prospective cohort study researching ovarian reserve test based prediction of spontaneous ongoing pregnancy in a cohort of healthy pregnancy women.

Chapter 8: Summarizes and discusses the results of the studies presented.



M. Depmann, M. J. Faddy, Y. T. van der Schouw, P. H. M. Peeters, S. L. Broer, T. W. Kelsey, S. M. Nelson, and F. J. M. Broekmans

Journal of Clinical Endocrinology and Metabolism, June 2015



Chapter 2

The relation between variation in size of the primordial follicle pool and age at natural menopause

Abstract

Context: Menopause has been hypothesized to occur when the non-growing follicle (NGF) number falls below a critical threshold. Age at natural menopause (ANM) can be predicted using NGF numbers and this threshold. These predictions support the use of ovarian reserve tests, reflective of the ovarian follicle pool, in menopause forecasting.

Objective: To investigate the hypothesis that age-specific NGF numbers reflect age at natural menopause.

Design and Setting: Histologically derived NGF numbers obtained from published literature (n=218) and distribution of menopausal ages derived from the population based Prospect-EPIC cohort (n=4037) were combined.

Participants: NGF data were from single ovaries that had been obtained post-natally for various reasons, such as elective surgery or autopsy. From the Prospect-EPIC cohort, women aged 58 years and older with a known ANM were selected.

Interventions: None

Main Outcome Measure(s): Conformity between observed age at menopause in the Prospect-EPIC cohort and NGF-predicted age at menopause from a model for age-related NGF decline constructed using a robust regression analysis. A critical threshold for NGF number was estimated by comparing the probability distribution of age at which NGF numbers fall below this threshold with the observed distribution of ANM from the Prospect-EPIC cohort.

Results: The distributions of observed age at natural menopause and predicted age at natural menopause showed close conformity.

Conclusion: The close conformity observed between NGF-predicted and actual age at natural menopause supports the hypothesis that that the size of the primordial follicle pool is an important determinant for the length of the individual ovarian lifespan and supports the concept of menopause prediction using ovarian reserve tests, such as anti-Müllerian hormone and antral follicle count, as derivatives of the true ovarian reserve.

Introduction

From conception to menopause, a dynamic decline in the primordial follicle pool occurs within the human ovary. It is postulated that the final menstrual period coincides with a decline of the follicle pool below a critical threshold [25,39-41]. At this threshold the ovary becomes insufficient for maintaining the production of mature oocytes within the framework of a menstrual cycle and menopause, defined as the cessation of menstrual cycles, occurs. Various models have been designed to determine the rate of decline of the follicle pool with advancing age, as well as to quantify the critical threshold for cycle cessation, however no consensus has been reached thus far [25,39-41].

In recent years research focused on providing methods for individualized predictions of the age at natural menopause (ANM) [18,21,42]. It has been postulated that these predictions could be extended to predict the end of natural fertility [5,43,44] and be used to identify women at greater risk of cardiovascular and neurological disease [45,46], osteoporosis [47], or breast and intestinal cancer [48] due to either early or late menopause.

Individual predictions of ANM have been based on markers indirectly reflecting the size of the primordial follicle pool such as anti-Müllerian hormone (AMH). [16,18,19,21,22] However, model imprecision limits their clinical utility at the current time. The biological basis for these prediction models is the assumption that the size of the primordial follicle pool present at birth or later on in life is the determining factor in age at menopause, and that ovarian reserve tests may accurately capture this pool size at any age. [30] The association between the primordial follicle pool number and age at natural menopause has not been examined in detail, and follow up studies will remain elusive.

The aim of this current paper is therefore to investigate the hypothesis that the size of the primordial follicle pool is the main determinant for the length of the individual ovarian lifespan. This is done by modelling the decline in numbers of follicles with increasing age. A model for age at menopause is then based on the number of primordial follicles declining below a critical threshold for cycle cessation. The distribution of predicted age at menopause from this modelling is compared with observed data on age at menopause. Demonstrating a clear link between the 'true' ovarian reserve and age at natural menopause could then provide some rationale for the use of ovarian reserve tests in clinical practice.

Materials and Methods

Subjects

Two databases were used to investigate the above formulated aim. For the purpose of creating predictions of age at menopause based on primordial follicle numbers, the non-growing follicle (NGF) database, as considered by Wallace et al. [9], was used to estimate a model for the age dependent decline in the number of primordial follicles. In this database the histologically derived number of NGFs in human ovaries of eight different cohorts [39,41,49-54] were combined to form one large database. From this database, a selection of cases was made for use in the present analysis.

First, a selection was made based upon age at specimen collection, with only data from ovaries obtained post-natally being used. Furthermore, in cases where a single person provided two ovaries, the mean number of NGFs from the both ovaries was used in order to prevent overrepresentation of these cases. This was in accord with results of Hansen et al [41], were no significant difference was apparent in the number of NGFs between left and right ovaries collected from a single person. Considering the fact that it is not practically possible to determine the amount of time an ovary has been void of follicles, providing an age at the time the ovary reached a NGF count of zero is not possible. Furthermore, it is not possible to determine that an ovary is actually devoid of follicles. For these reasons, it was decided to exclude zero counts rather than make arbitrary adjustments to these and other low counts as was done with the data used by Wallace et al. [9]. After applying these criteria, 218 cases provided NGF counts from the equivalent of a single ovary, with the original low counts being used without any adjustment.

The second database, providing the observed distribution of ages at natural menopause, is the Prospect-European Prospective Investigation into Cancer and Nutrition (Prospect-EPIC) database. A total of 17357 women, aged 50 to 70 years, were recruited between 1993-1997 from a nationwide breast cancer screening program in the Netherlands [55,56]. Via extensive questionnaires menopausal status as well as reproductive health was assessed. Menopause was defined according to the World Health Organization definition: the absence of spontaneous menstrual bleeding for more than 12 months. For the current study, a cross-sectional sample of women aged over 58 years at study recruitment with a recorded natural menopause was selected in order to prevent overrepresentation of women who reach menopause at an early age. Including only women of age over 58 years will ascertain that the full

normal range of menopausal ages has been recorded. Furthermore, exclusion criteria were the use of medication interfering with menstrual cycles, ovarian abnormalities or surgery, and surgery on the uterus prohibiting menstrual cycles. Applying the above mentioned selection criteria reduced the total of available participants to 4037.

Analyses

The age dependent decline in NGF number was modelled using a robust regression methodology [57], as previously applied [58]. A natural logarithmic transformation of the NGF counts was used to stabilize the residual variation in NGF numbers. The regression model for NGF counts contains linear and quadratic (in age) components, giving it considerable generality in allowing for accelerating rates of exponential loss of follicles with increasing age (negative quadratic term) and other possibilities such as different rates of follicle loss between subjects (positive quadratic term), as well as being similar to other modelling. [25,57] Furthermore, the more general (than normal) skew-t distribution permits a more critical assessment of the residual variation of NGF counts about the estimated mean.

Under the hypothesis that menopause occurs when the number of NGFs falls below a critical threshold, this regression based model for NGF decline enabled the construction of a probability distribution for age at menopause for a given threshold to be derived from the following relationship:

Probability that NGF count at age *y* is below threshold
= probability that menopause has occurred before age *y*.

For this derivation of a distribution of ages at menopause, the residual standard deviation in the regression model (for NGFs) was adjusted to allow for the possibility of excess variation in NGF counts compared with variation in menopausal ages, as was necessary in the previous modelling of AMH and menopause [58].

This gives linked models for both age at menopause and declining NGF count with increasing age. Maximum likelihood estimation was then used with the Prospect-EPIC menopausal age data and the NGF count data to estimate all model parameters, including the critical threshold NGF number.

Finally, individual prediction of age at menopause could be made from a nomogram of age specific NGF percentile bands constructed from the estimated NGF regression model (i.e. a very low age specific NGF count would place a

woman below the 5th percentile, whereas a high age specific NGF count would place her in a higher percentile band). Prediction of age at menopause then follows from the corresponding percentiles of the estimated distribution of menopausal ages.
All calculations were done using MATLAB® numerical software.

Results

As described, the selection criteria reduced the size of the NGF database from 325 cases to 218 (Table 2.1), and the size of the Prospect-EPIC cohort from 17357 to 4037 women.
Details of all the studies in the original NGF database, and the number of remaining specimens in the adjusted NGF data after applying our selection criteria are depicted in Table 2.1.
The follicle counts in the NGF data ranged between 9 and 402018 per ovary. The age range of subjects at the time of obtaining the ovary specimens was 0 to 51 years, with more details given in supplementary Table 2.1. The mean age at menopause in the Prospect-EPIC data was 50.2 years (AD 4.2 years), mean age at inclusion was 63.05 years (SD 3.4 years).
Textbox 2.1 shows all parameter estimates from the model fitting.

Regression model for NGF count and age:
mean of log(NGF count): 12.29(0.16) + 0.020(0.015) × age – 0.0028(0.00029) × age2
standard deviation of log (NGF count): exp{–0.45(0.12) + 0.019(0.0042) × age}

Derived distribution of menopausal ages:
critical NGF count threshold: 498(88)
standard deviation of log (NGF count): exp{–0.81(0.19) + 0.019(0.0042) × age}

Both models:

skewness of log(NGF count) residual distribution: –0.33(0.037)

Textbox 2.1: Maximum likelihood parameter estimates (with standard errors in brackets) of components of the model for declining NGF count with increasing age and the derived distribution of menopausal ages from the NGF count falling below a critical threshold.

Author (year)	N 2010	Age at count	Orirgin of specimens	Counting method	N 2015	Reason for exclusion
Block (1952)	86	6-44y	Accidental death, suicide, acute illness/ infection	Model based	43	2 ovaries per woman
Block (1953)	19	Fetal	Abortion	Model based	0	Fetal specimens
Baker (1963)	11	Fetal	Spontaneous abortion, unknown	Beaumont & Mandl	0	Fetal specimens
Richardson (1987)	9[a]	45-51y	Elective surgery	Model based	13	Women after menopause
Gougeon (1987)	52	19-50y	Surgery for various conditions	Linthern-Moore	43	2 ovaries per woman
Bendsen (2006)	11	Fetal	Abortion	Fractional/ optical dissector	0	Fetal specimens
Forabosco (2007)	15	Fetal-1y	Abortion, maternal disease	Caveilleri	1	Fetal specimens
Hansen (2008)	122	0-51y	Oophorectomy, donor	Fractional/ optical dissector	118	Fetal specimens /NGF count 0
Total	325				218	

Table 2.1: Outline of Case Selection from the NGF Database

N 2010 is the number of specimens included in the original database; N 2015 is the equivalent number of single ovaries of cases aged 16 years or older included in the current analyses. [a] Original data referred to more specimens than the 2010 database.

Figure 2.1 demonstrates the fit of the quadratic regression model of log(NGF count) on age to the NGF count data, showing the reducing mean and increasing residual standard deviation with increasing age. The mean function of age from the earlier NGF analysis [9] is shown for comparison.

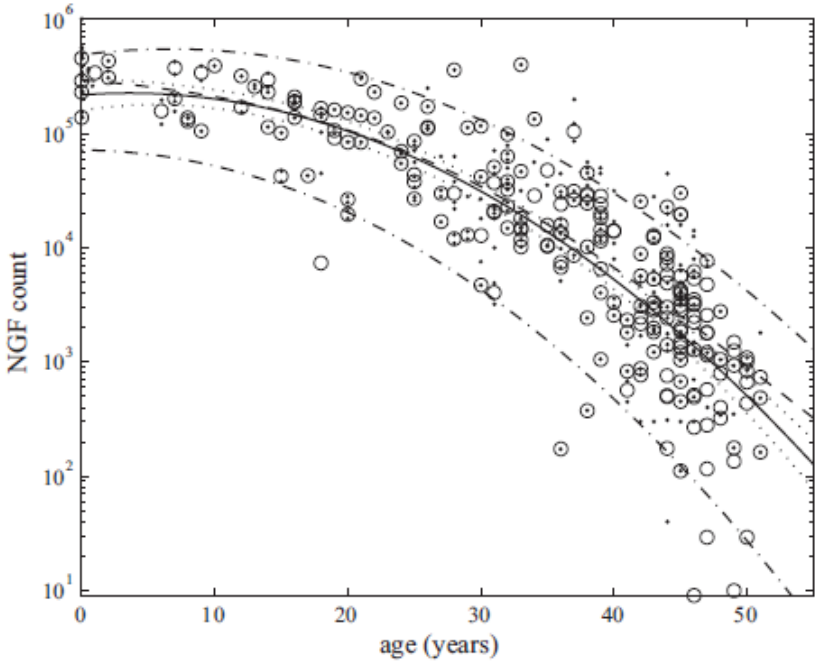


Figure 2.1 NGF counts (circles) from the 2015 data and from the 2010 data (dots). The quadratic regression (solid line) fitted to the 2015 data with 95% confidence intervals (dotted lines) and 90% prediction intervals (dashed dotted lines). The Wallace-Kelsey model as fitted to the 2010 data is also depicted (dashed line).

The two mean functions are similar for ages up to 45 years (with the Wallace et al.[9] mean within 95% confidence limits of the quadratic regression), but diverge somewhat after that age due to the different treatment of low counts in the data (R^2 values 0.72 from the current analysis and 0.70 from the 2010 analysis). There was significant left-skewness in the distribution of the residual variation from the regression of log(NGF count) on age (p-value < 0.001), suggesting that these residuals were not normally distributed (i.e. more observations above the estimated mean than below it). The negative quadratic component of the regression was significant (p-value < 0.001), corresponding to accelerated decline in later years. There was also a significant (p-value < 0.001) increase in the residual standard deviation of log(NGF count) with increasing age.

This shows that the logarithmic transformation had overcompensated for the heterogeneous variation in the untransformed NGF data (demonstrated in Figure 2.1 by the gradual broadening of the 90% probability range with increasing age).

The estimate of the value of the critical threshold of number of non-growing follicles from a single ovary for the occurrence of menopause was 498 NGFs.
Figure 2.2 shows the fit of the derived distribution of menopausal ages, predicted by the number of non-growing follicles falling below this critical threshold, compared with the observed menopausal ages in the Prospect-EPIC data.

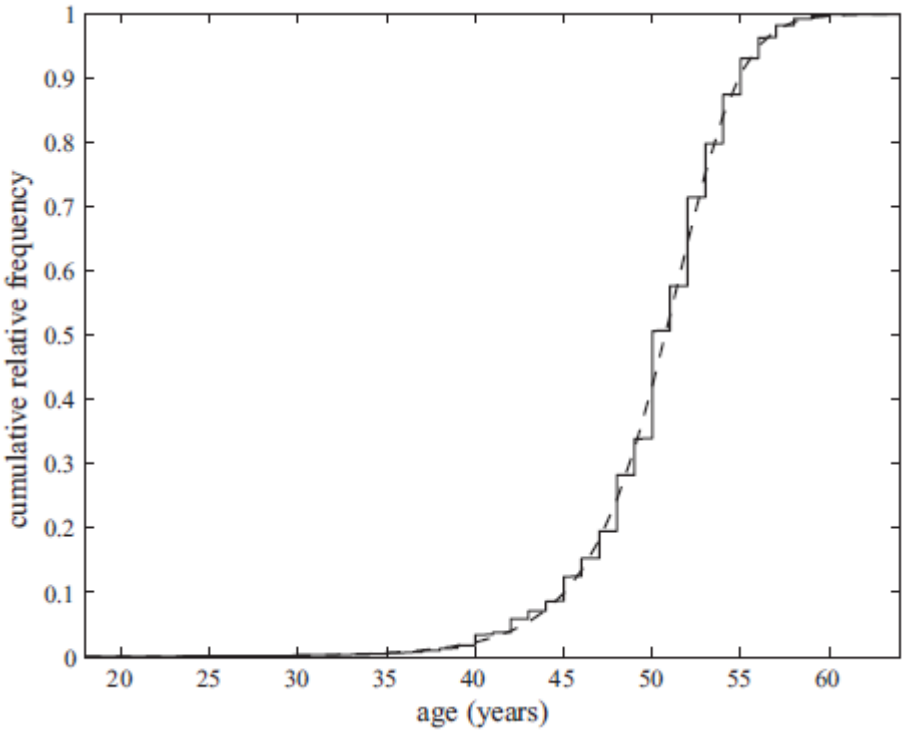


Figure 2.2 Observed distribution of menopausal ages from the Prospect-EPIC data (solid line, n=4037) compared with the distribution (dashed line, n=218) derived from NGF count falling below a critical threshold, in which close agreement is apparent.

Good agreement was observed between observed and predicted ages at menopause. There was significantly less variation (p-value < 0.001) needed to explain the variation in menopausal ages than that apparent from the regression analysis of log(NGF count) on age, with the residual standard deviation from the regression analysis reduced by an estimated factor of 0.70.
Figure 2.3 shows a nomogram in which the 5th, 10th, 25th, 50th, 75th, 90th and 95th age-specific percentiles of NGF counts are depicted.
The estimated critical threshold in NGFs at which menopause occurs is also indicated in this figure (498 follicles, see Textbox 2.1).

Furthermore, the corresponding percentiles of the distribution of age at menopause derived from the NGF count are shown adjacent to the NGF percentiles.

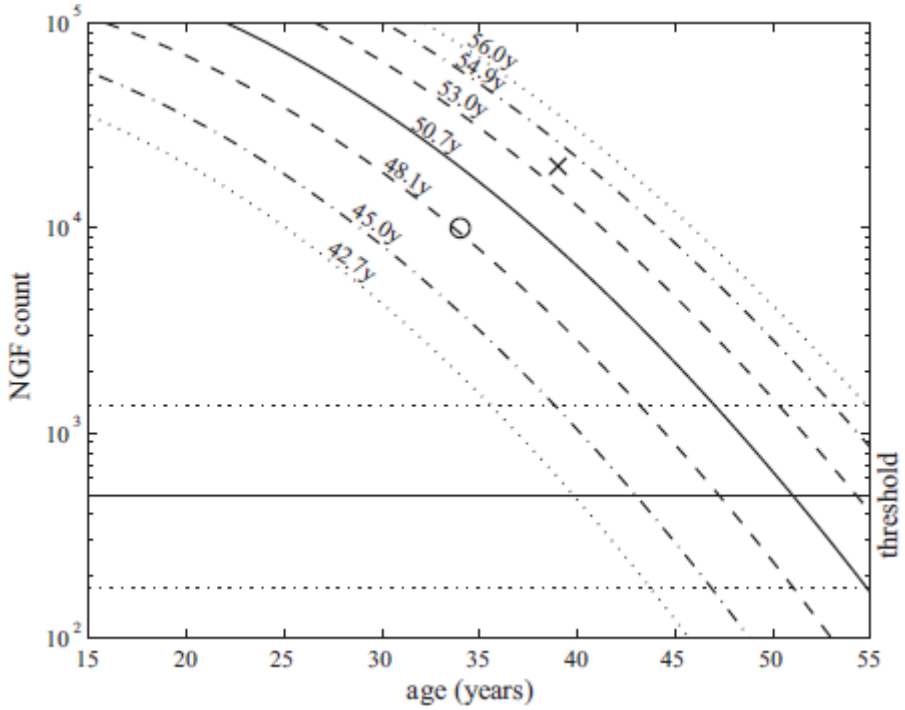


Figure 2.3 Nomogram showing 5% and 95% (dotted lines), 10% and 90% (dashed dotted lines), 25% and 75% (dashed lines), and 50% or median (solid line) age-specific quantiles for NGF counts from the fitted regression model. The corresponding percentiles of age at menopause derived from NGF count falling below a critical threshold are shown adjacent to these quantiles. The critical threshold (498 NGFs) is depicted by the faint solid horizontal line, whereas the faint dotted lines indicate an approximate interquartile range when the threshold is allowed to vary. As depicted in this figure, a 34-year-old woman with 10 000 NGFs is destined to experience menopause at approximately 48.1 years (circle), whereas a 39-year-old woman with an NGF count of 20 000 will become menopausal at an age between 53.0 and 54.9 years (plus sign).

Discussion

In this study we have demonstrated that the predicted age at menopause based on the decline of the primordial follicle pool shows close conformity with the observed age at menopause. This close conformity supports the hypothesis that the size of the primordial follicle pool is an important determinant for the length of the individual ovarian lifespan. These results strengthen the current

interest in the prediction of age at menopause using markers reflecting the ovarian pool of primordial follicles such as AMH or AFC.

This study is the first to assess the relation between predicted age at menopause based on the non-growing follicle pool and the observed age at menopause. Wallace et al. [9] did model NGF decline in relation to age using the original NGF dataset. A NGF threshold of 1000 follicles was used for the occurrence of menopause assuming a mean age at menopause of 51 years.

The threshold of 498 NGFs present in a single ovary estimated in the current analysis is significantly lower (p-value <0.001) than the threshold of 1000 follicles used in this previous publication. The difference here most likely stems from our sharpened inclusion criteria for the ovarian specimens in cases providing low NGF counts or NGF counts of zero, which were set at an artificial minimum higher than the estimated NGF count in the data used by Wallace et al [9]. In the current analysis, NGF counts of 0 were excluded but low NGF counts were kept at their original numbers. This resulted in the inclusion of a subset of lower NGF counts, thereby pulling down the mean NGF count. This in turn resulted in a lower estimate of the NGF threshold at which menopause is expected to occur. The effect of the inclusion of lower NGF counts is demonstrated in Figure 2.1 were the mean function of age is depicted for both the current model and the 2010 model [9]. As depicted, the two mean functions are similar up to 45 years of age. However after this age the lines diverge due to the lower NGF count pulling down the mean in the current analysis. We feel that by including NGF counts at their original number, the threshold as currently calculated more accurately represents the "true threshold".

In order to get good agreement between the observed and predicted ages at menopause (as depicted in Figure 2.2), the residual standard deviation from the regression model of log(NGF count) had to be reduced by a factor of 0.70 when constructing the NGF predictive distribution of age at menopause. This excess variation in NGF could be due to differences in NGF counting methods. Block et al. [49] counted non- growing follicles in one slice of ovarian tissue per 200 slices with a slice thickness of 20-40μm. Richardson et al. [39] counted one slice per 100 with a slice thickness of 10μm whereas Gougeon et al. [52] counted an unknown number of slices with a thickness of 10μm. Forabosco et al. [54] consecutively sliced each ovary in one 1-μm thick and 10 100-μm thick slices and counted every 1μm thick slice, lastly, Hansen et al. [41] sliced the total ovary into slabs of 1mm, selected 8 of these slabs to slice them further into slices of 25 μm thickness and eventually counted follicles in one in every 10 of these slices. Another potential source of variation in NGF number is the difference

in counting accuracy that could have originated from technical developments. In the first papers included in the NGF data [39,49,52] manual counting methods were used, whereas in the papers by Forabosco et al. [54] and Hansen et al. [41] an automatic counting method was applied. However, since the first paper in the NGF database was published in 1951 and the most recent in 2008, the occurrence of some variation in counting methods is almost inevitable.

The effect of excess residual variation in NGF counts on prediction of menopausal ages can be incorporated in the modelling by having a variable critical threshold at which menopause occurs that is positively correlated with NGF counts. This would result in a higher critical threshold for women with high NGF counts and a lower threshold for women with low NGF counts. This is indicated in Figure 2.3 by an inter-quartile range for this variation which additional calculations suggest could be as much as 175 - 1357 non-growing follicles.

A possible limitation of this study is the fact that data from two different datasets were combined in order to estimate the level of agreement between the observed and predicted ages at menopause. The Prospect-EPIC study is a population-based sample of healthy women, and is likely to represent the distribution of age at menopause of a Caucasian population. Unfortunately, no baseline characteristics are available for the NGF dataset, hence limiting the possibility of checking the comparability of the two datasets. Virtually all the ovarian specimens, although macroscopically described as normal, were derived post mortem or during gynecological surgery. As for the cases in the post mortem group, causes for demise may be considered as not affecting ovarian reserve. However, the group from which ovaries were obtained due to gynecological conditions, may introduce some bias.

To investigate if reasons for ovarian removal affected the model estimates, a sensitivity analysis was performed with the specimens categorized as being obtained from post-mortem cases (n=88) when death had occurred "suddenly" (for instance in a motor vehicle accident), from women providing specimens after gynecological surgery (n=129) or for "other" reasons (n=1). This sensitivity analysis was done by deleting subsets of data categorized according to reason for ovarian removal from the full dataset and then re-doing of the parameter estimation (note that afterwards, the deleted data were replaced before deletion of the next subset of data). The variation between the different estimates so obtained was assessed and found to be a little more than might be expected from random deletions (supplementary Figure 2.1). There is some confounding here in that the second (post-surgery) category contained a preponderance of older ages, while a preponderance of younger ages was

apparent in the first (post mortem) category, which would contribute to some of these differences. A good level of qualitative similarity was apparent among the different estimates. Menopausal threshold estimates were comparable, as was the significant left-skewness in the distribution of (log) NGF residuals.

Also, the increasing standard deviation of these residuals with increasing age was consistently apparent, as was the observation that there was more variation in NGF counts than necessary to describe variation in menopausal ages. This sensitivity analysis suggests that the NGF data are sufficiently homogenous without any obvious biases to justify the linking of these data with the Prospect-EPIC data.

Another possible limitation of this study is the fact that data on age at menopause were based on self- reporting, making this prone to recall bias, as for many included cases the life year in which menopause occurred had been a long time ago. However, a sufficient validity and reproducibility of estimating age at menopause based on these questionnaires has been reported in several studies. [11,12] In one study comparing age at menopause derived from repeated questionnaires with a 7-9 year interval, the agreement between initial and delayed recordings was shown to be high. [11]

The strength of the current study is the fact that the model provides some unique evidence regarding the prediction of age at menopause from what is generally considered to be the true ovarian reserve.

Due to the inherent necessity to remove an ovary to quantify the primordial follicle pool, it is not feasible to assess, in-vivo, the relationship between the true ovarian reserve and natural age at menopause. Hence, models have been designed to predict age at natural menopause using derivatives of this true ovarian reserve such as AMH or the AFC, or proxy variables such as oocyte yield after ovarian stimulation for IVF. [14-18,22,42,58-61] Yet no research has been designed to investigate the assumption made in these models that the size of the ovarian follicle pool is indeed the determining factor in the occurrence of menopause.

We do not see clinical utility for the NGF threshold/nomogram in the prediction of menopause. With the results of the present study, demonstrating the link between the size of the follicle pool and age at menopause, we have provided some further rationale for the use of derivatives of the true ovarian reserve in menopause prediction. Predictions that can potentially be used in the identification of women at risk of early menopause.

Conclusion

This study has shown that the age dependent decline in the number of non-growing follicles can be used to construct a distribution of age at natural menopause, which is close to the observed distribution.
This supports the hypothesis that the size of the ovarian follicle pool is an indicator for age at menopause. It thereby provides support for the concept of prediction of age at menopause using markers reflective of the true ovarian reserve, such as AMH or the AFC.

Supplementary material

Specimen source	Publication date	Age	NGF Count single ovary	Medical condition for ovary removal	Age group
Hansen	2008	0	458250	perinatal asphyxia	0
Hansen	2008	0	138880	diaphragmatic hernia	0
Hansen	2008	0	232031	congenital heart failure	0
Hansen	2008	0	290669,5	sepsis	0
Forabosco	2007	1	342095	cranial trauma	0
Hansen	2008	2	308235	asthma attack	0
Hansen	2008	2	434375	congenital heart failure	0
Block	1952	6	158500	sudden death	0
Block	1952	7	202500	sudden death	0
Block	1952	7	377500	sudden death	0
Block	1952	8	129000	sudden death	0
Hansen	2008	8	137951	pulmonary hypertension	0
Block	1952	9	342000	sudden death	0
Hansen	2008	9	105154	head injury	0
Hansen	2008	10	393000	hydrocephaly	0
Block	1952	12	170500	sudden death	0
Hansen	2008	12	319282	asphyxiation	0
Block	1952	13	258500	sudden death	0
Block	1952	14	295500	sudden death	0
Hansen	2008	14	114678,5	motor vehicle accident	0
Hansen	2008	14	231103	motor vehicle accident	0
Block	1952	15	42500	sudden death	0
Hansen	2008	15	101356	drug overdose	0
Hansen	2008	16	208980	intracranial hemorrhage	1
Hansen	2008	16	138202	motor vehicle accident	1
Block	1952	16	187000	sudden death	1
Hansen	2008	16	191230	motor vehicle accident	1
Hansen	2008	17	42892	motor vehicle accident	1
Hansen	2008	18	167361,5	cerebrovascular accident	1
Block	1952	18	7400	sudden death	1
Block	1952	18	145000	sudden death	1
Hansen	2008	19	161721	motor vehicle accident	1
Block	1952	19	106500	sudden death	1

Gougeon	1987	19	91750	surgery for breast cancer, cervical cancer, fibroadenomas	1
Hansen	2008	20	153021	motor vehicle accident	1
Block	1952	20	26500	sudden death	1
Block	1952	20	19500	sudden death	1
Hansen	2008	20	84297	motor vehicle accident	1
Hansen	2008	21	144731	motor vehicle accident	1
Hansen	2008	21	304369,5	motor vehicle accident	1
Hansen	2008	21	84471,5	pelvic pain/ endometriosis	1
Hansen	2008	22	136974	seizure	1
Hansen	2008	22	231875	motor vehicle accident	1
Block	1952	23	103000	sudden death	1
Hansen	2008	24	185652	pelvic pain/ endometriosis	1
Hansen	2008	24	54986	pelvic pain/ endometriosis	1
Block	1952	24	69500	sudden death	1
Hansen	2008	25	86287	cerebrovascular accident	1
Block	1952	25	27000	sudden death	1
Block	1952	25	38000	sudden death	1
Gougeon	1987	25	43950	surgery for breast cancer, cervical cancer, fibroadenomas	1
Hansen	2008	26	111316	head trauma	1
Hansen	2008	26	172794	dysmenorrhea, fibroids	1
Block	1952	26	114000	sudden death	1
Hansen	2008	27	16970,5	pulmonary embolism	1
Block	1952	27	30000	sudden death	1
Hansen	2008	28	361991	motor vehicle accident	1
Block	1952	28	12000	sudden death	1
Block	1952	28	30000	sudden death	1
Hansen	2008	29	112637	pelvic pain/ endometriosis	1
Block	1952	29	13000	sudden death	1
Hansen	2008	30	4702,5	pulmonary embolism	1
Hansen	2008	30	117082	pelvic pain	1

Hansen	2008	30	41910	pelvic pain/ endometriosis	1
Block	1952	30	12800	sudden death	1
Hansen	2008	31	51106	pelvic pain/ endometriosis	2
Block	1952	31	20500	sudden death	2
Block	1952	31	21000	sudden death	2
Block	1952	31	4050	sudden death	2
Gougeon	1987	31	37150	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	32	98900,5	carotid artery aneurysm	2
Hansen	2008	32	32504	pulmonary embolism	2
Hansen	2008	32	14870	pelvic pain	2
Block	1952	32	23000	sudden death	2
Block	1952	32	38000	sudden death	2
Block	1952	32	63000	sudden death	2
Hansen	2008	33	11703,5	head trauma	2
Hansen	2008	33	10230	pelvic pain/ endometriosis	2
Hansen	2008	33	46624	cerebrovascular accident	2
Hansen	2008	33	18134	menorrhagia/ endometriosis	2
Hansen	2008	33	402018	motor vehicle accident	2
Block	1952	33	15000	sudden death	2
Block	1952	33	14500	sudden death	2
Hansen	2008	34	134158	pelvic pain/ adenomyosis	2
Gougeon	1987	34	28650	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	35	10553,5	pelvic pain/ endometriosis	2
Hansen	2008	35	10298	pelvic pain	2
Gougeon	1987	35	15700	surgery for breast cancer, cervical cancer, fibroadenomas	2

Gougeon	1987	35	47750	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	36	15896,5	pelvic pain/ endometriosis	2
Hansen	2008	36	173	cerebral aneurysm	2
Hansen	2008	36	30910	endometriosis	2
Hansen	2008	36	13388	endometriosis	2
Hansen	2008	36	6762	pelvic pain	2
Block	1952	36	7450	sudden death	2
Gougeon	1987	36	24150	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	36	13400	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	37	8590,5	motor vehicle accident	2
Hansen	2008	37	31396	dysmenorrhea, pelvic pain	2
Hansen	2008	37	26153	cerebrovascular accident	2
Block	1952	37	104000	sudden death	2
Hansen	2008	38	376	endometriosis, fibroids	2
Hansen	2008	38	10242	pelvic pain, fibroids	2
Hansen	2008	38	2432	pelvic pain/ endometriosis	2
Hansen	2008	38	45277	dysmenorrhea/ adenomyosis	2
Hansen	2008	38	28782	pelvic pain/ adenomyosis	2
Block	1952	38	29500	sudden death	2
Gougeon	1987	38	24650	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	39	4046	endometriosis/ adenomyosis	2
Hansen	2008	39	11500	menorrhagia	2
Hansen	2008	39	1058,5	pelvic pain/ adenomyosis	2

Hansen	2008	39	18315	pelvic pain/ endosalpingitis	2
Hansen	2008	39	6546,5	pelvic pain/ endometriosis	2
Hansen	2008	39	12353	pelvic pain/ adhesions	2
Hansen	2008	39	14555	cerebrovascular accident	2
Hansen	2008	39	19965	motor vehicle accident	2
Gougeon	1987	39	28400	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	39	24050	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	40	2563	cerebral aneurysm	2
Block	1952	40	14000	sudden death	2
Block	1952	40	3350	sudden death	2
Gougeon	1987	40	14150	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	41	1818	cerebral aneurysm	2
Hansen	2008	41	831	menorrhagia/ fibroids	2
Hansen	2008	41	2343,5	menorrhagia/ fibroids	2
Block	1952	41	565	sudden death	2
Hansen	2008	42	771,5	menorrhagia/ fibroids	2
Hansen	2008	42	5662	pelvic pain/ fibroids	2
Hansen	2008	42	873	menometrorrhagia/ fibroids	2
Hansen	2008	42	25511	pelvic pain/ menorrhagia	2
Hansen	2008	42	8828	menorrhagia/ fibroids	2
Hansen	2008	42	3084	menorrhagia/ fibroids	2
Block	1952	42	2450	sudden death	2
Gougeon	1987	42	2150	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	43	1973,5	pelvic pain/ endometriosis	2
Hansen	2008	43	1840	pelvic pain/ endometriosis	2

Hansen	2008	43	3323,5	dysmenorrhea/ fibroids	2
Hansen	2008	43	1224	menorrhagia	2
Hansen	2008	43	12256	menorrhagia/ fibroids	2
Block	1952	43	5300	sudden death	2
Block	1952	43	3000	sudden death	2
Gougeon	1987	43	12900	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	43	2825	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	44	1418,5	menorrhagia/ fibroids	2
Hansen	2008	44	8831	cardiac arrest	2
Hansen	2008	44	2430	menorrhagia/ adenomyosis	2
Hansen	2008	44	3039	hydrosalpinx, fibroids	2
Hansen	2008	44	4953	menorrhagia/ fibroids	2
Hansen	2008	44	22784	pelvic pain	2
Hansen	2008	44	7161	menorrhagia/ fibroids	2
Block	1952	44	175	sudden death	2
Gougeon	1987	44	8200	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	44	510	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	44	500	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	44	760	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	44	5800	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	45	3947	cerebrovascular accident	2
Hansen	2008	45	1040,5	menorrhagia	2
Hansen	2008	45	30210	endometrial hyperplasia	2
Hansen	2008	45	19491,5	menorrhagia/ fibroids	2

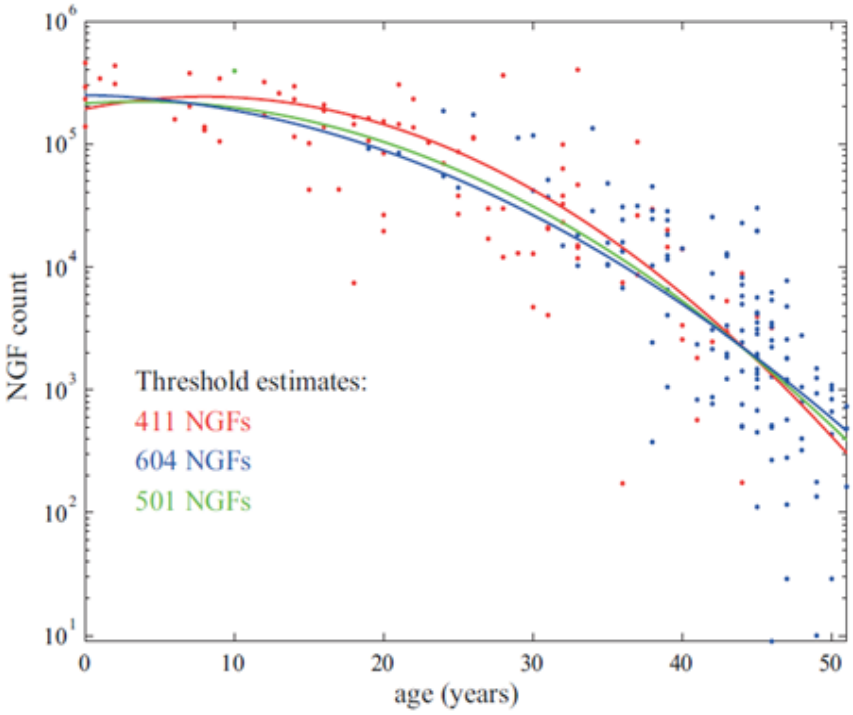
Hansen	2008	45	1219	pelvic pain/ endometriosis	2
Hansen	2008	45	111,5	menorrhagia/ fibroids	2
Hansen	2008	45	3535	pelvic pain/ fibroids	2
Hansen	2008	45	677	menorrhagia	2
Hansen	2008	45	453,5	menorrhagia/ fibroids	2
Hansen	2008	45	4157,5	menorrhagia/ fibroids	2
Hansen	2008	45	1842	menorrhagia/ fibroids	2
Hansen	2008	45	4255,5	pelvic pain	2
Richardson	1987	45	1958	elective hysterectomy/ oophorectomy	2
Richardson	1987	45	1472	elective hysterectomy/ oophorectomy	2
Gougeon	1987	45	3125	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	45	19650	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	45	2850	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	45	1350	surgery for breast cancer, cervical cancer, fibroadenomas	2
Gougeon	1987	45	5650	surgery for breast cancer, cervical cancer, fibroadenomas	2
Hansen	2008	46	3188,5	sepsis	3
Hansen	2008	46	1281	menorrhagia/ adenomyosis	3
Hansen	2008	46	494	menorrhagia/ fibroids	3
Richardson	1987	46	9	elective hysterectomy/ oophorectomy	3
Richardson	1987	46	515	elective hysterectomy/ oophorectomy	3
Richardson	1987	46	269	elective hysterectomy/ oophorectomy	3

Gougeon	1987	46	5375	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	46	3250	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	46	2225	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	46	3500	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	46	2525	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	46	6180	surgery for breast cancer, cervical cancer, fibroadenomas	3
Hansen	2008	47	7728	pelvic pain/ fibroids	3
Hansen	2008	47	1210	incontinence, fibroids	3
Richardson	1987	47	116	elective hysterectomy/ oophorectomy	3
Richardson	1987	47	2570	elective hysterectomy/ oophorectomy	3
Richardson	1987	47	29	elective hysterectomy/ oophorectomy	3
Gougeon	1987	47	570	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	47	1150	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	47	1800	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	47	1815	surgery for breast cancer, cervical cancer, fibroadenomas	3

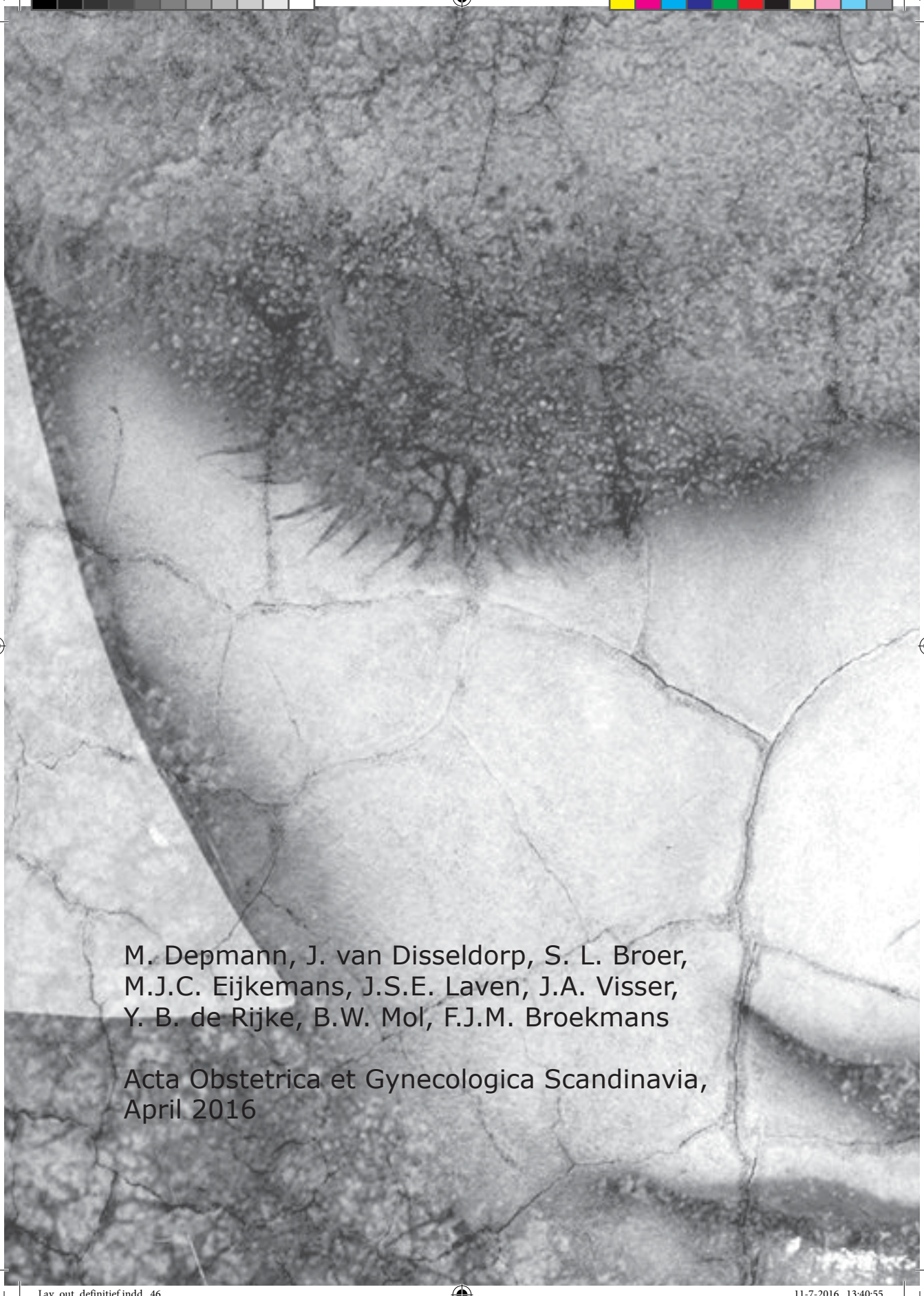
Gougeon	1987	47	280	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	47	4775	surgery for breast cancer, cervical cancer, fibroadenomas	3
Hansen	2008	48	324	menorrhagia/ fibroids	3
Hansen	2008	48	1053	menorrhagia/ fibroids	3
Hansen	2008	48	2770	pelvic pain/ menorrhagia	3
Gougeon	1987	48	400	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	48	795	surgery for breast cancer, cervical cancer, fibroadenomas	3
Hansen	2008	49	177	pelvic pain/ adenomyosis	3
Hansen	2008	49	933	menorrhagia/ fibroids	3
Richardson	1987	49	1499	elective hysterectomy/ oophorectomy	3
Richardson	1987	49	135	elective hysterectomy/ oophorectomy	3
Richardson	1987	49	10	elective hysterectomy/ oophorectomy	3
Gougeon	1987	49	1250	surgery for breast cancer, cervical cancer, fibroadenomas	3
Hansen	2008	50	1087,5	abnormal uterine bleeding	3
Hansen	2008	50	841	dysmenorrhea/ endometriosis	3
Richardson	1987	50	29	elective hysterectomy/ oophorectomy	3
Gougeon	1987	50	435	surgery for breast cancer, cervical cancer, fibroadenomas	3
Gougeon	1987	50	665	surgery for breast cancer, cervical cancer, fibroadenomas	3

Gougeon	1987	50	1005	surgery for breast cancer, cervical cancer, fibroadenomas	3
Hansen	2008	51	162,5	menorrhagia/ fibroids	3
Hansen	2008	51	484	menorrhagia/ adenomyosis	3
Richardson	1987	51	734	elective hysterectomy/ oophorectomy	3

Supplementary Table 2.1: Outline of the NGF dataset (Age group (as used in the subgroup analyses): 0=0-15, 1=16-30, 2=31-45, 3=46-60)

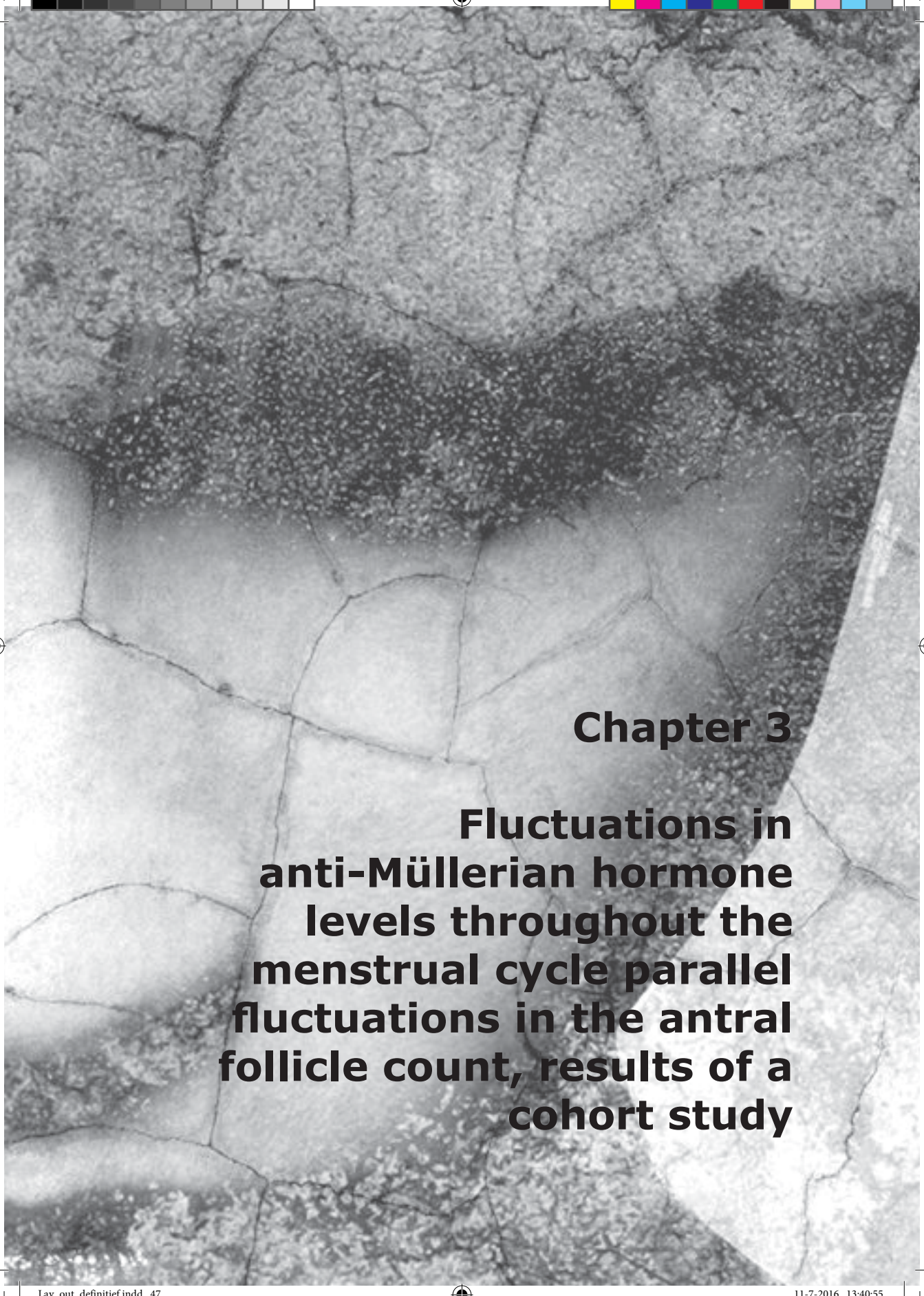


Supplementary Figure 2.1: NGF counts grouped by the different reasons for specimen collection: sudden death (red dots), gynecological surgery (blue dots) and other (green dot); the coloured curves represent the quadratic regressions fitted to data *without* those from the corresponding sub-group.



M. Depmann, J. van Disseldorp, S. L. Broer, M.J.C. Eijkemans, J.S.E. Laven, J.A. Visser, Y. B. de Rijke, B.W. Mol, F.J.M. Broekmans

Acta Obstetrica et Gynecologica Scandinavia, April 2016



Chapter 3

Fluctuations in anti-Müllerian hormone levels throughout the menstrual cycle parallel fluctuations in the antral follicle count, results of a cohort study

Abstract

Introduction: In this prospective cohort study we aimed to investigate the hypothesis that fluctuations in anti-Müllerian hormone (AMH) levels stem from fluctuations in the number of antral follicles (AFC).

Material & methods: Repeated measurements of AMH and AFC (follicles 2-8 mm) were performed in 44 women with a regular cycle, during one menstrual cycle. If our hypothesis that AMH fluctuations stem from fluctuations in the AFC is correct, a fluctuation in the number of antral follicles would result in an equal and parallel shift in AMH. Hence, the difference between AFC and AMH would remain constant over time.

A mixed model analysis (MMA), assessing the stability between AMH and AFC, was performed using the difference between $_{log}$AFC and $_{log}$AMH. Cohen's D was calculated for the largest of fixed effects in order to assess stability in relative distance between AFC and AMH. To assess if fluctuation in AMH or AFC originated from between subject fluctuation, or from within subject fluctuation, the Intra Class Correlation Coefficient (ICC) was calculated.

Results: MMA and Cohen's D (0.12) confirmed the stability of the difference between $_{log}$AFC and $_{log}$AMH and thus confirmed our hypothesis. The good ICC (0.73) indicated a small contribution of within subject variation to AMH fluctuations.

Conclusion: Fluctuations in AMH levels parallel fluctuations in AFC, suggesting that AMH levels are closely linked to variation in the AFC. This knowledge adds to the basic understanding of the origin of AMH and could aid in interpretation of individual AMH levels.

Introduction

In recent years research became available regarding anti-Müllerian hormone (AMH) and its capacity in predicting ovarian response to controlled ovarian hyperstimulation [62], the age at natural menopause [15,16,18,19,21,22,42,58,60], the prediction of the end of natural fertility [18], and its potential to aid in the diagnosis of polycystic ovarian syndrome [63].

Reports on the stability of AMH levels throughout and between menstrual cycles indicated that AMH does not follow any classical endocrine cycle pattern and remains stable over time [64-66]. Other studies however, emphasized AMH levels to fluctuate considerably, depending on reproductive stage [67], female age category [68] or cycle phase [69]. Moreover, AMH levels in general are influenced by body mass index (BMI), smoking, genetic factors, polymorphisms of AMH and the AMH receptor, vitamin D status or ethnicity [28]. Furthermore, AMH fluctuations may result from analytical variation, sample instability during storage, or complement binding [28].

Peripheral AMH levels stem from granulosa cells of ovarian antral follicles. Therefore, an important potential factor causing fluctuations in AMH levels within a menstrual cycle is the fluctuation in the number of antral follicles. The intra cycle variability of the antral follicle count (AFC) was documented by van Disseldorp et al. [66] In this paper a 34% variance in AFC within one menstrual cycle was described. Since the majority of circulating AMH originates from follicles ranging up to 8 mm in size [34,35], one could postulate that AMH levels shift parallel to fluctuation in AFC (2-8mm) during the menstrual cycle. In spite of extensive research on the intra- and intercycle variability of both AMH and AFC [64-66,68], no attempt to link these fluctuations within the menstrual cycle has been published.

The aim of this current paper is to investigate the hypothesis that fluctuations in AMH levels across the menstrual cycle originate from fluctuations in the AFC. This information could add to the basic understanding of the origin of AMH and can potentially aid in the clinical framework of interpreting individual AMH levels.

Material and Methods

Participants

In 1996 and 1997 healthy female candidates were recruited via newspaper advertising. Women were considered eligible when experiencing a regular

menstrual cycle (21-35 days), when having a biphasic temperature chart and if proven natural fertile (defined as having experienced at least one term pregnancy and conceiving this pregnancy within one year of trying to conceive). Exclusion criteria were evidence of endocrine disease and ovarian surgery or – abnormalities on ultrasound. Hormonal contraception needed to be discontinued at least two months prior to inclusion. A full description of the cohort can be obtained in previous publications from this study group [64,68,70].
Institutional Review Board approval was obtained in the UMC Utrecht in The Netherlands (Reference number U-99-062, date of approval April 26, 1999). Written informed consent was obtained from all participants. A monetary compensation was received for participation.

Study design

Women were asked to fill out a temperature chart. One menstrual cycle was evaluated for analysis from the start of the luteal phase of the first cycle, to the peri-ovulatory phase of the subsequent cycle. The start of the luteal phase was pinpointed on the day of the temperature rise on the temperature chart. Visits to the clinic were planned during the early-, and late luteal phase of the first cycle, and in the second cycle during the early-, mid-, and late follicular phase and the peri-ovulatory phase. Visits were scheduled at an interval of 2-3 days initially and daily when the dominant follicle had reached a size of 14 millimeters. The last measurements were performed on the day the ovulation occurred, this event was defined as the day of complete disappearance or a reduction in size of the dominant follicle of at least 5mm (Figure 3.1).
At every visit, a transvaginal ultrasound was performed for measurement of the antral follicle count. All measurements were performed by a single observer (G.J.S.) using a 7.5-MHz transvaginal probe on a Toshiba Capasee SSA-220A (Toshiba Medical Systems Europe BV, Zoetermeer, the Netherlands). The ovary was scanned from outer to inner margin, follicle diameter was calculated from two or three perpendicular measurements depending on the size of the follicle (≤6 or >6mm). All follicles ranging 2-8 mm were counted in each ovary, the sum of both counts constituted the AFC. The documented inter-observer reproducibility of the antral follicle count in the Utrecht University Medical Center is strong (mean difference= -0.46, p-value=0.23) [33]. During the same visit, a blood sample was drawn for AMH assay.

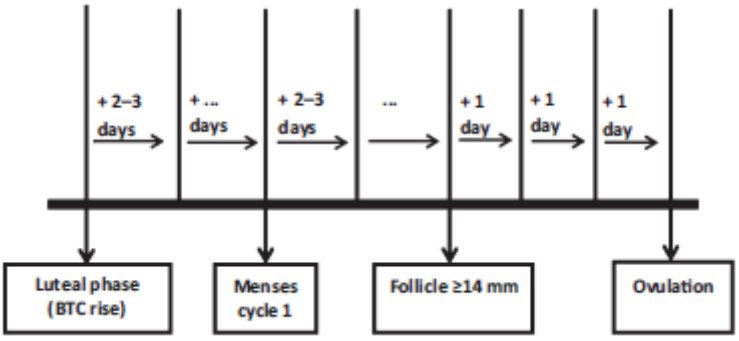


Figure 3.1: Schedule of AMH and AFC measurements

Each vertical bar represents 1 measurement; 6 cycle phases were defined as follows: early follicular= first day of menstruation to menstruation +4 days, mid follicular= day of ovulation -9 to day of ovulation -6, late follicular= day of ovulation -5 to day of ovulation -2, peri ovulatory= day of ovulation-1 to day of ovulation +1, early luteal: day of ovulation +2 to day of menstruation -7, late luteal: day of menstruation -6 to day of menstruation -1; Measurements were divided per cycle phase as follows: early follicular N=43, mid follicular N=31, late follicular N=36, peri- ovulatory N=44, early luteal N=12, and late luteal N=39.

Hormone assay

Serum and plasma were separated within 5 hours of sampling and stored at -20C until processed. AMH levels were measured using the enzyme-immunometric assay (Diagnostic System Laboratories, Webster, TX) during a single assay run. Inter- and intra-assay coefficients of variation were less than 5% at the level of 3ug/L and less than 11% at the level of 13ug/L. The limit of detection for AMH was 0.026ug/L. Repeated freezing and thawing of the samples or storage at 37 C for 1 h did not affect results of the assay.[71]

Analyses

Measurements were performed at fixed time intervals independent of cycle duration and measurement interval was adjusted according to the size of the dominant follicle. Therefore, women could provide more than one measurement per cycle phase. In order to correct for this, a mean AMH or AFC value per cycle phase was calculated for women providing these multiple measurements in a single cycle phase.

A mixed model analysis was performed to test the hypothesis that a correlation between fluctuations in AMH levels and fluctuations in the AFC exists. This mixed model analysis was performed on the difference between AMH and AFC, since

the distance between these variables remains constant when a fluctuation in the AFC is followed by an equal and parallel fluctuation in peripheral AMH levels. First, and in order to correct for differences in scale and distribution of AMH and the AFC, standardizing of $_{log}$AMH and $_{log}$AFC was performed by calculating the Z scores for both parameters per cycle phase. Since a Z-Score is a statistical measurement of a variable's relationship to the mean in a group of observations (a Z-score of 0 for instance indicates that the variable measurement is zero times the SD different from the mean), the influence of scale or distribution of a variable is undone when using Z-scores. Next, the difference between Z-score(AFC) and Z-score(AMH) per cycle phase was calculated. As previously described, this difference remains constant throughout the menstrual cycle, if AFC and AMH fluctuations are concordant. A statistically significant difference as outcome in this mixed model analysis would therefore imply that none or only incomplete concordance exists between AMH and AFC values over time.

While a mixed model analysis can only describe the presence or absence of fluctuation, it cannot describe the possible origin of fluctuation. Fluctuation can originate either from variation between subjects, or from variation within a subject. Variation between subjects is comprehensible, since the distance between AMH and the AFC can vary between women, but fluctuation within a subject would conflict with our hypothesis. In order to detect the origin of fluctuation, the Intra Class Correlation Coefficient (ICC) was calculated. The ICC is defined as the ratio of the between subjects variance to the total (=between + within subjects) variance. As such, a high ICC would imply that the distance between AMH and the AFC for a given case remains stable throughout the menstrual cycle.

For additional testing on the stability of the difference between Z-score(AFC) and Z-score(AMH) throughout the menstrual cycle, the mean differences between Z-score(AFC) and Z-score(AMH) per cycle phase were compared. Cohen's D, indicating the standardized difference in means, was calculated for the largest difference present between two cycle phases. A lower Cohen's D will therefore reflect a more stable relation between AFC and AMH. A value for the Cohen's D between 0.2 and 0.4 was considered small, between 0.5 and 0.7 medium, and ≥ 0.8 as large. [72]

Finally, since BMI and smoking have been suggested to influence the levels of circulating AMH [73], we performed a subgroup analysis comparing smokers to non- smokers and comparing women with a normal BMI (<25 kg/m^2) to women with an elevated BMI (≥25 kg/m^2) by incorporating these variables in the mixed model analysis.

For visualization purposes, individual AMH and AFC fluctuations were depicted and box plots were constructed depicting the fluctuations throughout the menstrual cycle in AMH and AFC values as well as the fluctuations in the difference between Z-score(AFC) and Z-score(AMH) as used in the mixed model analysis (displaying the stability in the distance between AMH and AFC) for the entire cohort.
Variables of interest were checked on normal distribution. Mixed model residuals were graphically depicted in a QQ-plot, in order to allow for testing on models assumptions by graphically reviewing the linearity. Data analysis was performed using SPSS version 21 (SPSS Inc., Chicago, IL) and R version 3.0.3 (http://www.r-project.org).

Results

44 women aged 25-46 years (mean 37.66 years) were included in this study. Women provided 5-14 AFC and AMH measurements, depending on cycle length and size of the dominant follicle, during one full menstrual cycle, this resulted in 396 data points. Variables were checked on normal distribution, where appropriate standardization was applied. Moreover, models residuals were depicted in a QQ-plot and the linearity assumption was met. After calculating a mean value for the AFC and AMH for those women providing more than one data point per cycle phase, 205 data points remained available. AMH levels ranged from 0.1-3.10 ng/ml (median 0.48 ng/ml) and AFC ranged from 0-37 follicles (median 5 follicles). Individual AMH and AFC fluctuations are depicted in Figures 3.2 and 3.3.

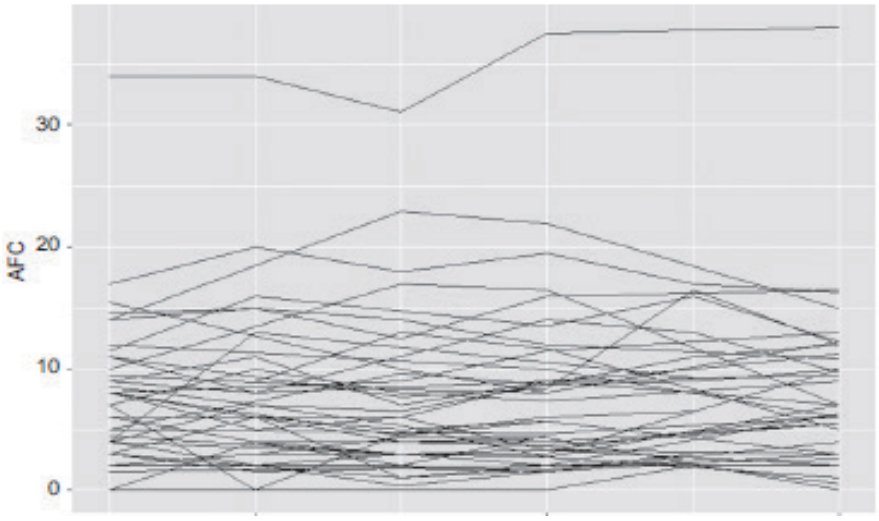


Figure 3.2: Fluctuations in the number of antral follicles per individual. X-axis represents cyclephase.

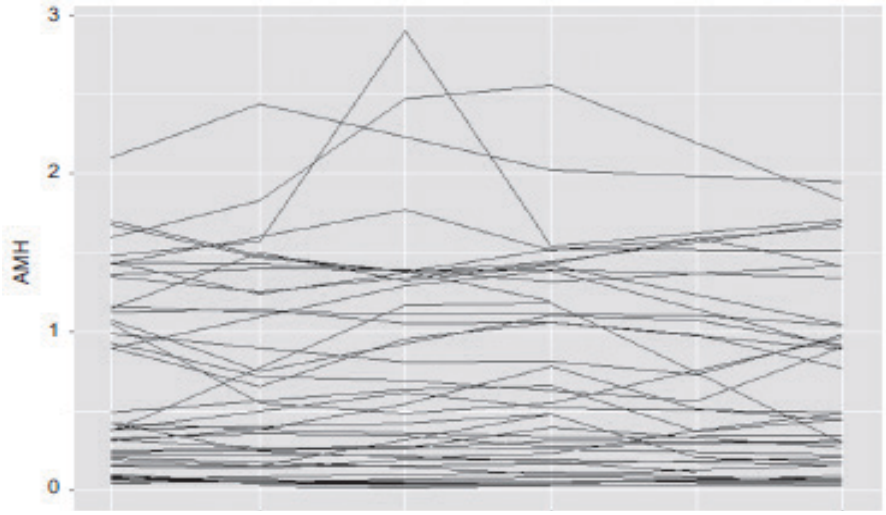


Figure 3.3: Fluctuations in AMH levels per individual. X-axis represents cyclephase

In Figures 3.4-3.6 box plots showing crude AMH and AFC values, and the difference between Z-score(AFC) and Z-score(AMH) per cycle phase are depicted.

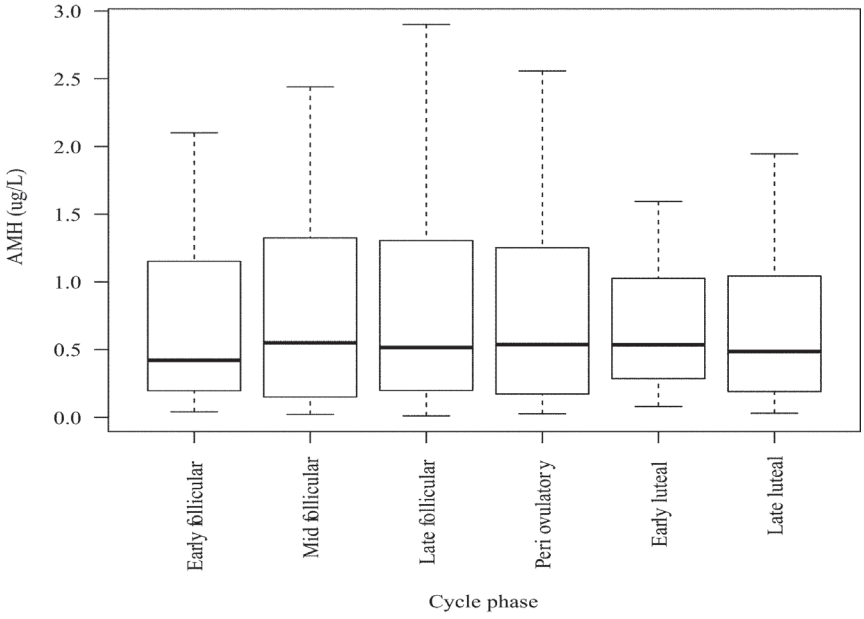


Figure 3.4: Boxplot showing AMH values throughout the menstrual cycle

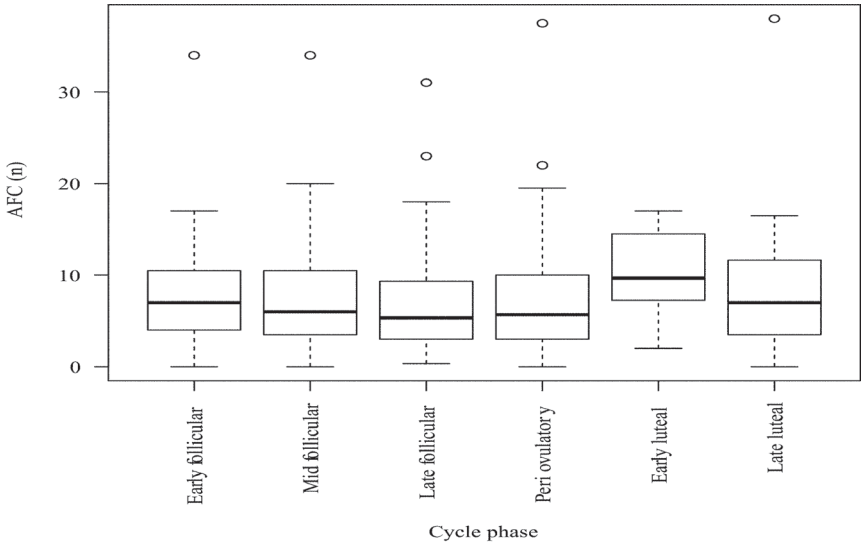


Figure 3.5: Boxplot showing AFC values throughout the menstrual cycle

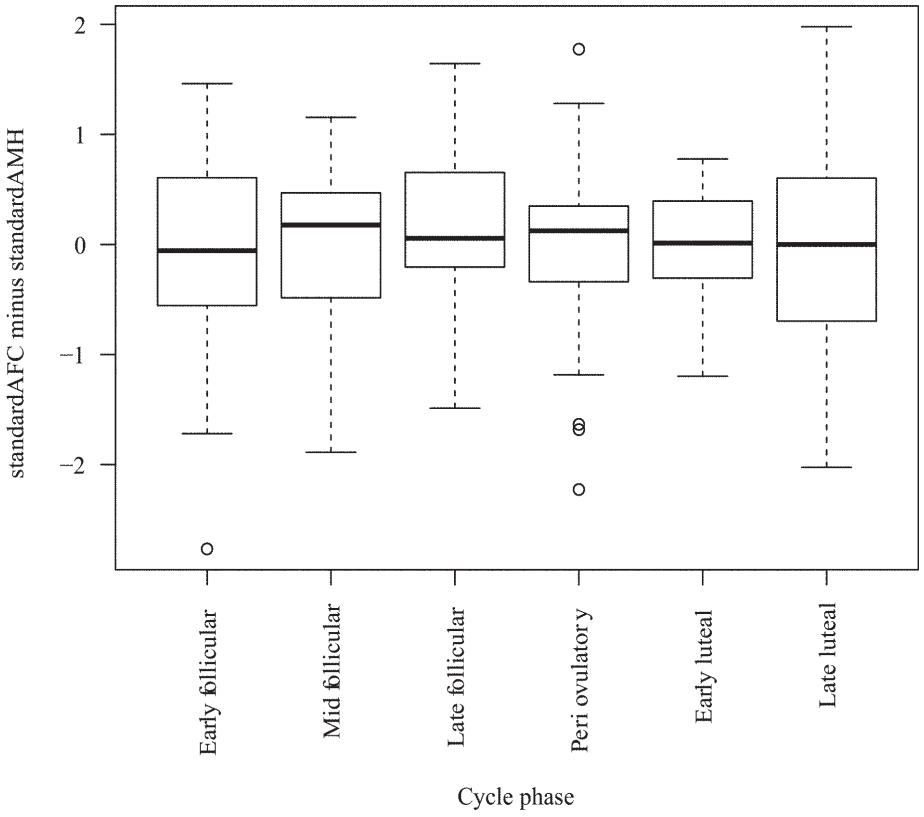


Figure 3.6: Boxplot showing the value of the difference between Z-score(AFC) and Z-score(AMH) throughout the menstrual cycle. The Z-Score is a statistical measurement of a variable's relationship to the mean in a group of observations (a Z-score of 0 for instance indicates that the variable measurement is zero times the SD different from the mean)

MMA calculating the difference between Z-score(AFC) and Z-score(AMH) did not show a statistically significant variation (p=0.91) throughout the menstrual cycle. The ICC, depicting the origin of fluctuation in the difference between Z-score(AFC) and Z-score(AMH) was good (0.73).

The stability throughout the different cycle phases was furthermore reflected by the very small and non-significant effect size, with a Cohen's D of 0.12. MMA comparing smokers (n=16) to non- smokers (n=13) revealed that the parallelism between AMH and the AFC did not depend on smoking status (p=0.85). MMA incorporating the influence of BMI (<25 n=21 or ≥25 n=12) did not alter the results (p=0.07) from the original analysis.

Discussion

The current study revealed that fluctuation in peripheral AMH levels throughout a menstrual cycle parallels fluctuation in AFC. This implies that much of the short term variation in circulating AMH is caused by changes in the number of antral follicles, specifically follicles sized 2-8 mm. As stated, the non- significant results of the MMA, together with the small value of Cohen's D and the good ICC, reflect the firm stability in the relative distance between the AFC and AMH over time. This confirms our hypothesis that any fluctuation in the number of antral follicles is accompanied by a proportional shift in AMH levels.

The stability of AMH throughout the menstrual cycle has been extensively researched. However, research centered on the correlation between AFC and AMH is scarce. The parallelism between the AFC ranging 2-8 mm and peripheral AMH levels as observed in the current paper corroborates the notion made in the paper by Jeppesen et al. [34]. In this study it was stated that the main source for circulating levels of AMH is the pool of antral follicles ranging 2-8mm. These conclusions were based on measurements of AMH in both serum and follicular fluid which were compared to the size and number of antral follicles. This observation is in line with analyses performed by Fanchin et al. [36] where a positive and steady correlation was observed between peripheral AMH levels and small antral follicles (3-11mm), but no correlation was observed between AMH levels and follicles ≥12mm. Much in the same line, Weenen et al [35] performed AMH staining of human ovary tissue obtained after oophorectomy and demonstrated AMH staining to become undetectable in follicles larger than 8mm.

A previous analysis by the present research group [66] concentrated on the fluctuation of AMH and the AFC across the early follicular phase of four consecutive menstrual cycles and noted a positive association between AMH and the AFC. Streuli et al. [65] performed analysis to assess concordance of AMH and AFC on the second day of the menstrual cycle, but failed to show any correlation. The cohort consisted of 24 women of whom 14 were using hormonal contraceptives. The use of oral contraceptives is known to result in lower age specific AMH levels [73], a potential consequence of the suppression of FSH levels [74].

Additionally, Fanchin et al. [75] and Eldar- Geva et al. [76] reported on the parallelism between AMH and AFC. In both papers, a strong relation between AFC and AMH levels measured on the third or fourth day of the menstrual cycle was observed. The current paper demonstrated this correlation to be present

throughout the entire menstrual cycle. Streuli et al. [65], Fanchin et al .[75] and Eldar-Geva et al. [76] did not provide data allowing for repeated measurement analysis, as such the current data have further refined our knowledge on the source of the AMH fluctuation.

It might be suggested that not all short term variation in peripheral AMH levels is caused by changes in the number of antral follicles sized 2-8mm. Variation in repeatedly measured AMH could also origin from assay dependent variation [28]. However, since all AMH measurements were performed in a single laboratory using a single assay under the same conditions, the assay specific influence on AMH fluctuation is expected to be minimal.

Furthermore, individual AMH levels or individual AFC in general might be influenced by genetic or lifestyle factors such as smoking and BMI. Reports demonstrate that smoking results in lower age specific AMH levels [73]. Our data allowed for a subgroup analysis investigating the influence of smoking on the level of parallelism between AMH and AFC but showed no effect. It is possible however that the reduced sample size in this subgroup analysis (from 44 to 29 women), resulted in a lack of power to detect any effect. The effect of BMI on peripheral AMH levels and the AFC is subject to research as well, though no consensus could be found in papers regarding this subject [73,77]. The subgroup analysis incorporating BMI in the mixed model analysis showed no effect of this variable on the parallelism between the AFC and AMH.

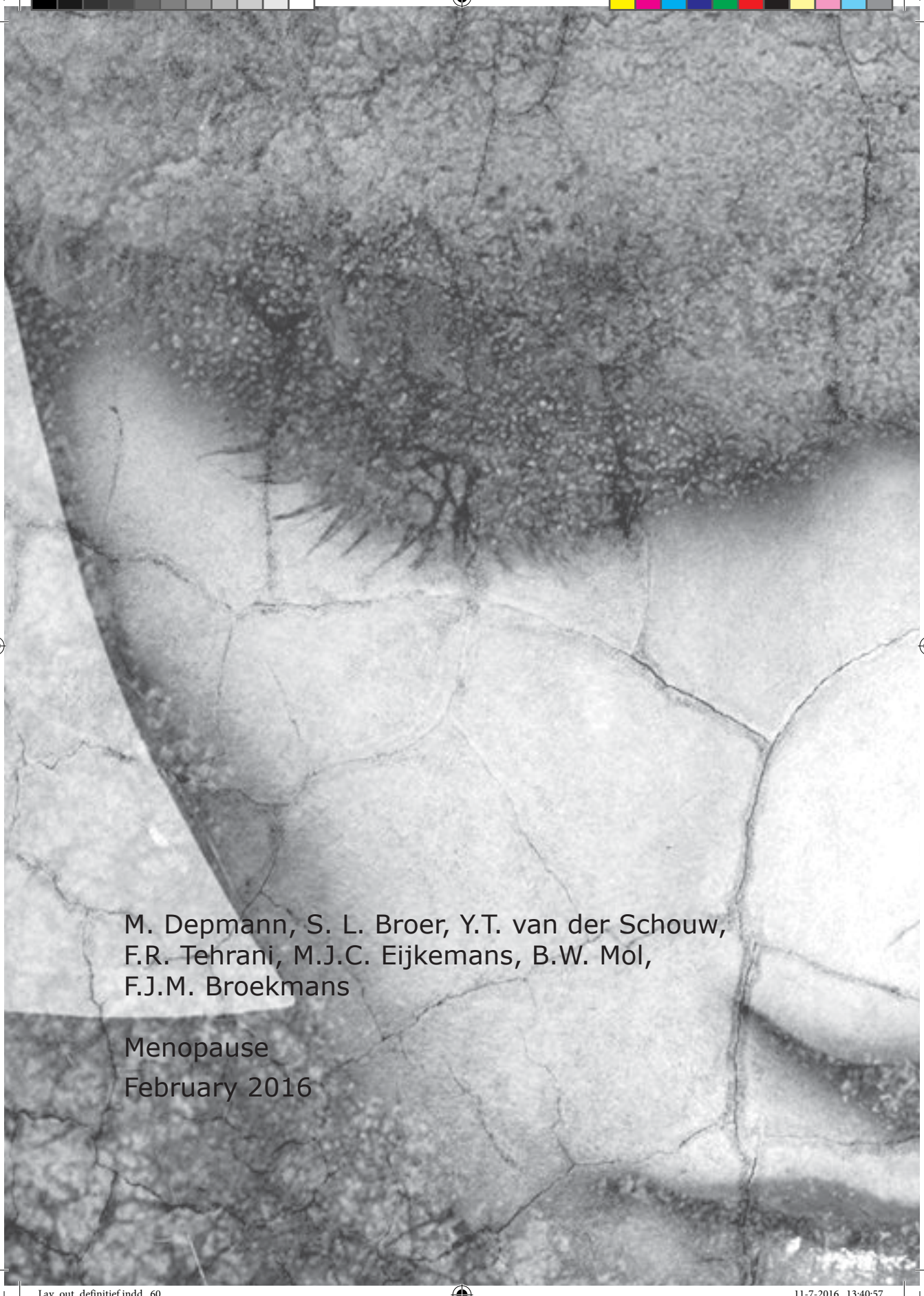
Our database consists of repeated measurements of AMH and the AFC performed in two subsequent cycles merged into one cycle for the purpose of the current analyses. Although significant differences might occur in the AFC and in AMH levels in subsequent cycles [68], when evaluating the parallelism in the distance between AMH and the AFC, as executed in the current analyses, potential "between cycle" fluctuations of these ORTs do not influence calculations made. As such, we feel that evaluating the fluctuations in levels of both AMH and the AFC in two subsequent cycles is justified.

Measurements for this study were performed between 1996 and 1997, in order to assess the relation between ORTs and reproductive age [70]. It was the recent paper by Jeppesen et al. [34] stating follicles ranging up to 8 mm in size are the main source for circulating AMH, that prompted us to look in this existing cohort. One could argue that using an early AMH assay, or a relatively outdated ultrasound machine might have influenced the outcomes. However, and as previously mentioned, AMH measurements were performed in a single highly experienced laboratory and in a single assay run and AFC measurements were performed by a single observer. We are therefore confident that the applied

technology for the measurements will not have jeopardized the present analysis on the interrelationship between the two variables.

A potential limitation of this study lies in its sample size of 44 women. However, when looking at our results it becomes clear that the fluctuation of the difference between $_{log}$AFC and $_{log}$AMH is highly non-significant (p=0.91) as is the Cohen's D. It is therefore not expected that a larger study population would alter these results. Moreover, our cohort reflects a representative selection of the fertile female population, therefore no large alterations in results are to be expected with a larger sample size. The strength of this study lies in the uniqueness of our data which allowed for repeated measures analyses on AMH and AFC. The parallelism observed between the fluctuation in the number of follicles ranging 2-8mm and fluctuations in peripheral AMH levels identifies the antral follicle pool as the main source of circulating AMH. Furthermore, the parallelism observed provides a solid base for the understanding of intended and potential use of AMH in clinical practice. For instance, the use of AMH in the diagnosis of polycystic ovarian syndrome is comprehensible since women with this syndrome demonstrate an elevated AFC and thus a corresponding elevated AMH level [63], whilst women facing menopause due to ovarian depletion have low peripheral AMH levels reflecting this cycle state [18].

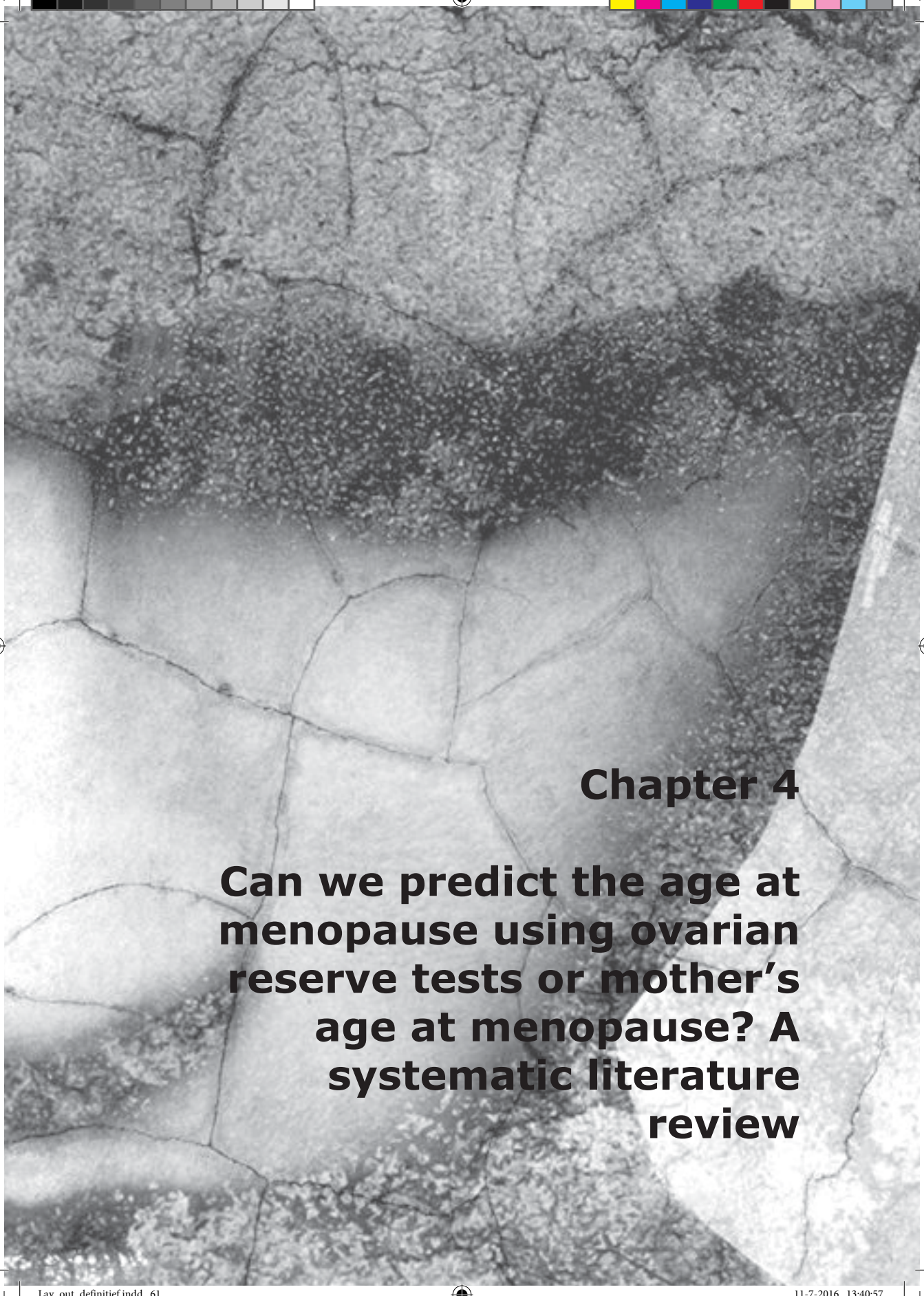
In conclusion, we have demonstrated that fluctuations in the levels of circulating AMH throughout the menstrual cycle parallel fluctuations in the AFC. This confirms the statement that AMH levels are mainly determined by the number of antral follicles sized 2-8 mm and identifies the antral follicle pool as the most likely source for serum levels of AMH. The parallelism observed provides a solid base for the basic understanding of AMH origin and provides a rational for intended and potential use of AMH in the clinical practice.



M. Depmann, S. L. Broer, Y.T. van der Schouw, F.R. Tehrani, M.J.C. Eijkemans, B.W. Mol, F.J.M. Broekmans

Menopause

February 2016



Chapter 4

Can we predict the age at menopause using ovarian reserve tests or mother's age at menopause? A systematic literature review

Abstract

Objective: This review aimed to appraise data on prediction of age at natural menopause (ANM) based on anti- Müllerian hormone (AMH), the Antral Follicle Count (AFC) and mother's ANM in order to evaluate clinical usefulness and identify directions for further research.

Methods: Three systematic reviews of literature were conducted identifying papers regarding menopause prediction based on either AMH, the AFC or mother's ANM, corrected for baseline age.

Results: The six papers selected in the search for AMH all consistently demonstrated AMH capable of predicting ANM (HR 5.6-9.2). The single paper reporting on mother's ANM stated this variable capable of predicting ANM (HR 9.1-9.3). Two studies provided analyses on AFC and yielded conflicting results, making this marker less strong.

Conclusion: AMH is the most promising marker currently available in ANM prediction. The predictive capacity of mother's ANM demonstrated in a single paper makes this marker a promising contributor to AMH in menopause prediction. Models, however, do not predict the extremes of menopausal age very well and have wide confidence intervals. Improvement is clearly needed before using these markers for individual prediction of menopause in the clinical setting. Moreover, potential limitations for such usage result from different AMH assays used and from lack of correction for factors or diseases affecting AMH levels or ANM. Future studies should include women of a broad age range, irrespective of cycle regularity and should base predictions on repeated AMH measurements. Furthermore, currently unknown candidate predictors need to be identified.

Introduction

In recent years an increasing amount of research has been invested in the prediction of age at natural menopause. The mean age at this final menstrual period is 51 years, with menopause occurring in a broad range of ages between 40-60 years [3]. Population based studies regarding the distribution of age at the end of natural fertility have suggested a fixed interval of ten years between the end of natural fertility and menopause, both occurring with a similar age range. [4] Since the postponing of childbearing has led to an increase in age-related infertility, a marker accurately assessing the individual limits of the fertile lifespan could reduce the number of women confronted with infertility. If women with a reduced reproductive lifespan could be identified timely, interventions such as timely family planning or cryopreservation of oocytes may help to prevent involuntarily childlessness. However, there is currently no marker capable of predicting age at the end of natural fertility. Therefore the final menstrual period has been used as a proxy variable for the end of natural fertility, based upon the presumed fixed time relation.

Individualized forecasts of the expected age at menopause are mostly studied using age in relation to cycle status. This combination builds a comprehensible predictor, a 35-year-old woman experiencing a regular cycle will most likely become menopausal in the normal age ranges, whilst a 45-year-old women still experiencing a regular cycle will reach menopause relatively late in life. However, age does not differentiate well for young women. A regular cycle at age 20, 25 or 30 for instance, does not provide information as to at what age these women would become menopausal. In order to enhance the precision of age at menopause predictions, the search for more specific predictors or a combination of predictors is ongoing.

Follicle Stimulating Hormone (FSH) has been demonstrated to accurately reflect the current reproductive state, and therefore is still a keystone in the classification of the current phases of reproductive ageing according to STRAW [37]. However the capacity of FSH in the prediction of the timing of future changes in reproductive state (i.e. the occurrence of menopause or menopause transition) is likely to be weak. [18,78] Therefore, interest is increasingly kindled towards other markers reflecting the ovarian reserve, of which anti- Müllerian hormone (AMH) and the antral follicle count (AFC) are the most promising ones.

The variation in levels of AMH or in the number of antral follicles amongst women of the same age- group and cycle state, has led to the hypothesis

that these ovarian reserve tests (ORTs) could aid in the prediction of age at menopause.

Furthermore, research designed to identify genetic markers responsible for the occurrence of the potential complex genetic trait menopause is ongoing. Although linkage analysis has detected few regions of interest and Genome Wide Association Studies have detected potential genetic loci [79], no dominant allele or alleles responsible for ovarian depletion have been discovered as of today [80]. Derivatives of genetic factors, though, as represented by mother's age at menopause (mother's ANM) seem promising, as they have underlined the high degree of heritability of menopause [22,81,82]. Consequently, this information, together with AMH and the AFC, is suggested to be used as a marker for the decline of natural fertility.

The aim of this review is therefore to appraise available data on AMH, the AFC and mother's ANM as to their capacity to predict time to natural menopause or age at natural menopause in women experiencing a regular cycle. This could aid in assessing their potential clinical usefulness and to identify the directions for further research. Based on the fixed 10-year interval between menopause and the end of natural fertility, the information on expected age at menopause could be used to identify women at risk for a reduced fertile lifespan timely, so that early family planning or cryopreservation of oocytes might prevent permanent childlessness. Furthermore, women at risk for diseases related to early or late menopause could also be identified."

Methods

A systematic review was conducted according to the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analysis) guideline [83]. No review protocol was previously published. Given the nature of this study, no institutional review board approval was required.

Eligibility criteria

There were no restrictions on types of studies to be considered eligible for inclusion in this review. Studies enrolling women with a regular cycle and a measurement of AMH or the AFC or a recorded mother's ANM were eligible for inclusion. Analyses performed required to model time to menopause (TTM), or age at menopause, using the ORT of interest or mother's ANM and needed to correct for age at baseline. Since both TTM and ANM were desirable targets for analysis, there was no preferred outcome measure that limited the search or

inclusion of studies and if different outcome measures were indeed presented by the papers selected, a narrative comparison was to be provided.

Information sources

We performed three extensive searches in the PubMed, Embase and Cochrane databases. All searches aimed for the prediction of menopause, the first in relation to AMH, the second in relation to AFC and the third in relation to mother's ANM, details of the full electronic searches can be found in appendix 1. The three searches were confined to papers published up to September 2014, there was no language constriction.

Study selection

The titles and abstracts of the publications available were manually screened, duplicate papers were discarded. Titles and abstracts of the remaining papers were screened to determine if they met the inclusion criteria of this review. When doubt existed, the full paper was examined for a more detailed assessment by a second author (FB) and irrelevant papers were excluded. Where more than one publication of a database existed, we included the publication with the most complete dataset. The reference lists from publications identified included in this review were scrutinized for additional studies that could prove to be of value.

Results

AMH and the prediction of menopause

After excluding duplicates, 419 titles were retrieved as depicted in the flow chart (Figure 4.1). Predefined inclusion criteria (Figure 4.1.) and referenced papers were screened.

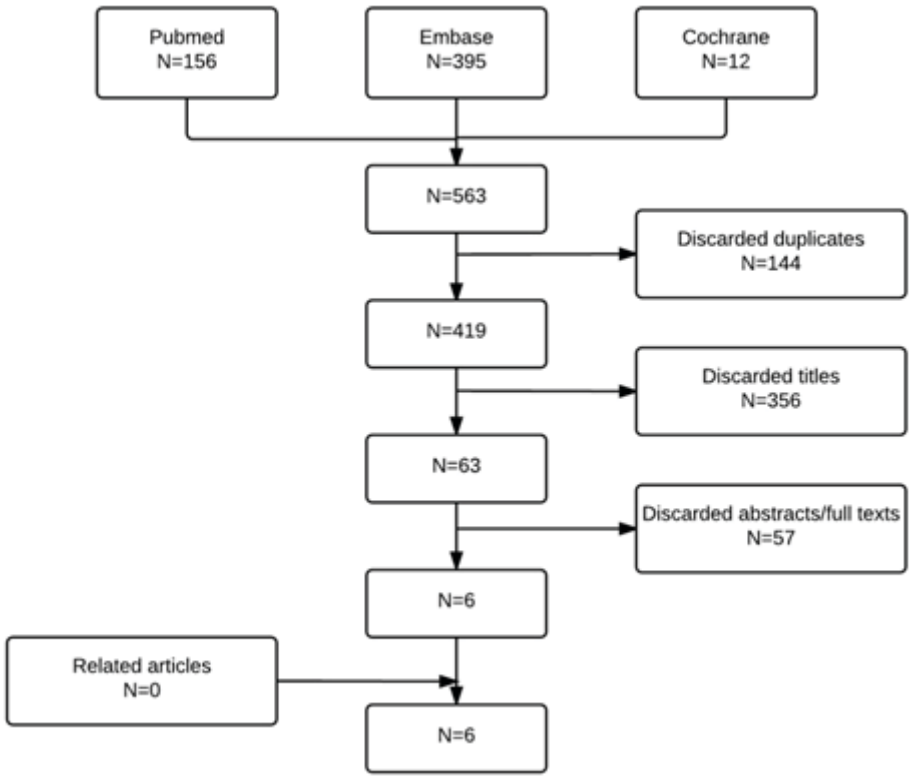


Figure 4.1: flowchart of search for AMH and menopause prediction; Inclusion criteria: regular cycle, baseline measurement of AMH, analyses modelling time to menopause (or age at menopause) using AMH and corrected for age at baseline.

This process resulted in 6 papers that could be included in this part of the review [16-19,21,22]. The papers selected consisted of five prospective and one cross-sectional study. The characteristics of these papers are depicted in Table 4.1.

Sowers et al. performed a TTM analysis using data from the Michigan Bone Health and Metabolism Study [16], 50 menopausal women, who experienced a regular cycle at inclusion 13 years earlier, were enrolled. Levels of circulating AMH (ELISA, Diagnostic System Laboratories, Webster, TX) were measured in frozen blood samples drawn at baseline and during six consecutive years. For the levels of AMH above the detection limit, Generalized Estimating Equations were used for analysis. For those below the detection limit, a mixed model analysis was applied.

Author Year	N (#MP at FU)	Outcome variable	Analytical Method	Results		
				[a] HR (95% CI) P-value	[b]C-stat	Other
Sowers 2010	50 (50)	TTM	Generalized Estimating Equations and Mixed Model Analysis			$_{log}$AMH 1 unit lower, age of menopause 1.75 y earlier (±0.14)
Tehrani 2011	266 (63)	ANM	Accelerated failure time modelling and AUC			Acceptable agreement between observed and predicted ANM (bias -0.3 95% CI -4 to 3 yrs); individual ANM predictions
Broer 2011	281 (48)	TTM & ANM	Cox regression analysis + C- statistic	9.2 (2.5-34)[c] <0.001	90	
Freeman 2012	401 (198)	TTM	Cox regression analysis	5.6 (4.7 -6.7)[d] <0.001		
Tehrani 2013	1015 (277)	ANM	Accelerated failure time modelling and AUC + C-statistic		90	Good agreement between observed and predicted ANM; individual ANM predictions
Dólleman 2014	150 (46)	TTM	Cox regression analysis + C-statistic	5 (1-22) <0.001	91	

Table 4.1: Characteristics of included studies for AMH and menopause

AMH= anti-Müllerian hormone; TTM, time to menopause; ANM, age at natural menopause; AUC, area under the curve. [a] Hazard ratio is the percent increase in the chance of reaching menopause during follow-up per unit decrease in AMH. [b] C-statistic is the percentage of women correctly predicted to reach menopause during follow-up (in an ANM analysis); or the capacity of the model to discriminate women with a short TTM from women with a long TTM (in a TTM analysis). [c] One unit of AMH equals 0.89 ng/mL. [d] One unit equals 1 SD logAMH.

AMH was significantly associated with TTM (1 unit decline in $_{log}$AMH resulted in reaching menopause 1.75 years earlier, $p<0.001$) in a model correcting for age at baseline. Furthermore, when AMH became non-detectable, it was also significantly associated with time to menopause.

Tehrani et al. calculated individual ANM and derived data from an on- going prospective cohort study [17,21]. The primary objective of this study, for which inclusions started in 1998, was to explore the risk factors of non-communicable diseases [84]. Data was obtained from two different samples of this cohort of regularly cycling women. For AMH measurement, frozen serum was used which was drawn at an undisclosed moment during the menstrual cycle. The first paper published in 2011 [17] was based on a random selection of 266 participants aged 20-50 years of whom 63 had reached menopause during follow up. AMH was measured using an enzyme- immunometric assay (Diagnostic System Laboratories, Webster, TX). A Weibull regression analysis was performed to model ANM based on baseline AMH and age at AMH measurement. The predicted ages at menopause, as derived from this Weibull regression analysis, were compared to observed ages at menopause within the cohort and to menopausal ages obtained from Iranian national data [85] using the Bland and Altman method and a quintile-quintile plot, respectively. An acceptable conformity was demonstrated between the observed and predicted ages at menopause within the cohort (estimated bias -0.3, 95% CI -4 to 3 years), although the level of agreement declined in the extremes of ages. In the QQ plots, a good visual agreement was present between predicted distribution of ages at menopause and menopausal ages derived from Iranian National data. ANM predictions for the highest AMH values, and thus corresponding highest menopausal ages, were set at an undisclosed age beyond 60 years of age. Each predicted ANM was accompanied by a prediction interval which ranged 2-4 years.

The paper published in 2013 [21] was based on data from 1,015 women aged 20-50 years, women included in the first paper were excluded from analyses. AMH assay was again obtained from frozen blood samples. This time however, the Gen II kit for AMH measurement was used (Beckham Coulter Inc. Fullerton, CA). In line with the 2011 analyses, a Weibull regression analysis produced predicted ages at menopause, which were compared to observed ages at menopause in the cohort using the Bland-Altman method. The medians of differences between observed and predicted menopausal ages within the cohort were equal (0.51 yrs, range -5.4 to 9 yrs). This time women with the highest age specific AMH values were predicted to experience menopause at a mean age >65 years of

age, the number of years greater than 65 years of age was not specified in the paper, prediction intervals ranged nine years at the minimum. Additionally, the proportion of correctly predicted events, as expressed by the C-statistic, was calculated. The actual and predicted menopause free survival plots were similar, but when extending the predictions to the cohort of women that did not experience menopause, a poorer level of agreement became apparent. The model with age as a sole variable had a C-statistic of 84% for the occurrence of either early or late menopause. This was improved to 92% when AMH was added to the model.

Broer et al. [18] performed TTM analysis next to individual ANM predictions. For these analyses three highly similar cohorts were combined to form one cohort of 281 women aged 21-46 years. AMH was measured using both the enzyme-immunometric assay (Diagnostic System Laboratories Inc., Webster, TX) and the immune- enzymometric assay (Immunotech-Coulter, Marseille, France), a correction factor was used to compare AMH levels. A multivariate Cox regression analysis correcting for age at baseline showed AMH highly capable of predicting TTM. The Hazard Ratio (HR), reflecting the proportional decrease in the chance of becoming menopausal during follow up with "every unit increase in AMH", was 0.092 (95% CI 0.025-0.340, p<0.001). For a model with AMH next to age at baseline the C-statistic, calculating the capacity of the model to discriminate women with a short time to menopause from women with a long time to menopause was 90%, compared to 87% in a model with age as a sole variable. Moreover, individual predictions of ANM were calculated using a Weibull survival model. Predictions were made using a nomogram for age-specific AMH concentrations that reflected a percentile category (i.e. p5 for the low age-specific AMH value and p95 for the highest age-specific AMH value). This percentile category then placed a woman in a corresponding menopause category (i.e. an age-specific AMH level at the 5th percentile represents women experiencing a relatively early menopause, with a corresponding shift of the age at menopause interval, and the 95th percentile represents women becoming menopausal late in life). The earliest predicted ANM was 42.1 years, the latest 60.1 years, the prediction interval ranged from 11 to 12.5 years.

Freeman et al. [19] included 401 women aged 35-48 years. AMH was measured using the ELISA kit (Beckham Coulter Inc. Brea, CA). A Cox regression analysis was performed assessing TTM adjusted for age at baseline. AMH was added to the model using quartiles of the crude baseline values. In this analysis, AMH showed a significant capacity in predicting time to menopause (HR 0.56, 95% CI 0.47-0.67, p<0.0001).

In a cross sectional paper, Dólleman et al. [22] focussed research on predicting ANM using AMH on top of mother's ANM. Aside from performing analysis on this primary outcome, a subgroup analysis was performed assessing TTM using AMH and age at baseline. For this subgroup analysis, two databases were pooled, one being the previously discussed database described in the paper by Broer et al. [18], and one obtained from a study that assessed whether menopausal ages differed between women who did and did not have a history of a trisomy-21 pregnancy, in this last cohort AMH was measured using an immune- enzymometric assay (Immunotech-Coulter, Marseille, France). [86] A multivariate Cox regression analyses assessing TTM using AMH and age at baseline showed that AMH significantly predicted time to menopause (HR 0.05, 95% CI 0.01-0.22, p<0.0001). The C-statistic of 91%, again representing the correctly predicted events, was strong for this model.

AFC and the prediction of menopause

In this literature search we retrieved 150 titles after excluding duplicates (Figure 4.2). Screening on title, abstract and full text resulted in the selection of two papers [18,20] of which the paper of Broer et al. was previously also selected in our AMH search. The characteristics of the selected papers are depicted in Table 4.2.

Author Year	N (#MP at FU)	Outcome variable	Analytical method	Results	
				[a] HR (95% CI) P-value	[b]C- stat
Broer 2011	281 (48)	TTM	Cox regression analysis +C-statistic	5.6(2.6-12)[c] p=0.135	88%
Wellons 2013	705 (101)	TTM	Cox regression analysis	18.9 (11.9-30.2)[d] p<0.001	

Table 4.2: Characteristics of included studies for AFC and menopause

MP: reached menopause; FU: Follow Up; ANM: age at natural menopause; TTM: time to menopause; HR: Hazard Ratio (% increase in chance of MP occuring during FU per [unit] decrease of AFC); CI: Confidence Interval; C-Statistic (% of correctly predicted MP occuring during FU); [c] 1 unit= 6.94 antral follicles; [d] AFC dichotomized ≤4 or >4

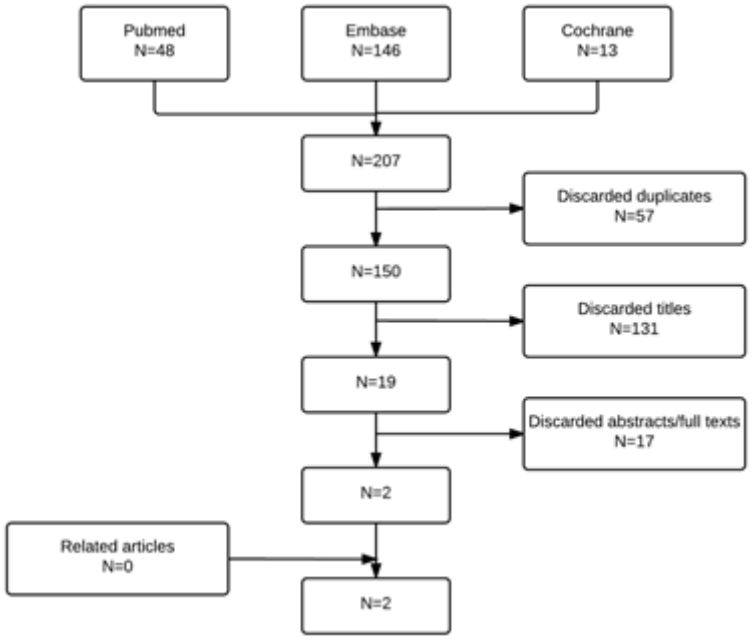


Figure 4.2: flowchart of search for AFC and menopause prediction; Inclusion criteria: regular cycle, baseline measurement of the AFC, analyses modelling time to menopause (or age at menopause) using the AFC and corrected for age at baseline.

Broer et al. [18] performed TTM analysis and used data from the cohort of 281 regular cycling women as previously described in the results regarding AMH and menopause. The AFC measurement was performed at cycle day 2-4. A Cox regression analysis for TTM with the baseline AFC as a single variable showed a statistical significant capacity for AFC in the prediction of TTM in the univariate regression (HR 0.13, 95% CI 0.068-0.230). However, in a multivariate regression correcting for age at initial screening, there was only a non-significant trend for added value of the AFC (HR 0.56, 95% CI 0.26-1.20; p=0.13).

Wellons et al. [20] performed TTM analysis using data on 705 women aged 18-30 years from the Coronary Artery Risk Development in Young Adults Study [87,88]. This longitudinal community-based study assessed the evolution of cardiovascular risk among young adults. A transvaginal ultrasound for AFC measurement was performed at an undisclosed time during the follicular phase of the menstrual cycle. Cox regression analysis was performed to predict TTM using a dichotomised value of AFC (≤4 and >4) as covariate and adjusting for baseline age and smoking. It proved AFC capable of predicting time to

menopause, the adjusted 7-year hazard of natural menopause was a near twofold higher (HR 1.89, 95% CI 1.19-3.02) for women with a low AFC (described as ≤4).

Mother's ANM and the prediction of menopause

The third extensive search was performed using synonyms for menopause, mother and prediction. After duplicates were excluded, 221 titles were retrieved (Figure 4.3).

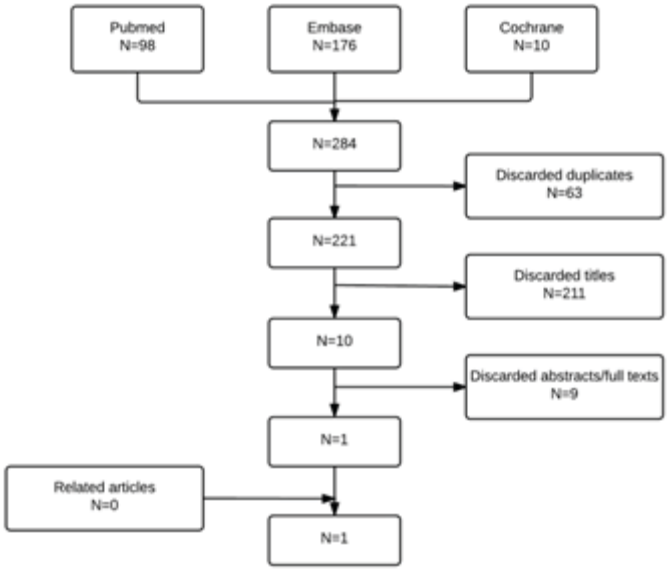


Figure 4.3: flowchart of search for mother's ANM and menopause prediction; Inclusion criteria: regular cycle, known mother's ANM, analyses modelling time to menopause (or age at menopause) using mother's ANM and corrected for age at baseline.

Screening on title, abstract and full text resulted in the selection of a single paper which performed a time to menopause selection, this paper was previously discussed in our search for AMH and menopause. [22] The characteristics of this paper are depicted in Table 4.3.

To assess the value of mother's ANM in predicting TTM, Dólleman et al. [22] added a third database to the previously described pooled database. This sample was selected from a cohort of women participating in a prospective study on determinants of the development of breast cancer [89].

Author Year	N (#MP at FU)	Outcome variable	Analytical method	Results	
				[a]HR (95% CI) P-value	[b]C-stat
Dólleman 2014 Cohort 1	164 (164)	TTM	Cox regression analysis + C-statistic	9.3 (9.0-9.6) p<0.0001	79%
Dólleman 2014 Cohort 2	150 (46)	TTM	Cox regression analysis + C-statistic	9.1 (8.4-9.7) p=0.01	85%

Table 4.3: characteristics of included studies for mother's ANM and menopause

MP: reached menopause; FU: Follow Up; ANM: age at natural menopause; TTM: time to menopause; HR: Hazard Ratio (% increase in chance of MP occuring during FU per [unit] decrease of mother's ANM); CI: Confidence Interval; C-Statistic (% of correctly predicted MP occuring during FU)

In a multivariate Cox regression analysis in the first cohort (as described in the section on AMH and menopause), mother's ANM was a significant predictor for time to menopause next to baseline age (HR 0.91, 95% CI 0.84-0.97 p=0.01). The C- statistics of 85%, reflecting the accuracy of the model used, was strong. Moreover, analyses in the newly added cohort for mother's ANM next to baseline age were also significant (HR 0.93, 95% CI 0.90-0.96, p<0.0001), while the C-statistics of 79% for this model were moderate.

Discussion

This review aimed to appraise the value of AMH, the AFC and mother's ANM in the prediction of menopause, an event that can be regarded as a proxy variable for the end of natural fertility. Published literature has consistently reported AMH capable to predict ANM. Moreover, AMH proved to be of added value to predictions based on female age. The scarce literature available regarding mother's ANM and the timing of menopause showed added value for this variable in a model next to female age. As for the AFC, though capable of forecasting menopause when comparing a low AFC to an arbitrarily chosen high AFC, no added value in a model using a continuous value for the AFC next to female age could be demonstrated. Therefore, it seems that AMH and mother's ANM are the more promising factors related to the process of ovarian ageing and the quantity decline of the follicle pool over time.

The paper selected in the search regarding genetic factors as represented by

mother's ANM draws a promising conclusion on the capacity of this variable in the prediction of ANM next to baseline age. [22] Though only this one paper reported longitudinal data and performed menopause predictions using mother's ANM, the potential value of this predictor has been extensively researched. Papers regarding mother's ANM [90-93] uniformly state that in women experiencing a relatively early menopause, their mothers, or daughters, are likely to have become menopausal early in life as well. Furthermore, daughters of women with an early menopause are reported to have relatively low levels of circulating AMH and a relatively low antral follicle count.[81,82,94], demonstrating the interrelation between genetic factors and quantitative ORTs. As such, knowledge of the mother regarding age at menopause may deliver crucial information for the daughter(s).

The results regarding the performance of AMH, the AFC and mother's ANM in the prediction of menopause depicted in this review are subjected to limitations present within the papers selected and should therefore be interpreted in that perspective. Within study limitations stem from the in- and exclusion criteria used, from the lack of accurate prediction for the extreme menopausal ages, from broad prediction intervals and from correction for known confounders. Furthermore, further restrictions arise from the incomparability of studies selected which stems from the use of different AMH assays and from the lack of uniformity in outcome measures described.

All papers selected performed analyses on women experiencing a regular cycle, this may imply an impediment of such studies into the general applicability of menopause forecasting. However, the addition of AMH, the AFC or mother's ANM however, is specifically desirable in cases that otherwise have "normal" expectations regarding the age to reach menopause, though incorporating women irrespective of cycle state would potentially provide added value if we wish to cover the full range of the normal menopause variation in the predictions.

For individual ANM predictions provided by papers focussing on AMH and menopause, two limitations became apparent. First, predictions of individual age at menopause did not cover the full age ranges, or may even exceed the limits of the normal distribution. In the paper by Broer et al. [18] the earliest age at menopause was 42.1 years (fig 2.). In the papers by Tehrani et al. [17,21] the other side of the spectrum is inadequately covered resulting in a mean ANM for women with an age specific AMH at the 95th percentile line which lies beyond 65 years of age. The inability in predicting the extremely early menopause in the paper by Broer et al. [18] potentially originated from the fact

that a regular cycle was used as an inclusion criterion and mean age at inclusion was relatively high. Therefore, women doomed to experience menopause at a young age (40-45 years), possibly reflected by the presence of irregular cycles earlier in life, were not included in this cohort. The predictions in the late menopause group exceeding normal age limits, as seen in the papers by Tehrani et al. [17,21], may have originated from an incomplete follow up. When modeling age at menopause, extrapolations based on the constructed model are made for women not yet having experienced the final menstrual period. With modest follow up time, women experiencing a late menopause are largely represented by the extrapolating model. As a result, predictions for these women may be relatively unreliable, as reflected by a mean age of menopause for this group which lies beyond 65 years of age. Furthermore, the remarkable difference between highly precise predictions for the early ANM and the broad age interval for women experiencing a late menopause might stem from some inconsistency in the model used, that can potentially be solved by a sufficiently extended follow up.

The other limitation in individual ANM predictions is the broad age interval of predictions made. [17,18,21]. This likely originated on the one hand from variation when measuring AMH, as a result of both assay and biological variation [28] and on the other hand from the fact that not all potential factors regulating the timing of menopause may have been represented by the factor AMH in the models used. The first problem can potentially be partially faded out by using repeated measurements of AMH over a short time period [66,68] and applying a rigid assay methodology, which is anticipated to become available soon [95]. The second problem can only be solved by identifying other factors related to menopause. The search for these other factors is still ongoing, and candidate markers are genetic factors or lifestyle characteristics, such as BMI.

Smoking at the time of AMH assay is a known confounder [73], resulting in lower age specific AMH values, furthermore it is associated with an earlier age at menopause [96]. Of the papers selected regarding AMH, none corrected for both smoking at follow up and smoking at baseline. Freeman et al. [19] corrected for smoking at the time of AMH measurement, and demonstrated that smokers were significantly more likely to reach menopause during the 14-year follow up period (HR 1.61, 95% CI 1.19-2.10, p=0.002), and reached menopause in a shorter time interval (9.52 years compared with 10.02 years. Sowers et al. [16] adjusted for smoking at baseline and concluded that smokers became menopausal at an earlier age and experienced a more rapid decline in AMH values (analysis made using six consecutive measurements of AMH with a

one-year interval) when compared to non-smokers. Broer et al. [18] added smoking at follow up to the Cox regression analysis and showed that current smoking did not add predictive value to the model with baseline age and AMH. One of the papers selected regarding the AFC corrected for current smoking, Wellons et al. [20] added smoking at follow up to the multivariate Cox regression analysis for time to menopause next to the dichotomized AFC (>4 or <4), age at baseline and "stable versus unstable menses". smoking at follow up was a significant predictor for TTM (HR 1.75; 95% CI 1.07-2.87; p=0.03). Since none of the papers selected corrected for both smoking at follow up and smoking at baseline it is imaginable that a potential subduing effect of baseline smoking on AMH on the one hand, or an earlier menopause resulting from current smoking on the other hand, might have compromised analyses performed.

As previously stated, comparison of results between papers is most difficult since different AMH assays are used and there is no uniformity in targets for analyses (i.e. TTM or ANM) and corresponding outcome measures described. TTM was used in three papers regarding AMH [17,18,22], and expressed using a HR in all these papers, reflecting the percentage decrease in the chance of menopause occurring during follow up with each "one unit reduction in AMH". However, the "one unit in AMH", as used in the HR, has been differently chosen in each paper, imposing an additional difficulty for pooling. The remaining three papers reported on ANM [16,19,21] and one paper calculated both TTM and ANM [18]. This difference in outcome measures used was observed in the papers regarding the AFC as well, where both papers selected reported on TTM [18,20], but again using different values reflecting "one unit decline in the AFC".

The papers on AMH used three different assays for analysis of circulating levels of AMH and measurements were performed in six different laboratories. Sampling timing in relation to the menstrual cycle also varied between papers. Due to large inter- assay variability on the one hand and the inevitable yet substantial variability caused by the fact that each laboratory provided its own value ranges and calibration on the other hand [97], no correction factor is available to unify these assays. Therefore, observations made are assay specific and cannot be extrapolated to data available from other AMH assays or other laboratories, thus making pooling of data for meta-analysis impossible.

Other factors make comparison of results between studies on the one hand and comparison of published data with the general population on the other hand difficult. One is the fact that AMH levels or age at menopause can be influenced by factors which were not addressed in the selected papers. The majority of data for instance originated from Middle Eastern women (59%;

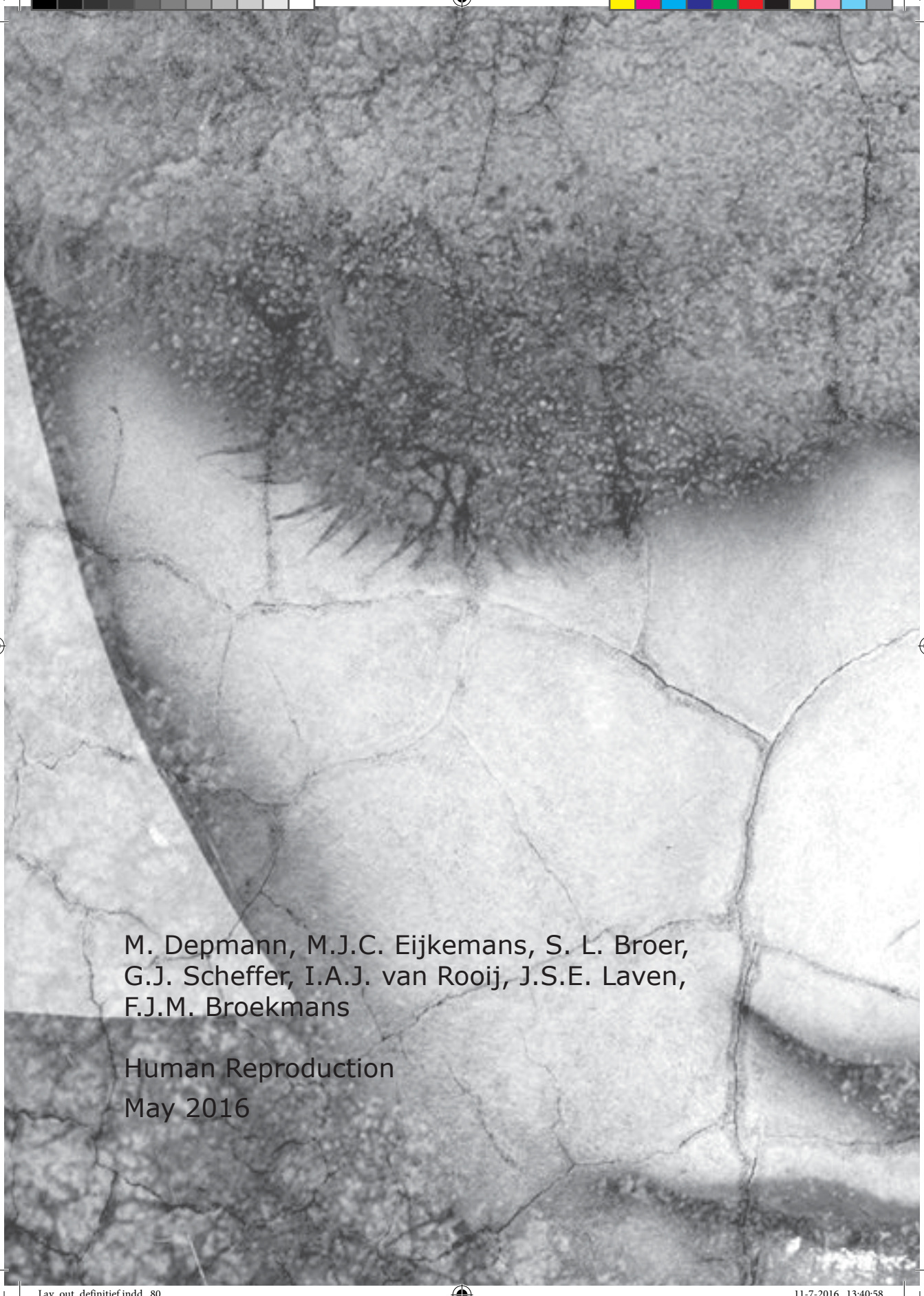
see supplementary Table 4.1), making extrapolation of menopause predictions perhaps less reliable to women of other ethnic descent. However, this depends on the question whether differences between ethnic groups exist regarding distribution of age ant menopause, or the decline pattern of primordial follicle numbers over time. The latter may come to expression in differences in AMH levels. Yet, so far effects of ethnicity have not been reported on cross-sectional studies. [98] Moreover, AMH levels are known to be subdued in women using oral contraceptives [73]. Since all selected papers excluded women using exogenous hormones, clinicians should take into consideration that translation of these findings into current clinical practice will not apply to women using oral contraceptives. Lastly, age at menopause may be influenced by a variety of illnesses such as treatment for malignant diseases, several genetic traits and possibly auto-immune diseases [99]. As depicted in supplementary Table 4.1, the inclusion criteria of studies selected mostly excluded these women from analysis. Analysis were therefore performed on otherwise normal women from the point of view of ovarian ageing, and should be interpreted in that context.
This systematic literature review is the first to assess the performance of AMH, the AFC and mother's ANM in the prediction of menopause. Our extensive search provided six papers addressing this subject that incorporated adequate correction for female age. The overview of literature provided in this review allows for critical evaluation of potential clinical use of ORT's and mother's ANM in menopause prediction and allowed for identification of directions for further research.
The main justification for ANM prediction is the prevention of the situation in which women remain childless due to the fact that they delay family planning. Naturally this implies that women predicted to experience menopause at a young age, are willing to take action. This would for instance result in starting a family relatively young when compared to their peers, or to decide on cryopreservation of oocytes in order to be able to conceive later on in life. Interestingly though, in spite of the growing amount of research on the prediction of menopause and the prediction of the end of natural fertility, scarce research is present investigating the need amongst women for a marker assessing the reproductive lifespan. [100,101] The gap between the doctor's wish to predict menopause in spite of not knowing if a need for this exists amongst the target group must be bridged by new research.
In summary, this systematic literature review is the first to assess the performance of AMH, the AFC and mother's ANM in the prediction of menopause. Our extensive search provided seven papers addressing this subject that

incorporated adequate correction for female age. The overview of literature provided in this review allows for critical evaluation of potential clinical use of ORT's and mother's ANM in menopause prediction and allowed for identification of directions for further research. This review demonstrated AMH and mother's ANM to be the most promising predictors currently available for possible use in daily clinical practice. Models lack capacity in the prediction of the extreme menopausal ages and provide wide prediction intervals, making AMH currently not applicable in the prediction of menopause or in the prediction of the end of natural fertility. If a large cohort of women with a broad range of ages and a known mother's ANM, irrespective of cycle state would undergo repeated baseline AMH and AFC measurements and a correction for smoking and baseline age was incorporated in predictions, a firm statement could be made on the predictive capacity of AMH, the AFC and mother's ANM towards the age of natural menopause. Information that could be used in family planning and to identify women at risk for developing diseases due to early or late menopause. Moreover, additional research should be aimed at identifying other potential contributors to age at natural menopause.

Supplementary material

Author Year	Population origin	Inclusion criteria	AMH Assay	AMH levels	Detection limit (ng/ml)
Sowers 2010	Caucasian N=50	25-46 yr, regular cycle, proven natural fertile, no endocrine disease, no ovarian/uterus abnormalities/ surgery, no hormone use	ELISA (Diagnostic System Laboratories, Webster)	0.66 (±0.50)*	0.05
Tehrani 2011	Middle- Eastern N= 266	20-50 yr, regular cycle, proven natural fertile, no endocrine disease, no ovarian/uterus abnormalities/ surgery, no hormone use	Enzyme- immunometric (Diagnostic System Laboratories)	<30 yrs 5.7 (±3.3)* 30-39 yrs 3.3 (±2.6)* ≥40 yrs 1.0 (±1.3)*	0.006
Broer 2011	Caucasian N= 281	18-46 yr, regular cycle, no ovarian/uterus abnormalities/ surgery, no hormone use	1) Enzyme-immunometric assay (Diagnostic System Laboratories) 2) Immune- enzymometric assay (Immunotech-Coulter)	Unknown	1) 0.026 2) 0.05
Freeman 2012	African-American (49.4%) Caucasian (50.6%) N= 401	35-48 yr, regular cycle, intact uterus, presence of 1 or 2 ovaries, no hormone use, no diabetes, liver disease or breast- or endometrial cancer	ELISA (Beckman-Coulter)	1.08 (±1.19)*	0.10
Tehrani 2013	Middle- Eastern N= 1015	20-50 yr, regular cycle, proven natural fertile, no endocrine disease, no ovarian/uterus abnormalities/ surgery, no hormone use	Gen II (Beckham Coulter)	1.65 (±1.81)**	unknown
Dólleman 2014	Caucasian N= 150	Group 1: 37-41 yr, mother's ANM available Group 2: 18-46 yr, regular cycle, no ovarian/uterus abnormalities/ surgery, no hormone use	1) Enzyme-immunometric assay (Diagnostic System Laboratories) 2) Immune- enzymometric assay (Immunotech-Coulter)	1.5 (0.5-2.9)***	1) 0.026 2) 0.05

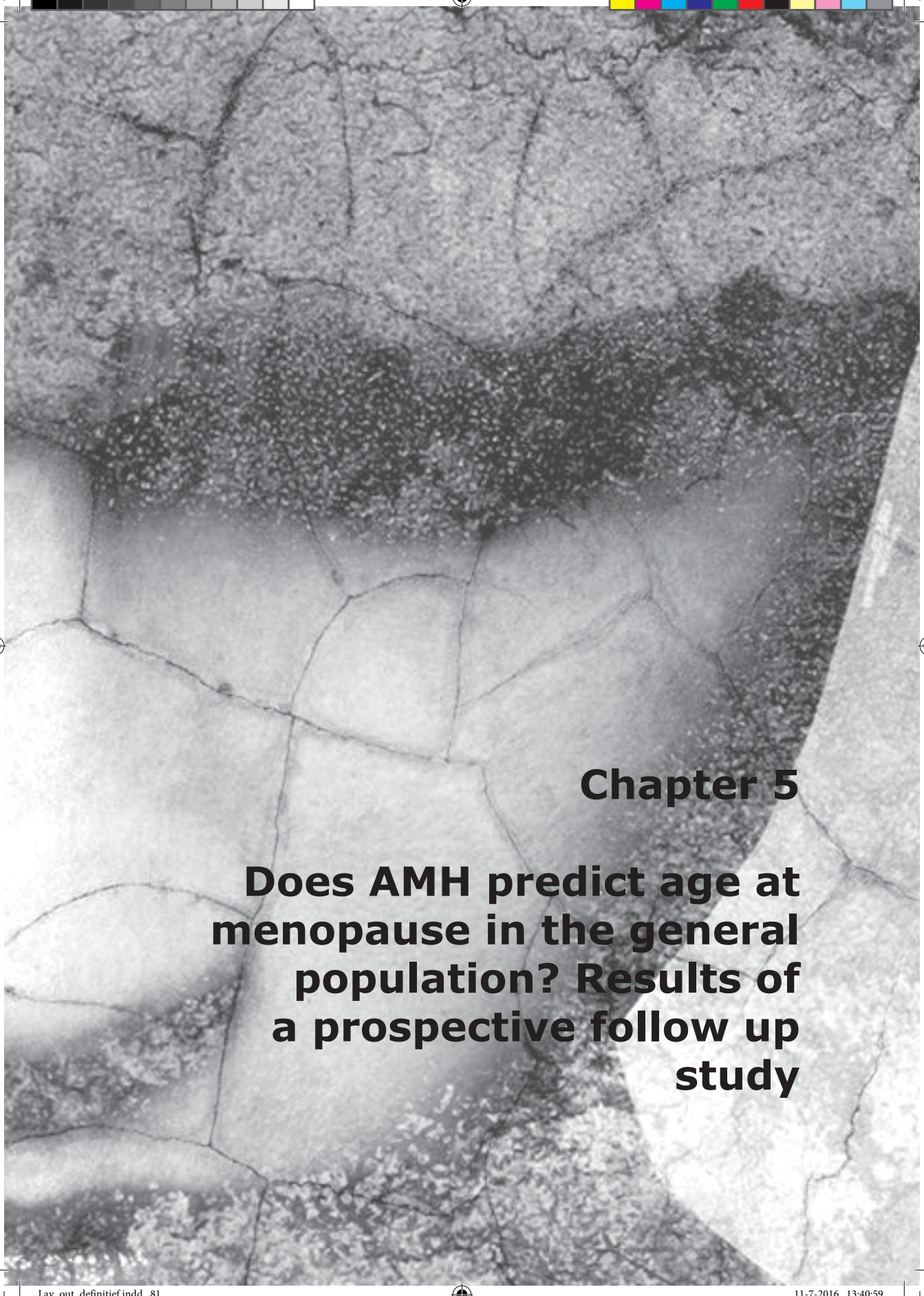
Supplementary Table 4.1: Baseline characteristics of studies included for AMH and menopause prediction (*mean (±SD) in ng/ml, **(*mean (±SD) in ng/dl, *** median (interquartile range) in ng/ml)



M. Depmann, M.J.C. Eijkemans, S. L. Broer, G.J. Scheffer, I.A.J. van Rooij, J.S.E. Laven, F.J.M. Broekmans

Human Reproduction

May 2016



Chapter 5

Does AMH predict age at menopause in the general population? Results of a prospective follow up study

Abstract

Study question: Do ovarian reserve tests (ORTs) predict age at natural menopause (ANM) in a cohort of healthy women with a regular menstrual cycle?

Summary answer: Of the ORTs researched, anti-Müllerian hormone (AMH) alone predicts age at menopause, it's predictive value however decreased with increasing age, prediction-intervals were broad and extreme ages at menopause could not be predicted.

What is known already: A fixed interval is hypothesized to exist between ANM and age at loss of natural fertility. Therefore, if it is possible to predict ANM, one could identify women destined for early menopause and thus at higher risk for age related subfertility. Of ORTs researched in the prediction of ANM, AMH is the most promising one.

Study design, study size and duration: A long-term, extended follow up study was conducted, results of the first follow up round were previously published. 265 normo- ovulatory women (21-46 yrs) were included between 1992-2001, 49 women (18.5%) could not be reached in the current follow up round.

Participants, setting, methods: 265 healthy normo- ovulatory women were included in an Academic hospital. We measured baseline AMH, follicle stimulating hormone (FSH) and the antral follicle count (AFC). At follow up (2009 & 2013), menopausal status was determined via questionnaires. Cox regression analysis calculated time to menopause (TTM) using age and ORT. A check of (non-) proportionality of the predictive effect of AMH was performed. A nomogram was constructed using a Weibull survival model in order to predict individual ANM.

Main results and the role of chance: In total, 155 women were available for analyses. 81 women (37.5%) had become post-menopausal during follow up. Univariable Cox regression analysis demonstrated age and ORTs to be significantly correlated with TTM. Multivariable Cox regression analysis, adjusting for baseline age and smoking, however demonstrated AMH alone to be an independent predictor of TTM (Hazard Ratio 0.70, 95% Confidence Interval 0.56-0.86, p-value <0.001). A (non-)proportionality analysis of AMH over time demonstrated AMH's predictive effect to decline over time.

Limitations, reason for caution: The predictive AMH effect observed became less strong with increasing age. Individual AMH based age at menopause predictions did not cover the full range of menopausal ages, but did reduce the variation around the predicted ANM from 20 years to 10.1 years.

Wider implications of the findings: Age specific AMH levels are predictive for ANM. Unlike in our previous publication however, a declining AMH effect with increasing age was observed. This declining AMH effect is in line with recent long term follow up data published by others. Moreover, the accompanying predictive inaccuracy observed in individual age at menopause predictions based on AMH, makes this marker currently unsuitable for use in clinical practice.

Introduction

Menopause, the final menstrual period, occurs at a mean age of 51 years, though a considerable variation in age exists for the occurrence of this event. Some women reach menopause at 40 years-of-age, while others experience their last menstrual period at age 60. [102] With an equally wide age distribution, but approximately ten years prior to menopause, a definite loss of natural fertility occurs, preceded by a period of gradually declining natural fertility [4,5]. The comparibility of the age distribution for the occurence of menopause and the final loss of natural fertility suggests the existence of a fixed time interval between these events for individual women.

In recent years, an increasing number of women is seeking expensive fertility treatment for age-related subfertility. Therefore, a need exists for a marker accurately assessing the limits of one's fertile lifespan and capable of identifying young women that cannot afford too much postponement of family building. In the absence of such a marker and based upon the assumed fixed time interval between fertility loss and menopause, age at natural menopause (ANM) is used as a proxy variable for the loss of natural fertility.

Menopause is considered to result from a drop in the number of primordial follicles below a critical threshold.[25,39-41,103] Since the number of antral follicles is related to the size of the primordial follicle pool [29], a marker correctly reflecting the number of antral follicles is potentially capable of predicting timing of menopause. Known potential markers, or ovarian reserve tests (ORTs), are anti-Müllerian hormone (AMH) [16-19,21-23], follicle stimulation hormone (FSH) and the antral follicle count (AFC) as measured by transvaginal ultrasound [14,18,20].

The aim of the present study is to investigate if single measurements of baseline endocrine or ultrasound markers are capable of predicting the occurrence of menopause in a group of normo-ovulatory female volunteers. This will be done by developing a multivariable prediction model predicting individual age at menopause. This study is an extension of follow up time of a previous study

that this group published[18], in which AMH alone was capable of predicting time to menopause in a cohort of 257 women of whom 48 had reached menopause. By extending follow up time, thus creating a larger group of women after menopause, a more robust analysis on the relation between ovarian reserve markers and menopause was to be performed.

Materials and Methods

The study group in this extended prospective follow up study consists of three different cohorts, previously pooled successfully [18]. The first cohort is a group of 172 healthy female volunteers, recruited in 1996 en 1997 [104]. Women met inclusion criteria when they were 25-46 years of age, had a regular menstrual cycle defined as a mean cycle length of 21-35 days and the next period predictable within a 7 day time-frame. Al women were proven natural fertile, meaning they experienced at least one pregnancy within one year of cessation of contraceptives, resulting in a normal term delivery. Hormonal contraceptives needed to be discontinued at least three months prior to inclusion. Ovarian surgery or – abnormalities were considered to be exclusion criteria.
The second cohort is a group of 90 healthy female volunteers, recruited between 1999 and 2001 [71,105]. Women were eligible if aged 18-46 years and when heaving a regular menstrual cycle of 21-35 days. Women attempting to conceive either their first or second child could be included. Hormonal contraceptives needed to be cessated for at least three months. Exclusion criteria were adnexal surgery, or a history of infertility.
The last cohort is a group of 40 women recruited between 1983 and 1992 [106-108]. Later on they were asked to participate in a prospective longitudinal study on ovarian function [109]. They met inclusion criteria when aged 20-35 years, having a regular cycle with a mean of 26-31 days, and having a body mass index between 19-26kg/m^2. Exclusion criteria were endocrine disorders, relevant disease or infertility treatment. Hormonal contraceptives were dicontinued at least three months prior to inclusion.

Ethical approval
All studies have been approved by the institutional review boards of the University Medical Center involved (i.e. Utrecht and Rotterdam). Written informed consent was obtained from each patient.

Design

The first visit to the clinic (T1, 1992-2001) was planned on the second, third or fourth day of the menstrual cycle. During this visit, the number of antral follicles (AFC, defined as the number of follicles in a range of 2-10mm) was determined by transvaginal ultrasound and a blood sample was provided. The AFC was performed by a limited group of well-trained physicians, applying a standardized approach in visualising the ovaries and counting and measuring of the antral follicles [104]. Serum and plasma were separated and stored at -20°C until assayed for AMH and FSH levels.

Between 2008 and 2010 (T2) and in 2012-2013 (T3) all women were contacted and asked to fill out a standardized questionnaire. The questionnaire contained questions on cycle regularity and reproductive and medical history, amongst other questions. It was designed to place women into one of the following five reproductive categories: regular cycle, menopausal transition, menopause, the use of exogenous estrogens or surgical removal of uterus and/or ovaries. The results regarding the questionnaire at T2 were previously published [18].

Definitions

A regular cycle was defined as a mean cycle length of 21-35 days and the next period predictable within a seven day time-frame. Menopausal transition was defined according to the STRAW criteria as a mean cycle length less than 21- or more than 35 days during the previous six months or longer, or a mean cycle length between 21-35 days but the next period not predictable within a seven day time-frame. Menopause was defined as no menstrual period in the previous 12 months or more. Women using exogenous hormones (including a Mirena intra- uterine device and progestagens) at the time of the most recent follow up were excluded from analysis, since no cycle state could be determined. Hormone use between baseline and last follow up was thus not regarded an exclusion criterion.

Moreover, women who underwent surgical removal of the uterus or one or both ovaries, either at baseline or at the most recent follow up were excluded from analyses.

Hormone assay

Hormone concentrations were measured in plasma (FSH) and serum (AMH). Specimens were stored at -20 C until processing. Concentrations of FSH were measured with the use of the MEIA technology on a fully automated AxSYM immune-analyzer (Abbott Laboratories, Abbott Park, IL.). The World Health Organization Second International Reference Preparation for human

FSH (78/549) was used as a standard in the FSH assay. For FSH, inter-assay coefficients of variation were found to be 5.7%, 5.7%, and 7.8% at the levels of 5, 26, and 79 IU/liter, respectively (n=80). The detection limit for the FSH assay was 0.03 IU/liter.

Two different AMH assays were used in our cohort. The first is the enzyme-immunometric assay (Diagnostic Systems Laboratories, Inc., Webster, TX.). Inter-assay and intra-assay coefficients of variation were less than 5% at the level of 3.0 ng/ml and less than 11% at the level of 13.0 ng/ml. The detection limit of the assay was 0.026 ng/ml. The second assay used is the ultrasensitive immune-enzymometric assay (Beckham Coulter, Marseille, France). The limit of detection (defined as blank +3 SD of blank) was 0.05 ng/ml. Intra and inter-assay coefficients of variation were less than 5% and less than 8%, respectively. For both assays, values below the detection limit were artificially set at half the detection limit.

In order to compare our data with recent publications and to pool our two different AMH assays, a two-step conversion factor was developed. In the first step, AMH levels measured with the enzyme-immunometric assay (Diagnostic Systems Laboratories, Inc., Webster, TX.) were multiplied by a factor 2 in order to transform them to the ultrasensitive immune-enzymometric assay (Beckham Coulter, Marseille, France)[15].

In the next step, the Beckham Coulter levels (those obtained in the previous conversion step and the original levels as available for part of the cohort) were transformed to Gen II assays by multiplying them by a factor 1.564 [110,111]. It should be noted that the Gen II conversion factor corresponds to the first generation Gen II assay levels measured before July 2013. In July 2013, a new manufacturer's instruction for Gen II assay sampling became available instructing to pre-dilute serum samples. Assay levels obtained with undiluted serum samples, thus before July 2013, were proven to be significantly lower [112]. Therefore, one should be cautious when comparing AMH levels measured in our study to AMH levels measured with other assays, or second generation Gen II values.

Analyses

Baseline characteristics of women lost to follow up were compared to women included in the analysis using the Mann-Whitney U test. Next, baseline characteristics and ORTs were analyzed for women divided into subgroups according to their cycle state at T3, using the Kruskal-Wallis or Chi square test. Univariable and multivariable Cox regression analysis for time to menopause or

time to T3 as time axis and menopause as event were performed to assess the capacity of the different ORTs in predicting time to menopause with a correction for age, smoking at baseline and current smoking (i.e. smoking at follow up) in the multivariable analysis. Prior to these analyses, single imputation was performed for cases with missing data.

The C-statistics were calculated to measure the capacity of the model to discriminate between women with a short or a long time to menopause. More precisely, the C-statistic reflects how accurate the model can distinguish between two randomly chosen women who have different times to menopause..

For each covariate that proved to be significant in the multivariable analysis, a check of (non-) proportionality for the risk of becoming menopausal with increasing age was performed.

Moreover, for covariates that proved to be significant in the multivariable analysis, a Weibull survival model was composed with age at the time axis (with delayed entry for age at T1) and age specific percentages of the ORT as a single covariate of the prediction of age at natural menopause. A Weibull survival model uses observed events (i.e. menopause) in order to fit a statistical distribution of predicted events. Participants were divided into percentiles for their age-specific ORT level. Per age category, ORT levels corresponding with the different percentiles and medians and ranges of predicted age at menopause distribution were depicted in a nomogram.

Moreover, two additional analyses were performed. The first was to assess whether the use of two different AMH assays in our cohort, had affected the results. The second is a subgroup analyses on complete data in order to determine if imputation of missing values could have affected the outcomes.

Data analysis was performed using SPSS 20 (SPSS Inc., Chicago, IL) and R version 3.0.3 (http://www.r-project.org).

Results

As shown in Figure 5.1, of the 265 women eligible, 216 responded to the questionnaire. The remaining 49 women (18.5 %) could not be reached despite repeated attempts to contact them (via (e)mail and telephone and after performing an address check at the municipal base administration).

Of the 216 women available for analysis, 81 women (37.5%) had experienced their last menstrual period and were thus classified postmenopausal, 31 women (14.4%) were in the menopausal transition and 43 women (19.9%) still experienced a strictly regular cycle.

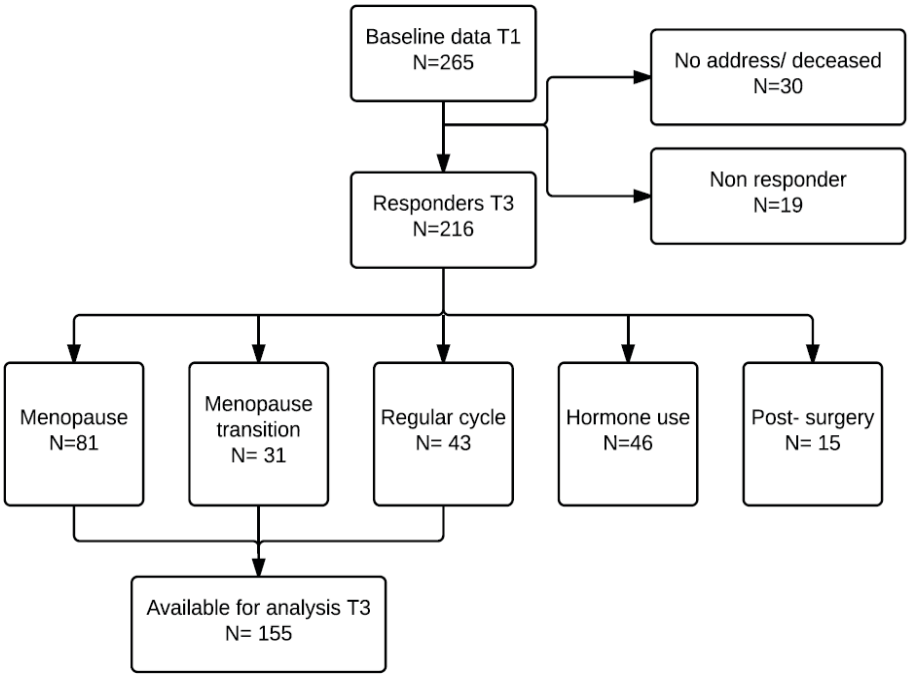


Figure 5.1: flowchart of inclusions

T1= baseline, T2= first follow up round (2009), T3= current follow up

A total of 46 women using hormones (21.3%) and 15 women post-surgery of uterus or ovaries (6.9%) were discarded from analysis.

The patient characteristics of the 49 women that were lost to follow up were compared to those of the 216 women that did respond to the questionnaire. Women lost to follow up were significantly younger than women included in the analysis (mean age 32.0 years vs. 36.0, $p<0.001$), which was also reflected in their ovarian reserve test results (AMH 4.2 vs 2.6 and AFC 14.1 vs 10.7, $p=0.001$ and $p=0.002$ respectively). Patient characteristics and ORTs at T1 were compared between the subgroups based on cycle status at T3 (Table 5.1). In line with the previously published data [18] and with increasing loss of a regular cycle (i.e. progression from a regular cycle, via menopausal transition to menopause), a significant difference in age upon initial screening and a longer follow up time was observed between the subgroups.

Moreover, the level of basal FSH was significantly higher with increasing loss of menstrual cyclicity, whilst AMH and AFC were significantly lower. Mean age at menopause was 50.0 years (range 35.5-56.5 years).

	Cycle Status at T3			
	Menopause	**Menopause transition**	**Regular cycle**	**P-value**
	(n=81)	**(n=31)**	**(n=43)**	
Age at T1 (yrs, SD)	41.63 ± 3.39	33.58 ± 4.73	31.56 ± 3.83	**<0.001**
Age menarche (yrs, SD)	13.07 ± 1.97	12.69 ± 1.51	13.01 ± 2.08	0.83
BMI at T1 (kg/m², SD)	23.82 ± 3.46	24.05 ± 4.99	23.67 ± 3.26	0.88
BMI at T3 (kg/m², SD)	25.11 ± 3.66	25.05 ± 4.03	24.88 ± 4.54	0.65
Smoking at T1 (%)	33.3	36.7	26.2	0.61
Smoking at T3 (%)	17.6	6.5	4.7	0.12
Interval T1-T3 (yrs)	14.84 ± 1.21	14.76 ± 1.89	13.86 ± 1.40	**<0.001**
FSH (IU/L, SD)	9.77 ± 5.97	6.96 ± 2.91	6.21 ± 2.22	**<0.001**
AMH (ng/ml, SD)	1.12 ± 1.31	3.01 ± 3.11	4.12 ± 2.87	**<0.001**
AFC (N, SD)	6.08 ± 3.89	12.13 ± 5.55	15.38 ± 7.82	**<0.001**

Table 5.1: Baseline characteristics comparing women according to cycle state at T3; T1= baseline, yrs=years, T3=current follow up, SD= standard deviation, BMI= body mass index, ORT=ovarian reserve test, FSH=follicle stimulating hormone, AMH=anti-Müllerian hormone, AFC=antral follicle count (follicles 2-10mm)

For women with missing data, single random imputation was performed as scheduled. Missing data occurred on AFC (n=8), FSH (n=3), AMH (n=5), baseline smoking (n=17) and current smoking (n=47). Three women with missing data on age at menopause were excluded from analysis.
Results of the Cox regression analyses are depicted in Table 5.2.

	HR	95% CI	P value	C-statistic
Univariate analysis				
Age at baseline	1.39	1.30-1.49	**<0.001**	**85%**
FSH (IU/L)	1.11	1.07-1.15	**<0.001**	66%
AMH (ng/ml)	0.53	0.43-0.65	**<0.001**	78%
AFC	0.79	0.74-0.84	**<0.001**	79%
Multivariate analysis				
- corrected for age at baseline-				
FSH (IU/L)	1.03	0.99-1.07	0.19	85%
AMH (ng/ml)	0.68	0.56-0.84	**<0.001**	**86%**
AFC	0.97	0.90-1.05	0.49	85%
Multivariate analysis				
-corrected for age and smoking at baseline-				
FSH (IU/L)	1.02	0.97-1.06	0.43	85%
AMH (ng/mg)	0.70	0.57-0.86	**<0.001**	**87%**
AFC	0.98	0.91-1.05	0.54	85%
Multivariate analysis				
-corrected for age and smoking at follow up-				
FSH (IU/L)	1.02	0.98-1.06	0.42	86%
AMH (ng/mg)	0.70	0.56-0.86	**<0.001**	87%
AFC	0.97	0.90-1.05	0.46	85%

Table 5.2: Menopause prediction; FSH= Follicle Stimulation Hormone; AMH=anti-Müllerian hormone; AFC=Antral Follicle Count; HR= Hazard Ratio calculated using Cox regression analysis and reflecting the % increase in chance of MP occurring during follow up per [unit] decrease/increase of the variable added to the model); 95% CI=95% Confidence Interval; C-Statistic=calculated using Cox regression analysis and thus reflecting the ability of the model used to discriminate between women with a short or long time interval to menopause

The univariable analysis indicated that age upon initial screening, AMH, FSH and AFC were all significantly correlated with time to menopause. The C-statistics were strong for age at baseline (0.85), reasonable for AMH and AFC (0.78 and 0.79 respectively) and poor for FSH (0.66).

In the multivariable analysis correcting for age at the time of ORT measurement, only AMH remained independently capable of predicting time to menopause and improved the C-statistic to 0.86. In the multivariable analysis correcting for both age and smoking at baseline or current smoking, the predictive capacity of AMH remained statistically significant, with an improved C statistic of 0.87. Since AMH alone was independently related to the time to menopause when

corrected for baseline age, the (non-) proportionality for this ORT was assessed and depicted in Figure 5.2.

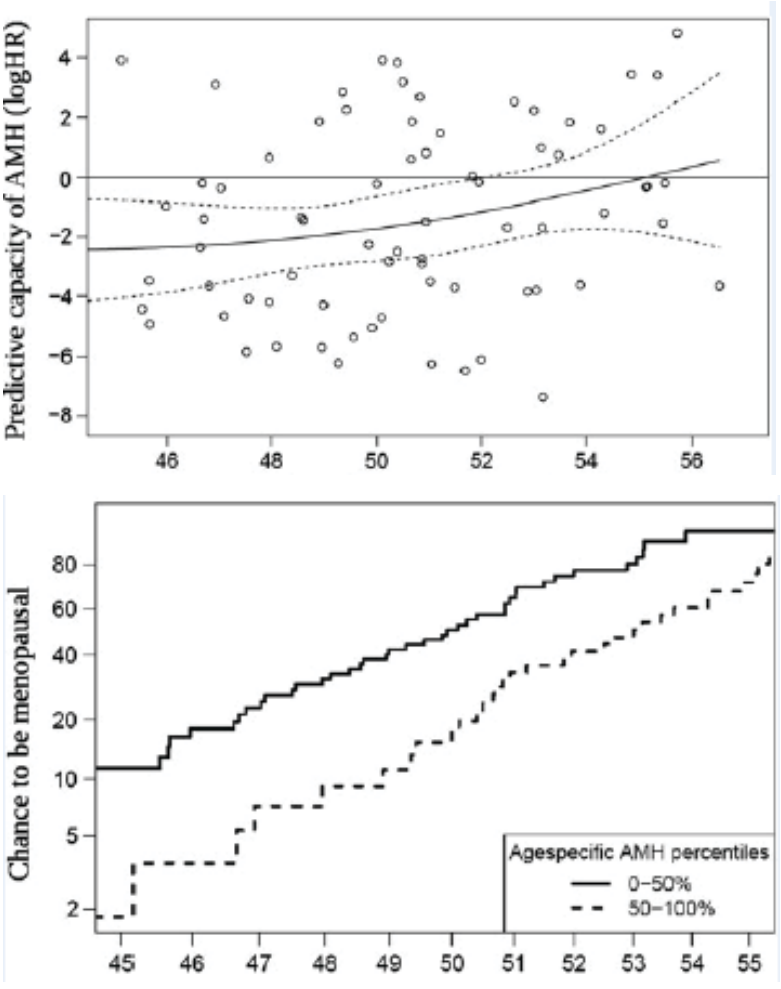


Figure 5.2: (non)-proportionality of the predictive capacity of AMH in menopause prediction

Upper panel: lines represent the AMH Q-score (i.e. the predictive capacity) with increasing age. Dots represent menopausal ages. The X-axis represents female age. A reduced predictive capacity of AMH with increasing age is reflected by the line indicating the level of prediction crossing zero (zero=no effect).

Lower panel: Kaplan Meier curves reflecting the chance of becoming menopausal for the low age-specific AMH percentiles categories (p<50) and the high age-specific AMH percentiles (p>50). The X-axis represents female age.

The upper panel indicates whether the predictive capacity of AMH remains constant throughout time. The predictive effect of AMH is less strong (i.e. the

line approaches zero) with increasing age. This is further depicted in the lower panel, in which the Kaplan-Meier curves for the low age-specific AMH percentiles (p<50) and the high percentiles (p>50) are depicted. With increasing age, the Kaplan Meier curves for both percentile groups approach each other thus showing a reduced discriminative effect of AMH in the prediction of early versus late menopause.

Lastly, a nomogram (Figure 5.3) predicting individual age at menopause based upon AMH and age at baseline was constructed using a Weibull survival model. As depicted, women were assigned an age-specific AMH percentile category (i.e. p5-p95; upper panel), next, the distribution for age at menopause was plotted using these percentile lines (lower panel). This nomogram thus shows age specific AMH percentile lines and their corresponding range of predicted ages at menopause.

As previously discussed, two additional analyses were performed. The first additional analysis assessed whether the use of two different AMH assays influenced our main outcome. AMH assay type was added to the multivariable Cox regression model next to AMH and baseline age. AMH assay did not affect the results of AMH based menopause prediction. Secondly, a subgroup analysis was conducted on complete data (i.e. without missing data) in order to assess if imputation of missing values could have affected the outcomes. Imputation of missing data did not affect the outcome measured.

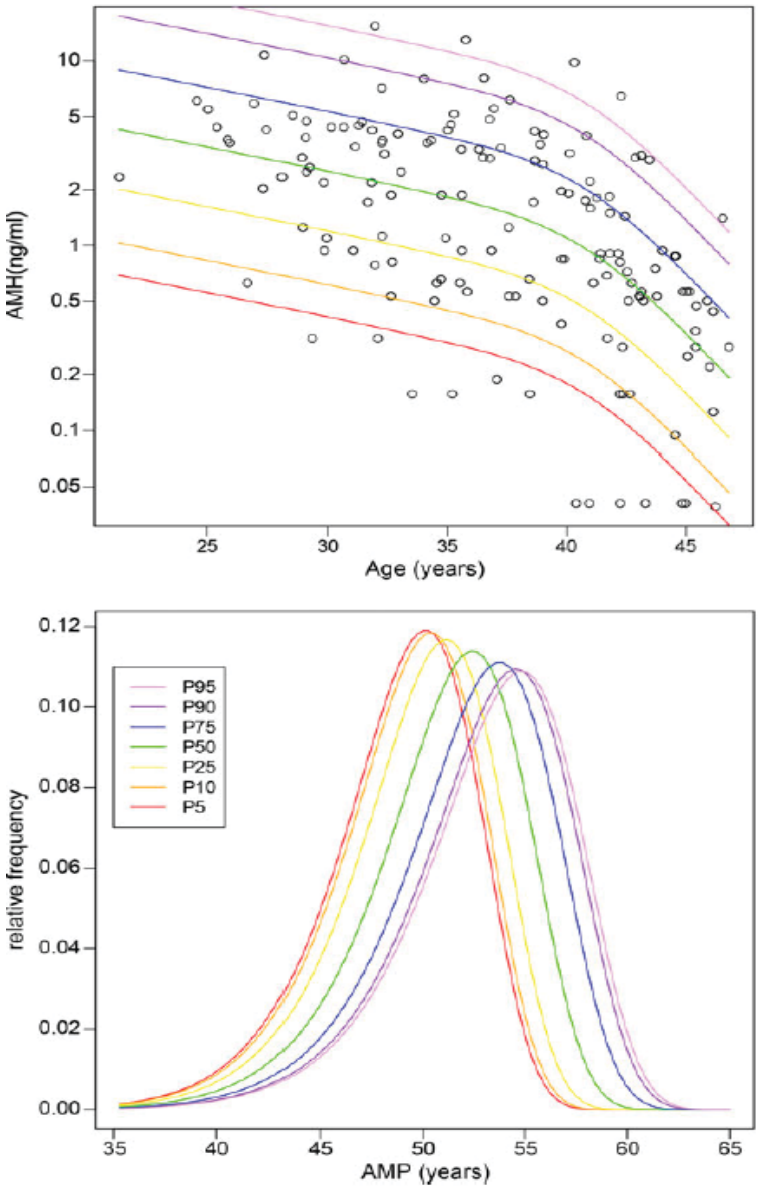


Figure 5.3: nomogram for age specific AMH percentiles

Upper panel: lines represent the different percentiles according to age specific AMH value. Lower panel: estimated distribution of age at menopause for each percentile line. The left side of the figure for instance represents low age specific AMH levels and thus shows that mean age at menopause (plus the age interval) shifts towards a younger age, while the right side represents a relatively high age specific AMH value and thus a later age at menopause.

Discussion

In this extended prospective follow up study, with a mean follow up time of 14 years, we aimed to determine whether the role for endocrine or ultrasound markers in the prediction of age at natural menopause could be demonstrated. First, the present study demonstrates that the true capacity of forecasting age at menopause by using AMH may only come from studies with an even more extended follow up time and using larger numbers of individuals that are followed from their early twenties. Still, the current study has underlined our previous conclusion [18] that AMH is capable of predicting menopause independent of baseline age. However, this predictive capacity became less strong with increasing age. Furthermore, predictions of age at menopause based on age specific AMH levels did not cover the full distribution of normal ages at menopause, and also lacked a high level of precision. As for other ovarian reserve tests of interest (i.e. AFC and FSH), no predictive capacity for age at natural menopause in models next to female age could be demonstrated. As of today, seven prospective databases have provided analyses on menopause prediction using AMH. [16-19,21-24,42]. Moreover, cross-linking studies have emerged comparing the predicted distribution of age at menopause, based on AMH decline models, with observed age at menopause from population based cohorts. [15,58,60] In both the prospective and the cross- linking analyses, the role for AMH as a quantitative marker of ovarian reserve in the individualization of menopause prediction has become consistently apparent. As such, our extended follow up data corroborates these notions.

Literature regarding AFC based menopause prediction is somewhat ambiguous. In a cross- linkage study, Broekmans et al. [14] demonstrated a link between declining antral follicle counts and reproductively significant events such as menopause, exemplified by the high degree of similarity among predicted and observed menopausal age distributions. In line with these findings, Wellons et al. [20] demonstrated the AFC capable of predicting time to menopause when a dichotomized value (≤4 or >4) was added to a model next to female age in a prospective cohort study. In contrast with these findings, our previous analysis published by Broer et al. [18] presented findings comparable to those in the present analysis, with the AFC losing its predictive value when female age at baseline was introduced into the model.

Lastly, two papers reported on FSH based menopause prediction. Freeman et al. [19] stated that FSH became non-significant when AMH was added to the prediction model. In line with the present results, the study by Broer et al.[18]

demonstrated that FSH lost its predictive value in a model next to female age. Taken together, the results of the present study and published data strongly suggest that AMH is the better candidate when it comes to providing individual corrections on the general forecast for the occurrence of menopause in women with a normal regular cycle.

As previously discussed, the final menstrual period occurs between age 40-60. [102] In our previous study [18], model predictions of ANM incorporated nearly all menopausal ages (range 42-60 years) within this known age range. In the present analysis however, predicted ages at menopause ranged from 43.6-58.2 years. This narrowed prediction range, when compared to our previous analysis, results from a more limited predictive capacity of AMH which was observed with increasing age. In our previous publication (data not shown), the predictive effect of AMH was constant over time. However, when looking at Figure 5.2 it becomes clear that, in the present analysis, the predictive capacity of AMH became less strong with increasing age and that the overall predictive effect, which is the average effect across all ages in the data, became closer to zero. This finding seems in accordance with a recent publication from the CARDIA project [24]. In this publication too, a non- proportional hazard was observed for the association between AMH based menopause prediction and time. This was resolved by reporting the average hazard ratio in three-year intervals. A reducing hazard ratio over time was presented (0-3 years 6.1;3-6 years 2.2; 6-9 years 1.8) again confirming our observation of a reduced predictive effect of AMH with increasing age.

One can only speculate about the origin of the declining predictive capacity of AMH with increasing age and the corresponding centering of the age at menopause predictions around the general mean. It is well known in statistical literature that non-proportional effects result from unobserved heterogeneity present within the population. Within each AMH percentile group, women vary in their individual age-specific hazard of becoming menopausal. Therefore, women with the highest hazard of menopause within each percentile group would have reached menopause at the first follow up round. The average hazard of menopause at the second follow up round will thus be lower than the hazard of the whole group at the first round. The lower the AMH percentile, the stronger this prediction reducing effect will be. As a consequence, the average hazard of menopause in low age-specific AMH percentile groups is relatively more reduced than in the higher percentile groups, as depicted in Figure 5.2.

Another explanation for the fact that age at menopause predictions do not cover the full age spectrum of menopause can be found in the inclusion criterion

'regular cycle'. This criterion has filtered out women experiencing an irregular cycle due to imminent ovarian failure at a young age. Women experiencing menopause relatively late are also underrepresented since most of these women are premenopausal at the time of this current follow up. Lastly, the relative low frequency of extreme menopausal ages makes these women prone to underrepresentation in any cohort, especially in a cohort with a limited sample size of 155 women.

The limited potential for the prediction of the extreme ages at menopause is an obstacle present in all papers presenting individual predictions of ANM [16-19,21,22]. In these papers too, this most likely resulted from the fact that having a regular cycle was an inclusion criterion. Additionally, the relative small sample sizes have resulted in the underrepresentation of extreme menopausal ages.

Moreover, Figure 5.3 shows that prediction intervals accompanying the mean predicted age at menopause remain unsatisfactory wide, a finding that has been reported in other studies [21]. On the one hand, this might be due to imprecision of the predictive capacity of AMH in general or of the assay applied specifically. As discussed, the AMH assays used in the present analysis are currently outdated. With the introduction of newer, more stable and more precise AMH assays, a narrower prediction interval might be within reach. One the other hand, not all factors contributing to age at menopause might be identified as of today, making predictions less precise. Further research is needed to identify factors capable of narrowing the interval of predicted ages at menopause whilst repeated measurements of AMH, and the use of newer AMH assays, could undo the influence of intra- or intercycle AMH variability [28] and could even fade out laboratory specific variability [97].

Research demonstrated current smoking to result in lower age specific AMH [73] and AFC [113] levels, in higher FSH levels [113] and an earlier age at menopause [96,114]. However, unlike in the present study, no other prospective research is available in which corrections are made for both smoking at baseline and current smoking (i.e. smoking at last follow up). Our data demonstrated AMH to remain significantly capable of predicting time to menopause in models next to female age and current or baseline smoking. The C-statistics originating from the models containing smoking behavior next to age were slightly adjusted from a model correcting for age alone. Both the AFC and FSH lost their predictive capacity in models next to female age and smoking behavior. The question remains through which pathway smoking behavior might influence age at menopause and ORT levels. One could postulate that smoking induces an accelerated decline of the primordial follicle pool which could account for

both an earlier age at menopause and for lower levels of age specific ovarian reserve tests. This explains why models predicting age at natural menopause incorporating smoking behavior are inconclusive, since smoking effects on ovarian reserve tests will automatically become incorporated in the prediction models.

As indicated above, we have shown AMH to significantly aid in the prediction of menopause in a model next to age, conforming our previous predictions [18]. However, our prediction model failed to predict the extreme menopausal ages and prediction intervals remained unsatisfactory broad. Additionally, the predictive capacity of AMH decreased with increasing age. Currently, AMH is often used in clinical practice to identify women at risk for early decline of fertility, thus allowing early treatment or oocyte preservation. The present paper however, jeopardizes such a policy as we demonstrate that all papers regarding AMH based age at menopause predictions, including our current analysis, could not predict early age at menopause since cohort sizes are relatively small and early age at menopause is a rather rare event.

Therefore, we feel there is currently no ground for AMH based menopause prediction in the day-to-day clinical practice.

The ideal prediction model would not only cover the extreme age ranges, but would also produce a narrow prediction interval. As previously stated, prolonging follow up or including younger women with an irregular cycle might resolve the first problem, whereas repeated AMH measurements with newer AMH assays and adding currently unknown predictors, to the model could narrow prediction intervals. Factors of interest are mother's age of natural menopause [22], genetic markers or lifestyle factors such as BMI [73].

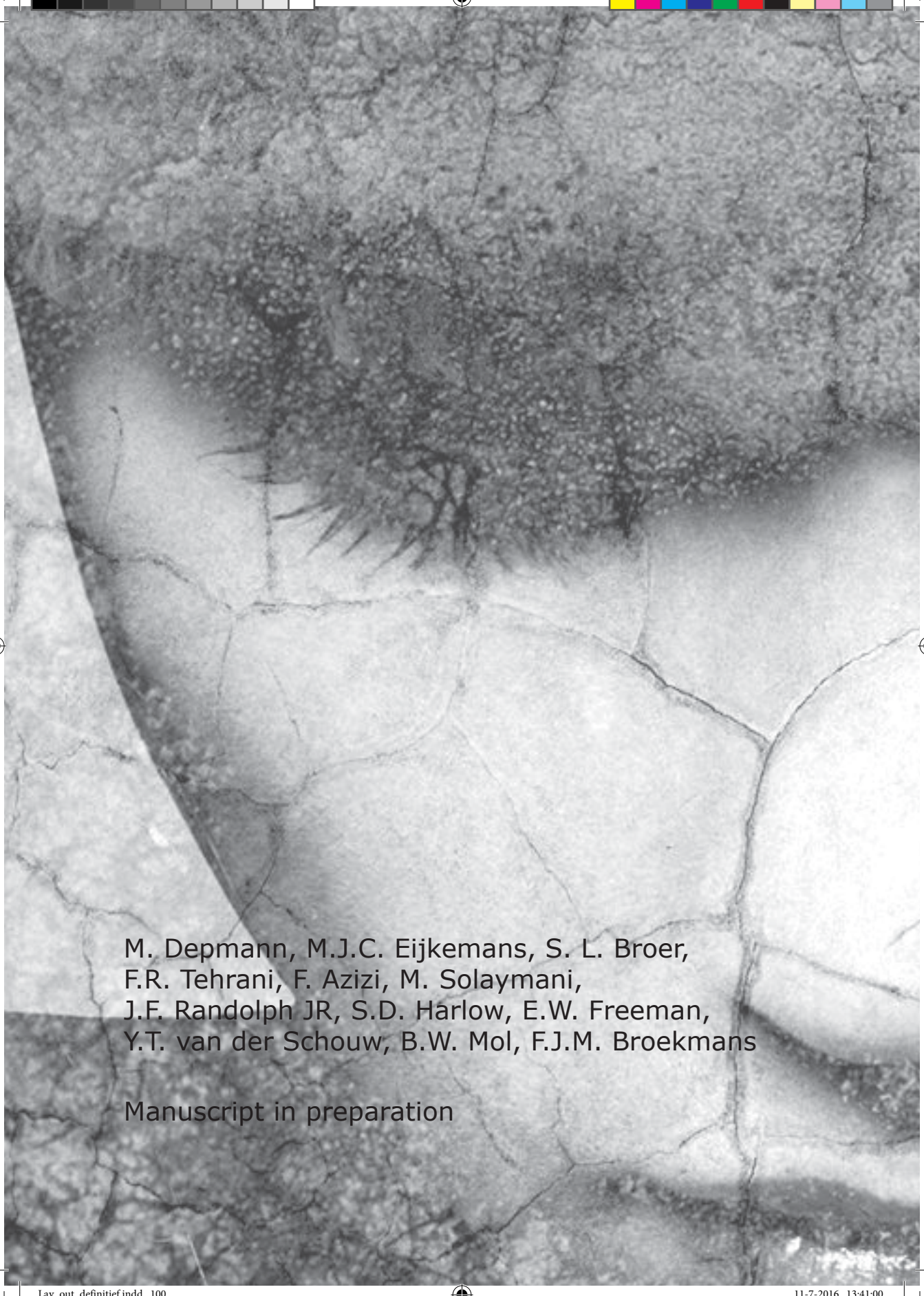
Furthermore, our hypothesis of a variable predictive AMH effect due to non-proportionality, needs further investigation and could be confirmed when other databases researching AMH based menopause prediction reach complete follow up. Additionally, in order to adequately predict menopause for the extreme ages, larger databases, or a meta-analysis of existing datasets, is required to cover these rare events.

In future research, analysis such as the present should be performed in large cohorts of women included at a young age. Women should be included irrespective of cycle regularity and should provide repeated AMH measurements. Lastly, new variables, such as mother's ANM and genetic- and lifestyle factors, need to be incorporated in the model used.

Conclusion

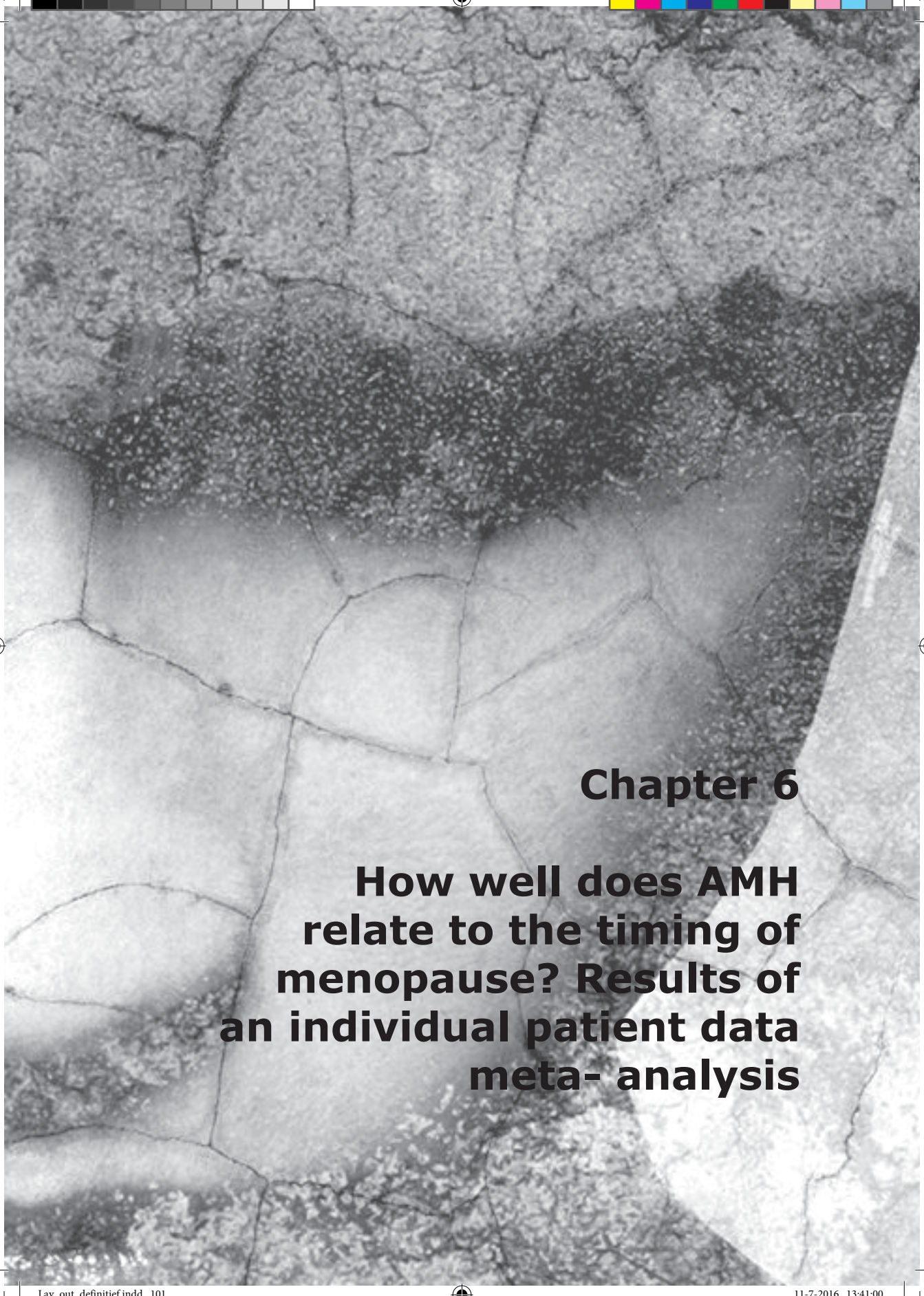
In line with our previous data [18] we have shown that age specific AMH is significantly capable of predicting time to menopause making this marker a possible candidate in the preventive management of age related infertility. However, a reduced predictive effect of AMH was observed with increasing age, and the variation ranges around predicted age of menopause remained wide and did not cover the extreme age ranges of menopause. Whether forecasting of age at menopause and thus the prediction of the end of natural fertility will become of use in reproductive health counseling remains to be firmly established by future research. AMH should therefore not be used to make individual forecasts of menopause, or the fertile life span, in the day-to-day clinical practice based on the current research.

5



M. Depmann, M.J.C. Eijkemans, S. L. Broer,
F.R. Tehrani, F. Azizi, M. Solaymani,
J.F. Randolph JR, S.D. Harlow, E.W. Freeman,
Y.T. van der Schouw, B.W. Mol, F.J.M. Broekmans

Manuscript in preparation



Chapter 6

How well does AMH relate to the timing of menopause? Results of an individual patient data meta- analysis

Abstract

Introduction: Research regarding the anti-Müllerian hormone (AMH) based forecasting of age at menopause has been promising. Still, individualized predictions of age at menopause appeared not to cover the full age range of menopause. Also, prediction intervals have shown to be wide and the predictive capacity of AMH tended to substantially drop with increasing age. We hoped to resolve the above mentioned problems in the interrelation between age specific AMH and timing of menopause by pooling individual data, creating the largest cohort to date.

Material & Methods: We performed a systematic literature search identifying all prospective follow up cohorts addressing the relation between AMH and menopause. Authors of publications that met the inclusion criteria were invited to share the published data in an individual patient data (IPD) meta-analysis. After harmonizing of the various databases, a check of heterogeneity arising from AMH assay used and study of origin was performed. The outcome of this check determined if age-specific AMH percentile categories, used in further analyses, were composed within the study of origin or in the pooled dataset. Uni- and multivariable Cox regression analyses were performed assessing time to menopause using female age, AMH and adjusting for study of origin. The C-statistics were calculated in order to assess the model's capacity to discriminate between women with a short and long time to menopause. Using age-specific AMH percentile groups, an individual age at menopause prediction was performed using a Weibull regression model. Lastly, a check of non-proportionality of the predictive effect of AMH was performed in order to assess whether the predictive effect of AMH on age at menopause would remain stable at higher age.

Results: Of the seven available cohorts known today five authors were willing to share data. The check of heterogeneity resulted in the fact that age-specific AMH percentile groups were composed within the study of origin. Univariable Cox regression analysis assessing time to menopause based on female age and correcting for study of origin demonstrated female age to be a significant predictor of time to menopause (HR1.25; 95% CI 1.23-1.26; C-statistic 0.83). AMH had a significant capacity in the prediction of menopause in the multivariable Cox regression model next to female age (HR 0.60; 95% CI 0.56-0.65). However, the added value of AMH on top of female age was poor (C- statistic 0.85). The individual predictions of age at menopause, using the Weibull survival model, did cover the majority of the age spectrum known for

menopause, but especially younger ages at menopause were not included. Precision of the forecasts remained quite poor, as the age intervals surrounding individual predictions of age at menopause remained wide. Finally, the reduction in the predictive effect of AMH with increasing age was clearly noted in this pooled data.
Conclusion: Although AMH aids in predicting the timing of age at menopause in models next to female age, it does not help in the identification of very early and very late menopause. These findings do not support the use of AMH in the prediction of age at menopause.

Introduction

In recent years more data is becoming available regarding the prediction of age at natural menopause based on peripheral anti-Müllerian hormone (AMH) levels [15-19,21-24,42,115,116]. It is suggested that these menopause forecasts can be extrapolated in order to predict the end of natural fertility. This extrapolation is based on the hypothesis that a fixed interval exists between the end of natural fertility and the onset of menopause [4]. Both events occur within a broad 20-year window, menopause between the ages 40 and 60, the end of natural fertility 10 years prior to this event. Considering this broad age interval for the end of natural fertility, it is understandable that a growing number of women are facing age related infertility at the time they hope to conceive [3]. This is especially the case since an increasing number of women are delaying childbearing due to an ever growing participation in the workforce and a wide availability of effective contraceptives [13].
In order to prevent age related infertility, there is need for a marker accurately assessing the limits of a woman's fertile lifespan. Currently, there is no marker available reflecting actual fertility, and the end of natural fertility is an event that goes by unnoticed, as the majority of women will still experience regular menstrual cycles at that time period in their lives. Thus, based on the presumed fixed interval between menopause and the end of natural fertility, menopause predictions are extrapolated to forecast the end of a women's individual natural fertility.
Menopause occurs when the number of primordial follicles falls below a critical threshold at which the ovary loses the ability to produce mature oocytes and is therefore unable to maintain a menstrual cycle [103]. Derivatives of the number of primordial follicles have been researched in order to assess their capacity in the prediction of menopause. Of these derivatives of the true ovarian reserve (i.e.

the primordial follicle pool) currently researched (the antral follicle count, follicle stimulation hormone and AMH), AMH has shown to be the most promising one. The studies providing longitudinal data on AMH based menopause prediction available today [16-19,21,23,24,86,116] all suggest that AMH is significantly capable of predicting age at natural menopause, and that AMH adds valuable information to the information obtained from the combination of age and cycle status. As unanimous as these suggestions may be, recently three problems in AMH based menopause prediction became apparent [115]. The first problem is the fact that predictions do not cover the extreme ages at menopause (i.e. the very early or late menopausal ages). The second is the fact that prediction intervals are rather wide making them possibly irrelevant for clinical usage. And finally, two recent studies [24,116] demonstrated the predictive capacity of AMH to decline with increasing age. It was hypothesized that the first problem emerged from small sample sizes resulting in the underrepresentation of women with either an early or late age at menopause. The lack of prediction of early menopause stems from the rareness of the condition and the lack of late ages at menopause is due to incomplete follow up. The second problem originated from imprecisions in models used, in AMH assays and in implementation of other factors contributing to the variation in age at menopause. The third and last problem was hypothesized to originate from a varying predictive effect of AMH due to non- proportionality of the predictive capacity of AMH. The above mentioned pitfalls in AMH based menopause prediction have prevented the buildup of sufficient evidence for clinical applicability.

We felt that pooling of data, hereby performing analysis on the largest possible cohort to date, might resolve some of the above mentioned problems regarding AMH based menopause prediction. The aim of this study therefore was to perform an individual patient data meta-analysis researching the capacity of AMH in the prediction of age at menopause. All available published data regarding AMH based menopause prediction were searched for and authors invited to participate in this project.

Material & methods

Method of systematic review of literature

In order to detect all papers regarding AMH based age at menopause prediction, we performed three extensive searches in the PubMed, Embase and Cochrane databases. Search criteria and inclusion criteria are listed in appendix 6.1. Searches were confined to papers published up to January 2016 and there

was no language restriction. Potentially eligible studies were studies reporting on the prediction of individual age at natural menopause using anti-Müllerian hormone, and subsequent follow up or retrospective use of stored blood samples. Titles, abstracts and full text papers retrieved were evaluated and, where necessary, consensus discussion determined eligibility of the paper.
Authors of studies that were considered eligible for inclusion in our Individual Patient Data meta-analysis (IPD) were invited to join our project and share their data. When willing to participate, authors received a study protocol, including a data request form, and a collaboration contract.
Data received was reformatted into a single format and variables were uniformly coded and harmonized.

Statistical analyses of Individual Patient Data meta- analysis
Baseline characteristics comparing women per cohort of origin were analyzed using the Kruskal-Wallis or Chi square test. Imputation was performed for cases with missing data on AMH, BMI and baseline smoking. Kaplan-Meier curves depicting the distribution of age at natural menopause were drawn for each cohort.
Next, in order to model individual AMH based time to menopause, the available cohorts needed to be pooled. Moreover, age- specific AMH percentile categories needed to be constructed in order to perform further analyses. Potentially substantial heterogeneity was foreseen arising from different types of AMH assays used in eligible studies, and from the differences present in the design of the various studies. In order to assess the heterogeneity and its sources, two checks of heterogeneity were performed prior to pooling of data.
The first assessment for heterogeneity addresses differences between studies in the age specific measured AMH levels. For this purpose, a regression analysis for AMH in relation to female age, using a flexible spline for non-linearity in the age effect on AMH, was applied. The regression model was performed adding the variable "study of origin", hereby making it possible to assess between-study heterogeneity of AMH levels. If heterogeneity turned out to be significant and relevant, thus making the pooling of raw AMH data unfeasible, age- specific AMH percentiles would be calculated per original dataset. These study- and age- specific AMH percentile categories were then to be used when pooling AMH data. If heterogeneity was not deemed to be an issue, the datasets would be pooled using raw AMH data and age-specific AMH percentiles would be calculated in the complete pooled dataset.
The second assessment for heterogeneity was performed in order to assess

whether time to menopause differed between studies. The Cox model for time to menopause was adjusted for female age, smoking and BMI at baseline, and "study of origin" was added to the model. As an alternative approach, the effect of "study of origin" was added as a frailty term with a Gaussian distribution to the Cox model. Hereby the differences between studies of origin were assumed to follow a Gaussian distribution on the log hazard scale, which was characterized by the between-studies standard deviation.

In order to assess if AMH predicts menopause, a Cox regression analysis assessing time to menopause was performed, this time using current study as a stratified baseline. Using this model, the C-statistics were calculated in order to reflect the capacity of the model to discriminate between women with a short or long time to menopause.

The age-specific AMH percentiles, (either pooled study- and age-specific AMH percentile categories or overall pooled age-specific AMH percentile categories) were used to calculate individual age at natural menopause using a Weibull survival model with age at the time axis. Left truncation at T1, the age at first inclusion, was used and the endpoint was either censoring at age at last follow-up or menopause at the age at natural menopause. The parametric Weibull survival model produces smoother curves when using age-specific AMH percentile lines than the semi-parametric Cox model. Therefore, this model is preferred over a Cox model. Per age-specific AMH percentile, the median predicted age at menopause and the corresponding age ranges were depicted in a nomogram.

Lastly, a check was performed of non- proportionality of the effect of AMH on the risk of becoming menopausal. This was done by plotting the Schoenfeld residuals of the model against age and adding a smoothed line of the time varying regression coefficient of the AMH percentiles (Beta(t)).

Data analysis was performed using SPSS 20 (SPSS Inc., Chicago, IL) and R version 3.1.3 (R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/). .

Results

Systematic review of literature and data acquisition

We performed a literature search in the PubMed, Embase and Cochrane databases. All papers reporting longitudinal data on AMH based menopause prediction were considered eligible, search terms and inclusion criteria are

listed in appendix 6.1. As depicted in Figure 6.1, 10 full text papers reported analyses regarding AMH based age at menopause (or time to menopause) prediction and were thus potentially eligible for inclusion in this IPD [17-19,21-24,86,116,117].

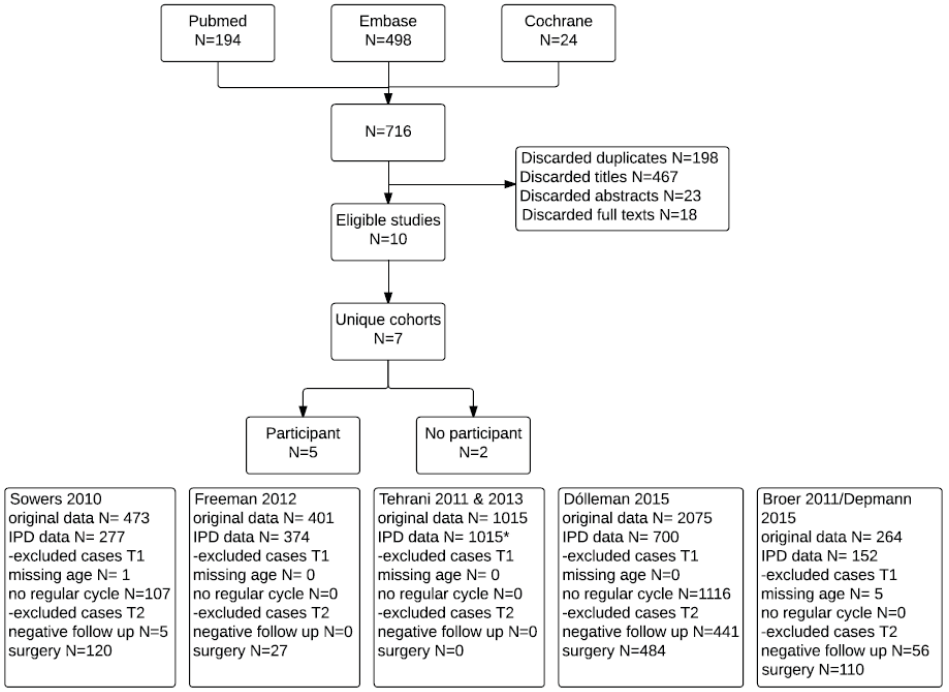


Figure 6.1: Flowchart of search for eligible studies to be included in this IPD

Excluded cases T1= reasons for exclusion on baseline criteria, Excluded T2 reasons for exclusion at last follow up. *Tehrani data applied the in- and exclusion criteria on original data, therefore no additional cases were lost.

These 10 papers reported on the prediction of individual age at menopause in seven different cohorts. Corresponding authors of the seven eligible cohorts were contacted and asked to join this IPD meta- analysis. Data was shared by five corresponding authors. Authors of the remaining two papers have not contributed data in spite of repeated attempts [24,86]. In Figure 6.1, the number of cases present in the source data, the number of cases lost due to in- and exclusion criteria and the final number of cases included in the IPD dataset were depicted.

Individual Patient Data meta- analysis

The five cohorts participating in this IPD provided data on 2,518 women all

experiencing a regular cycle at the time of AMH measurement. In order to match the largest cohort, women were excluded if they underwent surgery on uterus or ovaries during follow up. Moreover women using hormones either at baseline or at last follow up moment were excluded, since no cycle state could be determined for these women. Baseline characteristics of women available for analysis were depicted per cohort of origin in Table 6.1. As depicted in this table, statistically significant differences were present in all baseline parameters demonstrating the dissimilarity between the available cohorts. For women with missing data on AMH (n=20), BMI (n=97) and baseline smoking (N=107) single imputation was performed.

In Figure 6.2, Kaplan-Meier curves depicting the distribution of age at natural menopause were depicted for each of the participating cohorts. As illustrated, the distribution of menopausal ages is comparable although one study consistently reported a later age at menopause.

As stated, two checks for heterogeneity arising from different types of AMH assays used in eligible studies and from the differences in time to menopause present between the studies were performed. The first assessed the potential heterogeneity arising from the AMH assays used in the different studies contributing to this IPD. The results of this regression assessing AMH in relation to female age per study of origin are depicted in Figure 6.3. From this figure, it becomes clear that the shape of the curves is similar for al cohorts included in this IPD. However, although similarity in AMH levels is present between three of the five studies available for analysis, two studies differ in AMH levels measured for all ages. It was therefore decided to construct age-specific AMH percentiles within each cohort of origin for further analyses.

	Freeman (N=374)	Sowers (N=277)	Tehrani (N=1015)	Dólleman (N=700)	Broer/ Depmann (N= 152)	P-value
Characteristics baseline						
Age (yrs, SD)	41.4 (± 3.5)	38.2 (± 4.8)	36.7 (± 7.5)	41.7 (± 5.9)	37.2 (± 6.0)	<0.01
Age menarche (yrs, SD)	12.7 (± 1.8)	12.6 (± 1.3)	NA	13.2 (± 1.4)	13.0 (± 1.9)	<0.01
AMH (ng/ml, SD)	1.07 (± 1.2)	3.1 (± 2.8)	1.6 (± 1.8)	1.1 (± 1.4)	2.3 (± 2.6)	<0.01
AMH assay (%)	Gen II (100)	Gen II[a] (7.9) Gen II[b] (78.7) DSL (13.4)	Gen II (100)	Gen II (100)	Gen II (100)[c]	
BMI (kg/m², SD)	29.2 (± 7.5)	26.3 (± 5.6)	27.0 (± 4.6)	24.3 (± 3.9)	23.8 (± 3.7)	<0.01
Smoking						
Yes	38.6 %	27.1 %	NA	31.1 %	31.9 %	
Ever	NA	5.9 %	NA	NA	NA	
Characteristics at last follow up						
Cycle state						
MP	50.3 %	63.9 %	37.5%	57.1 %	52.0 %	
MT	26.5 %	9.0 %	0.0 %	19.7 %	20.4 %	
Regular	23.2 %	27.1 %	62.5 %	23.2 %	27.6 %	
BMI (kg/m², SD)	29.4 (± 7.5)	28.4 (± 6.4)	NA	NA	25.0 (± 4.1)	<0.01
Smoking						
Yes	38 %	19 %	NA	NA	8.6 %	
Ever	NA	20.9 %	4.6 %	NA	NA	

Table 6.1 Baseline data of IPD cohort: [a] Gen II assay measured in 2015, [b] Gen II assay measured in 2013, [c] conversion factor translating to Gen II assay.

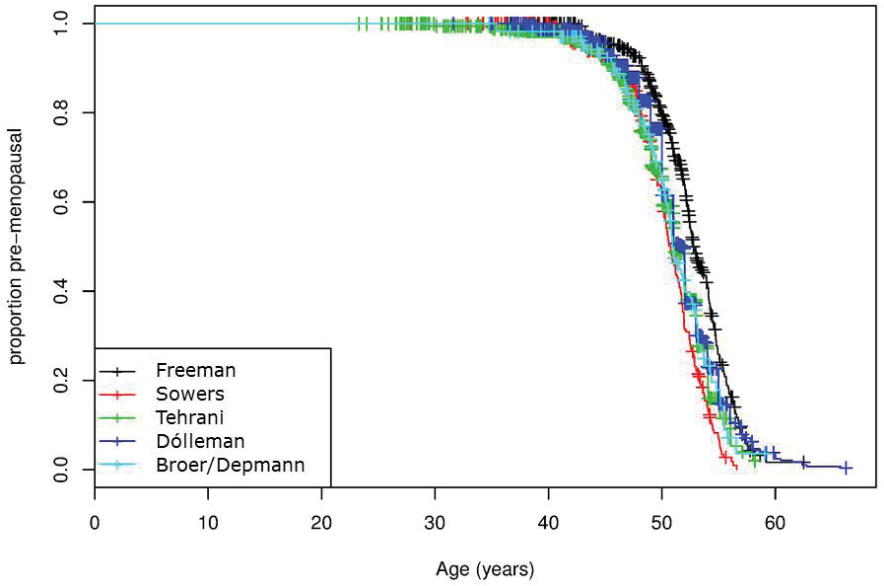


Figure 6.2: Distribution of age at natural menopause

The Y-axis represents the proportion of women that are premenopausal, the X-axis represents female age. Note that in this IPD cohort, menopause does not occur until after the age of approximately 40 years.

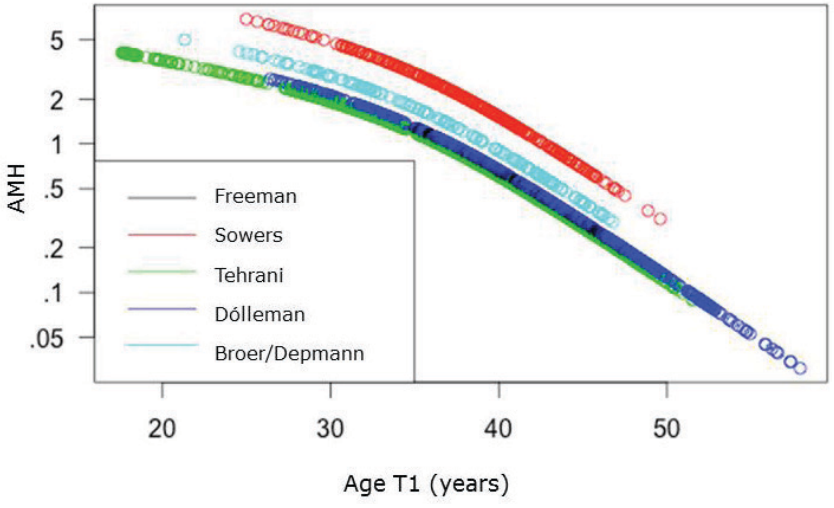


Figure 6.3 Regression curves of AMH per study of origin

The Y-axis reflects AMH levels, the X-axis age at AMH measurement. A flexible spline fitted for $_{log}$AMH ~ Age at baseline was performed per study of origin. Note that there is no variation between the shapes of the different curves, only in overall level of AMH. In order to correct for the difference in AMH levels, further analysis were performed using age- specific AMH percentiles created within the study of origin.

The second check of heterogeneity assessed the heterogeneity arising from study of origin in time to menopause, using two different analyses. From this check it became clear that significant heterogeneity arose from study of origin (variance of random effects 0.128, $p<0.0001$). Although the difference between studies was significant, the effect was small. Therefore, by using study- and age-specific AMH percentile curves (as explained above) pooling of data was deemed possible.

Univariable Cox regression analysis demonstrated female age to be an important predictor of time to menopause (Hazard Ratio (HR) 1.25; 95% Confidence Interval (CI) 1.23-1.26). The C- statistic of the univariable model was 0.83. The results of the multivariable model, where AMH based time to menopause prediction was performed correcting for female age, demonstrated AMH to be a significant predictor of time to menopause (HR 0.60; 95% CI 0.56-0.65). However, as reflected by a minor rise in the C-statistic (c = 0.85), the added value on top of female age was quite poor.

Age- specific AMH percentiles were calculated per dataset of origin and a age at menopause analysis was performed based on these age- specific AMH percentiles using a Weibull survival model (Figure 6.4). From this figure it becomes clear that for women with a low age- specific AMH percentile, menopause occurs at a relatively early age, whilst women with high age- specific AMH values experience menopause later in life. This figure clearly reveals that at younger age a more solid AMH effect is observed, whilst this effect is negligibly small at older ages. Thereby, at higher ages, the chance of becoming menopausal is no longer predicted by a woman's AMH percentile category.

Furthermore, a check of non-proportionality of AMH based age at menopause prediction was performed and outlined in Figure 6.5.

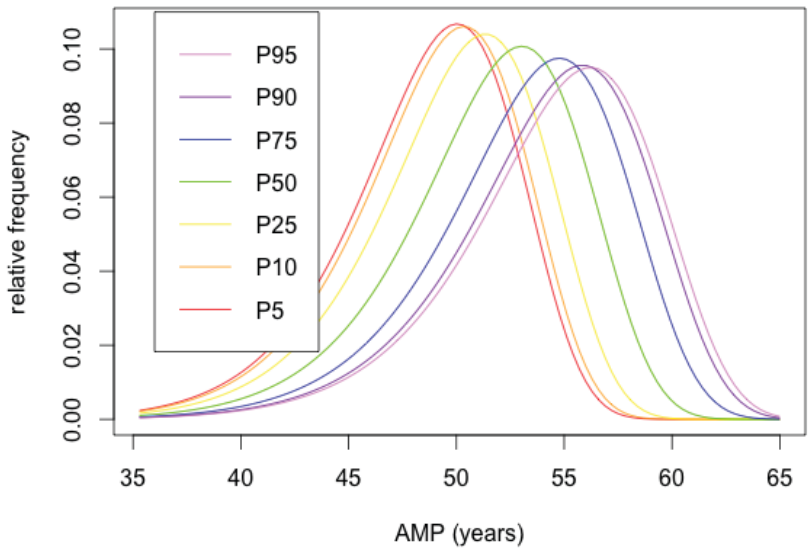


Figure 6.4: Nomogram depicting the estimated distribution of age at menopause per percentile line. The left side of the figure represents low age specific AMH levels and a corresponding mean age at menopause at a younger age. The right side of the figure represents high age specific AMH levels and corresponding later ages at menopause. From this figure it becomes clear that for women in the lowest age- specific AMH category, mean age at menopause shifts towards a younger age. AMP= age at menopause.

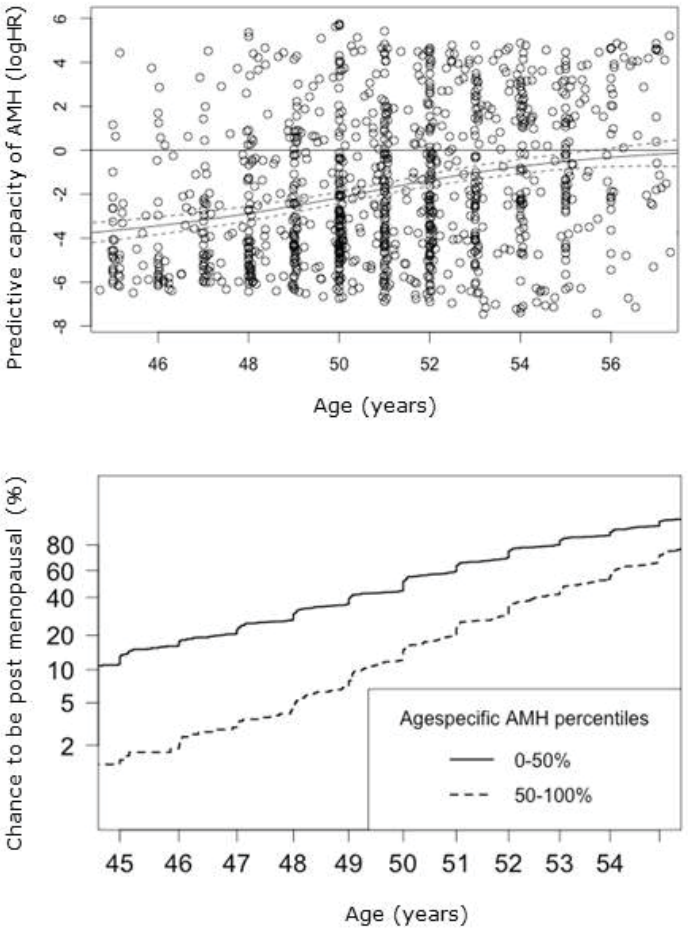


Figure 6.5: (Non-) proportionality of the predictive capacity of AMH in menopause prediction

Upper panel: lines represent the AMH Q-score (i.e. the predictive capacity) with increasing female age. Dots represent menopausal ages. A reduced predictive AMH effect is observed over time. With an increasing age, a strong decrease in the AMH effect is observed. This is reflected by the line starting at -4 (very strong effect) and eventually reaching 0 (no effect). This means that for instance for a woman aged 46 years, and not yet menopausal, the predictive effect of AMH for her instantaneous hazard of becoming menopausal is very strong. This effect no longer present if she reaches the age of 57 and is not yet menopausal.

Lower panel: Kaplan Meier curves reflecting the chance of becoming menopausal for the low age-specific AMH categories (p<50) and the high age-specific AMH categories (p>50). Due to a reduced hazard of becoming menopausal, which is stronger for the low age-specific AMH group, these lines approach each other at higher female ages.

Discussion

In this individual patient data meta- analysis, we confirmed the principal capacity of AMH to aid in the prediction of individual age at menopause. We also revealed that a larger sample size does not resolve the major problems present in AMH based menopause predictions. In the present analysis the prediction of age at natural menopause did cover women experiencing menopause at a later than average age, but identification of women experiencing early age at menopause did not improve. Moreover, prediction intervals remained wide and covered a range that makes clinical application troublesome. Finally, the role for AMH as added predictor next to female age becomes much weaker with increasing age.

As of today, seven prospective cohorts are available investigating the capacity of AMH in the prediction of age at menopause [18,19,21,23,24,86,116-118]. When looking more closely at the cohorts addressing AMH based menopause prediction, the three above mentioned pitfalls became apparent.

The first problem relates to the incapacity of AMH in the identification of the extreme ages at menopause, which was previously noted in other studies [17,18,116,119] and discussed in an extensive literature review [115]. It was postulated that the inadequacy in the prediction of the very early age at menopause stems from the rareness of the event combined with small datasets discussing the subject. However, when reviewing Figure 6.4, it becomes clear that a larger dataset did not solve this problem as the left side of the figure, where women experiencing an early age at menopause should be represented, is still empty. Another hypothesis posted in this review [115] regarding the origin of lack of predicting a young age of menopause was related to inclusion criteria and is perhaps the true reason for this problem. In our IPD, 'having a regular cycle' was an inclusion criterion in order to uniform data. However, by applying this inclusion criterion, women destined for early menopause expressed by the presence of irregular cycles at a young age, were never included in the first place. This is depicted in Figure 6.2, where it becomes clear that few women experience menopause at an early age in our cohort.

Upping numbers by performing an IPD did result in a more adequate prediction of the late menopausal ages. This is reflected by the fact that for women in the highest age specific AMH percentile category their predicted age interval for the occurrence of menopause ranged towards a forecasted mean age of ~56.5 years.

The second problem observed in AMH based menopause prediction was the fact that the prediction intervals surrounding the predicted age at menopause are wide, making them clinically inapplicable. It was hypothesized [115] that this originates from imprecisions resulting from the AMH assay used. Moreover, the broad prediction age accompanying the mean predicted age at menopause displayed in Figure 6.4 is suggestive of the fact that not all variation in age at menopause can be captured by AMH. Adding other factors potentially predictive for age at menopause, such as lifestyle factors like BMI or smoking or genetic factors, possibly represented by mother's age at natural menopause, might narrow this prediction interval. Unfortunately we could not incorporate all these factors, since, aside from BMI and smoking habits at baseline, this data was not available in most cohorts. Regarding the imprecision arising from the AMH assay used, when combining five cohorts, assay related problems increase rather than decrease. In the current dataset, four cohorts used the Gen II assay (Beckham Coulter Inc, Fullerton, CA), or a conversion factor leading to Gen II levels, for measurement of AMH. One cohort measured AMH using the Gen II assay at two different moments and used the enzyme immunometric assay (Diagnostic System Laboratories Inc) in small percentage of cases.

It is known that the older AMH assays available [28], which have been used in some cohorts [116,117], experienced stability problems resulting in the fact that AMH levels measured using these assays are potentially less accurate. And even though the relatively new Gen II assay was used in most of the contributing cohorts, substantial variety might occur due to the difference in timing of AMH sampling in relation to the menstrual cycle, and from the fact that each laboratory measuring AMH has its own value ranges ad calibrations [28]. We have attempted to bypass this problem by performing age at menopause prediction using age- specific AMH percentiles calculated within the dataset of origin. When evaluating Figure 6.2, depicting AMH levels per study of origin, it becomes clear that this method is most likely quite solid. As depicted, all studies follow the same distribution of AMH, variation between studies stems from variation in AMH levels. It is probably not coincidental that the two curves deviating in AMH levels are composed of the two datasets that used different AMH assays compared to the other. It is the variation in AMH levels that is undone when using study- and age specific AMH percentile lines.

Lastly, another potential source of AMH imprecision arises from the timing of AMH sampling in relation to the menstrual cycle. It is known that a significant fluctuation in AMH levels is present throughout the menstrual cycle [69] depending on female age. The datasets contributing to this IPD did not show uniformity

in timing of AMH sampling in relation to the menstrual cycle, and therefore timing remains a potential source of imprecision and may be further addressed in future studies. The third problem which recently became apparent [24,116] is the fact that AMH loses its predictive capacity with increasing age. In the current analysis we therefore performed a non- proportionality analysis for the predictive effect of AMH (Figure 6.4) and confirmed this interesting finding. When looking at the left side the upper panel in this figure, a very strong AMH effect is observed. This implies that for young women, AMH has predictive value. This predictive capacity is completely lost for older women. This is understandable if one keeps in mind that for women above 50-years-of-age the occurrence of menopause within the next ten years is inevitable and no level of AMH could alter this fact.

The lower panel represents the hazard of becoming menopausal for women with the lowest age-specific AMH values and for women with the highest age-specific AMH levels. While for young women these lines are far apart, thus implying a difference in risk of menopause based on AMH levels, this effect is no longer present later in life. It was stated in a recent study [116] that this latter effect is due to heterogeneity present within the group. When dividing women in age- specific AMH percentile categories, it is inevitable that women with the highest risk of becoming menopausal within their category reach menopause early in follow up, whilst women with a relatively low risk of becoming menopausal within their category, will experience menopause later on. The effect of heterogeneity, which is always stronger in women with the highest hazard of an event, results in the fact that the lines in Figure 6.4 approach as follow up time continues.

One of the strengths of this study is the fact that our method of using age-specific AMH percentiles in a Weibull regression analysis is a highly sophisticated way of performing complicated individual age at menopause prediction analyses. In recent years this method has adopted more followers [120] making it currently the method of choice for with the work on AMH based age at menopause predictions. One other feature of the present study lies within the fact that we have pooled (nearly) all available longitudinal data allowing for AMH based menopause prediction. By doing so, we were finally able to assess the predictive capacity of AMH in a cohort unaffected by small numbers.

The limitations of this study however lie in the same area. By pooling data we became unable to include all desired variables in the prediction model since this data was often unavailable in the largest cohorts. Within a cohort one could then opt for imputation of missing values, but imputing a variable for an entire

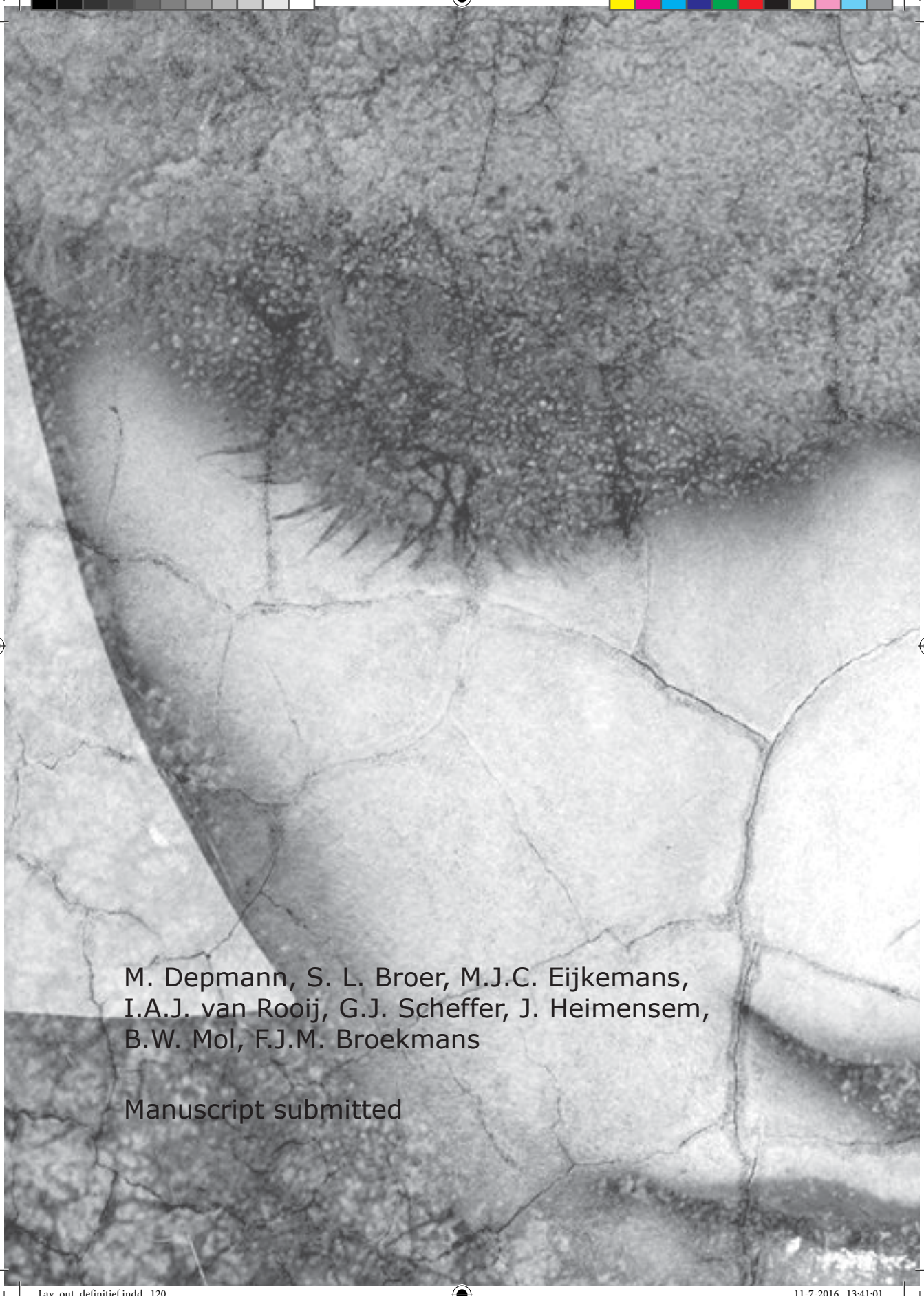
cohort is not possible. It was therefore that the largest cohorts, unfortunately containing the least number of variables, came to dictate the dataset.
This was also the case for some other important data. Since the largest cohorts excluded all women without a regular cycle at baseline or who underwent surgery during follow up, we needed to apply this criterion to all women in order to uniform the data hereby making our data less translatable to the general population.
Another potential limitation of this study is the fact that two eligible cohorts did not share data in spite of repeated attempts to involve the authors [24,86]. Although the contribution of these cohorts is desirable, we feel that combining the available cohorts as performed in the current analysis, provided solid evidence to support our analyses and conclusions.
As for clinical applicability of our results, the main goal of menopause prediction lies within the possibility to extrapolate predictions on age at menopause to predictions regarding the female fertile life span. As previously explained, the rationale behind this extrapolation is the fact that a presumed fixed temporal relation exist between age at menopause and age at the end of natural fertility. Using menopause predictions could therefore prevent age related subfertility by advising women towards timely family planning or cryopreservation of oocytes. However, and not unimportantly, women the prediction of late menopausal age is important as well, since women experiencing menopause at a later age are more prone to breast and intestinal cancer [48]. Unfortunately, when looking at the present results, even though AMH was proven to be a significant predictor of age at menopause, the problems surrounding age at menopause predictions make this marker clinically inapplicable.
With this study we were able to clearly identify directions for future research. The inability of AMH to identify women destined for early menopause might be resolved by including women in a prospective study at a very young age, irrespective of cycle state and allowing for a long follow up. The broad intervals surrounding age at menopause predictions could be narrowed by subduing the effect of AMH assays, this can be done by using a rigid assay methodology which is available soon [95]. Moreover, repeatedly measured AMH at a predetermined time in the menstrual cycle, preferably early follicular, could undo the within cycle fluctuation of AMH observed [68]. Lastly, adding of, currently unknown variables contributing to age at menopause (such as lifestyle factors or genetic factors), could narrow prediction intervals.

In conclusion, in this large individual patient data meta-analysis we have evaluated AMH based menopause prediction using nearly all data currently available worldwide. We have demonstrated that AMH is predictive of menopause, but that in view of the limited additional value of AMH to prediction with female age alone, there is at present no place for the use of AMH in clinical practice.

Supplementary material

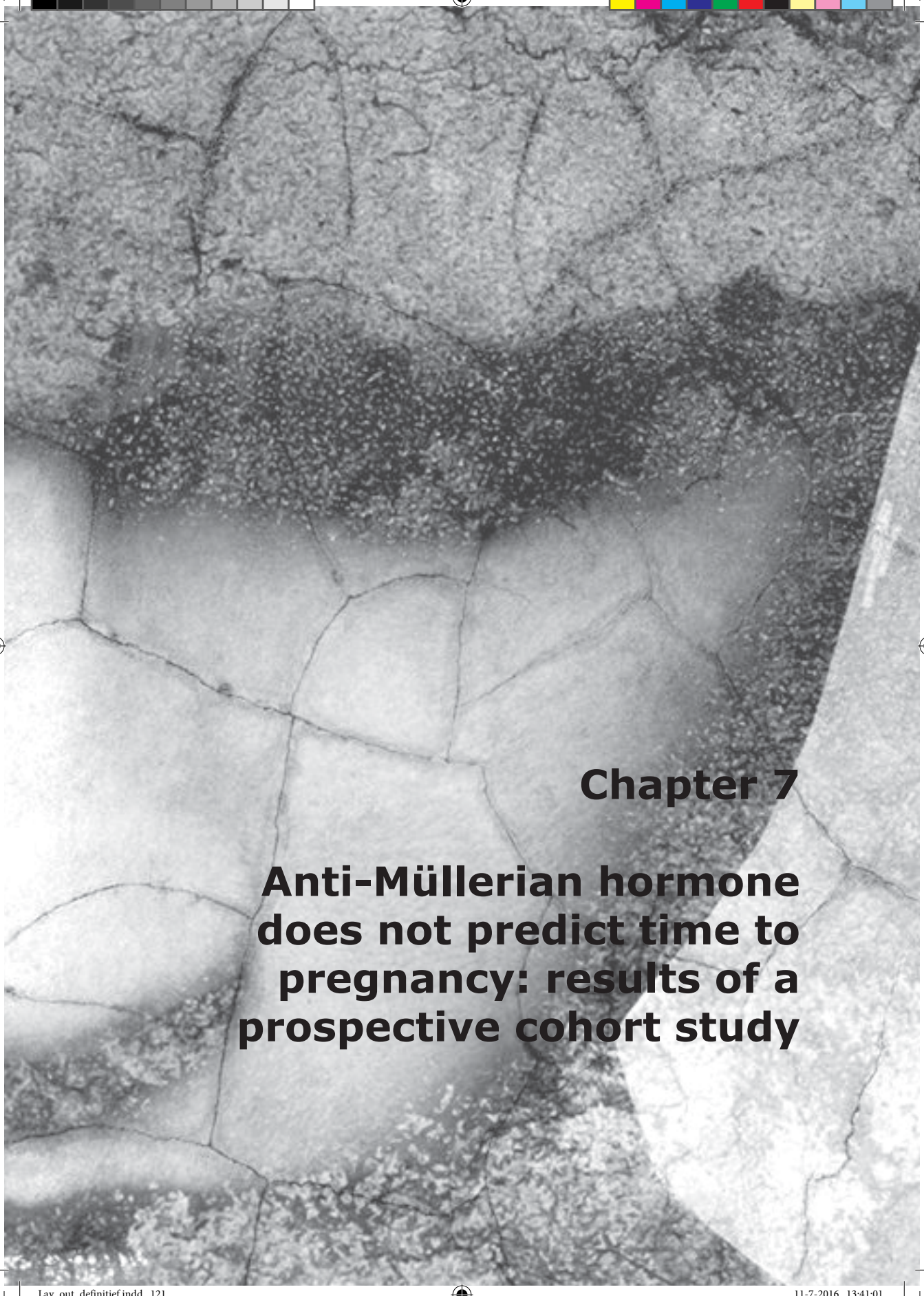
Search terms:
(anti mullerian hormone) OR (Anti-mullerian hormone) OR (AMH) OR (Mullerian inhibiting)
AND
(Menopau*) OR (Climacter*)

Inclusion criteria:
Availability baseline measurement of AMH (or retrospective use of stored blood samples)
Follow up data reporting age at menopause



M. Depmann, S. L. Broer, M.J.C. Eijkemans, I.A.J. van Rooij, G.J. Scheffer, J. Heimensem, B.W. Mol, F.J.M. Broekmans

Manuscript submitted



Chapter 7

Anti-Müllerian hormone does not predict time to pregnancy: results of a prospective cohort study

Abstract

Study question: Can ovarian reserve tests (ORTs) predict to time ongoing pregnancy in a cohort of healthy pregnancy planners?

Summary answer: A baseline measurement of Anti-Müllerian Hormone (AMH), Follicle Stimulating Hormone (FSH), or the antral follicle count (AFC) did not predict time to ongoing pregnancy.

What is known already: Studies are emerging regarding ORTs as time to pregnancy predictors. Of these ORTs, AMH is the most promising one. This has prompted AMH to be used in clinical practice or home fertility tests in order to assess fertility, despite lack of solid evidence.

Study design, size, duration: A prospective cohort study was conducted in an Academic hospital. 102 pregnancy planners were followed until ongoing pregnancy or until 12 months after cessation of contraceptives. Five couples were lost to follow-up before the study period of 12 months was finished.

Participants, setting, methods: 102 pregnancy planners provided a baseline measurement of AMH, FSH and the AFC. At the end of follow up, a semen analysis was performed and chlamydia antibody titers were assessed. Outcome measure was time to ongoing pregnancy. Predictive capacity of baseline ORT levels was assessed with Cox regression analysis, corrected for female age. The C-statistic reflected the capacity of the model to discriminate between women with a short or a long time to pregnancy.

Main results and the role of chance: 102 couples were included, of whom 97 completed the protocol. Of these 97 couples, 53 achieved an ongoing pregnancy during follow up time. In the univariate prediction model, both age and the AFC were significantly capable of predicting time to pregnancy (Hazard Ratio 0.92, 95% CI 0.87-0.98, p=0.01; 1.04, 95% CI 1.01-1.07, p=0.02 respectively). In the multivariate model however, we found AMH, nor basal FSH or the AFC to be predictive of time to ongoing pregnancy (Hazard Ratios 1.43, 95% CI 0.84-2.46, p=0.36; 0.96, 95% CI 0.86-1.06, p=0.43; 1.03, 95% CI 1.00-1.07, p=0.08 respectively). This was confirmed by the low C-statistic value.

Limitations, reasons for caution: Despite the fact that analyses were performed on a large cohort of pregnancy planners, some lack of power to detect an effect might have been present.

Wider implications of the findings: In a cohort of healthy pregnancy planners, baseline AMH, AFC or FSH levels did not have any relation to the time to ongoing pregnancy. These results limit the usability of these ORTs in assessment of current fertility.

Study funding: No external funds were used for this study

Introduction

Recent publications have resulted in a gain of territory for ovarian reserve testing in the field of fertility workup [121-124]. Moreover, an abundance of publications is available assessing the relation between markers reflective of the quantitative ovarian reserve, and live birth after assisted reproduction techniques [125,126]. Furthermore, ovarian reserve markers, such as anti-Müllerian hormone (AMH), follicle stimulating hormone (FSH) and the antral follicle count (AFC) are used to predict age at menopause [16,18,19,21,23,24,42,115], and to aid in the diagnosis of polycystic ovarian syndrome (PCOS) [63,127].

Since a trend amongst women towards delayed childbearing has resulted in a growing amount of women facing reduced fertility at the time they hope to conceive, the prediction of natural fecundability may be a promising field. If women facing reduced fecundability in the general population could be identified, it might be possible to refer this group to a fertility clinic sooner, or to advise them against delaying family planning.

However, the capacity of these ovarian reserve tests (ORTs) in the prediction of natural fecundability in the general population has been researched to a limited extent and no uniform conclusion can be drawn from the available research. While two studies demonstrated ORTs capable of predicting time to pregnancy, their conclusions were contradicted by three other studies demonstrating no predictive effect [121-124,128]. Moreover, these studies were either performed retrospectively on postpartum women [124], prospectively on women with a modest follow up time not exceeding six months [122,124], in women with a relatively high age [121], or in women experiencing repeated miscarriages [128]. The recent publications regarding ORT based time to pregnancy publications [121-124,128] followed by the use of home fertility tests prompted us to look in our existing cohort of pregnancy planners. In the present study we therefore evaluate the capacity of ovarian reserve tests (i.e. AMH, FSH and AFC) in the prediction of natural fecundability in a population based cohort of healthy couples.

Material and methods

Participants

Between 1999 and 2001 a prospective cohort study was performed in the

University Medical Center Utrecht in the Netherlands. Volunteers were recruited via advertisement in newspapers and magazines. Inclusion criteria were female age ranging between 18-46 years, the presence of 2 ovaries, no adnexal surgery in the past and the presence of a regular menstrual cycle (average interval between menses 21-35 days). Institutional Review Board approval was granted and informed consent was obtained from each of the participants.

Data collection

Participating women visited the clinic on day 2, 3 or 4 of their menstrual cycle (T1). A transvaginal ultrasound (Voluson-530D Kretztechnik, Zipf, Austria) was performed to obtain 3D scans of the ovaries for assessment of the number of follicles measuring 2-10mm. Furthermore, blood samples were obtained and plasma and serum were separated and stored at –20°C for later assessment of AMH and FSH. Couples returned when pregnant for a first trimester ultrasound or after 12 cycles (calculated from the time of cessation of contraceptives) if no pregnancy had occurred (T2). In both cases a semen analysis and Chlamydia Trachomatis antibody testing was then performed.

Hormone assay

Concentrations of FSH were measured with the use of the MEIA technology on a fully automated AxSYM immunoanalyzer (Abbott Laboratories, Abbott Park, IL) according to the manufacturer's instruction. The World Health Organization Second International Reference Preparation for Human FSH (78/549) was used as a standard in the FSH assay. For FSH, interassay coefficients of variation were found to be 5.7, 5.7 and 7.8% at the levels of 5, 26 and 79 IU/L, respectively (n=80). The detection limit for the FSH assay was 0.03 IU/liter.

AMH levels were measured with an ultrasensitive immune-enzymometric assay (Immunotech-Coulter, Marseille, France). The limit of detection (defined as blank +3sd of blank) was 0.05ng/ml. Values below the detection limit were artificially set at half the detection limits level (i.e. 0.025 ng/ml). Intra- and inter-assay coefficients of variation were less than 5% and less than 8%, respectively.

Outcomes

The end point was time to ongoing pregnancy. This was defined as a viable pregnancy of at least 11 weeks of gestational age. Time to ongoing pregnancy is defined in days since cessation of contraceptives.

Analyses

Baseline characteristics were depicted for all women enrolled in this study. Continuous variables were presented as means with the standard deviation (mean ± SD), categorical variables are presented as percentages.

Baseline characteristics of women lost to follow up were compared to women included in the final analysis. Furthermore, participants with an ongoing pregnancy were compared to participants who did not conceive during the 12 months follow up. A T-test or Chi- square test was performed to compare means between groups. Results were considered to be statistical significant at a p-value <0.05.

For cases with missing data, either single or multiple imputation was performed. Cases with a missing outcome measure were discarded from analysis.

A scatterplot of AMH versus age was constructed to differentiate between women that did or did not achieve an ongoing pregnancy during the follow up period. A logistic regression analysis was performed to assess if baseline ORT could differentiate between women who would conceive an ongoing pregnancy during follow up time and those who would not. Next, a multivariable logistic regression analysis was performed incorporating age, smoking status or BMI in a model next to baseline ORT.

Univariate and multivariate Cox regression analyses for time to pregnancy (or until 12 months after cessation of contraceptives for women that did not achieve an on-going pregnancy) as time axis and ongoing pregnancy as event were performed to assess the capacity of the variables of interest in predicting time to pregnancy. Time to pregnancy was censored at the last moment it was known that pregnancy had not occurred. Note that in order to correct for variation occurring in time between cessation of contraception and inclusion, left truncation was applied. Moreover, in order to better fit the nonlinear distribution of AMH data, a restricted cubic spline was used.

The C-statistics were calculated to measure the capacity of the model to discriminate between women with a short or a long time to pregnancy. In order words, the C-statistic reflects how accurate the model can distinguish between two randomly chosen women who have a different time to pregnancy. A model is considered reasonable when the C-statistic is greater than 0.7 and strong when this value is greater than 0.8.

Analyses were performed using SPSS 21 (SPSS Inc., Chicago, IL) and R version 3.0.3 (http://www.r-project.org/).

Results

We enrolled 102 women in this study. Baseline characteristics are shown in Table 7.1.

	Baseline parameters	Completed protocol		
	N=102	Ongoing pregnancy N=53	No ongoing pregnancy N=44	P-value
Female age (years; ± SD)	32.9 ± 4.4	31.9 ± 3.6	34.4 ± 5.0	**0.01**
Male age (years; ± SD)	35.5 ± 6.1	33.4 ± 5.1	36.1 ± 6.9	**0.04**
BMI women (kg/m²; ± SD)	24.3 ± 3.8	24.1 ± 3.6	24.7 ± 4.3	0.47
Previously pregnant –any outcome- (%)	26.5	22.6	30.6	0.36
Current smoking				
Women (%)	7.8	3.8	12.2	0.11
Men (%)	22.2	17.3	27.7	0.22
Frequency of intercourse (/week in %)				0.10
1	31.4	35.8	26.5	
2	35.3	41.5	28.6	
3	22.5	13.2	32.7	
>4	10.8	9.4	12.2	
ORT				
FSH (IU/L; ± SD)	6.2 ± 2.6	6.1 ± 2.4	6.6 ± 2.8	0.37
AMH (ng/L; ± SD)	2.1 ± 2.1	2.4 ± 2.3	1.7 ± 1.6	0.08
AFC (n; ± SD)	15.1 ± 8.2	16.1 ± 7.7	13.4 ± 8.6	0.10
Semen parameters[a]				
VCM (± SD)		123.0 ± 120.8	118.4 ± 167.9	0.90
Volume (ml)		3.5 ± 1.8	3.6 ± 1.7	0.70
Concentration (10^6)		62.6 ± 49.3	51.4 ± 82.5	0.49
Motility (type A + B)		53.6 ± 12.3	46.8 ± 46.1	0.16
Chlamydia antibody (% positive)	22.3	7	13	0.35

Table 7.1: Baseline characteristics and comparison of couples with and without an ongoing pregnancy. Values are mean ± standard deviation or n (%); BMI= body mass index (kg/cm²), [a]= the number of semen analyses in couples with and without an ongoing pregnancy is 41 and 25 respectively.

Mean age at baseline was 32.9 ± 4.4 years. Women usually had a normal body mass index (BMI, 24.3 ± 3.8 kg/m^2) and were primarily non-smokers (8% current smokers). One in four participants (26.5%) had had a previous pregnancy. Mean values of FSH and AMH were 6.2 IU/L (range 2.2 -16.9) and 2.1 ng/L (range 0.03-10.5) respectively, the mean AFC (2-10mm) measured 15.1 (range 3-48 follicles). Five couples were lost to follow-up before the study period of 12 months was finished. One couple was divorced after three months, one couple started fertility treatment elsewhere during the study period and three couples left the study without further notice.

Couples lost to follow up were compared to couples included in the final analysis. There was a significant difference noted for male smoking at baseline (60% for couples lost to follow up and 19.1% for the remaining study cohort, p=0.001). Furthermore, couples lost to follow up had a higher frequency of intercourse (p=0.003). In Figure 7.1, a flowchart of inclusions is depicted.

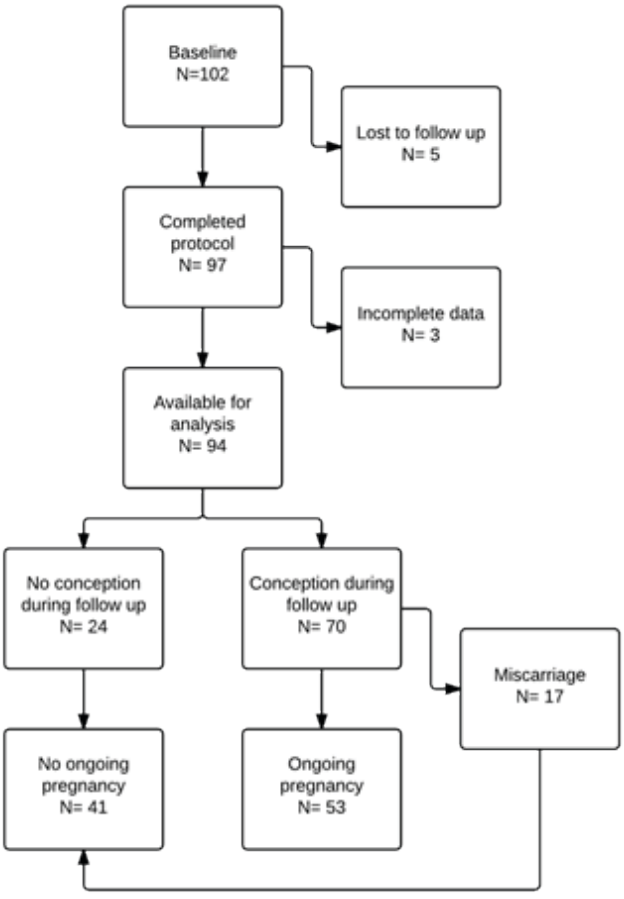


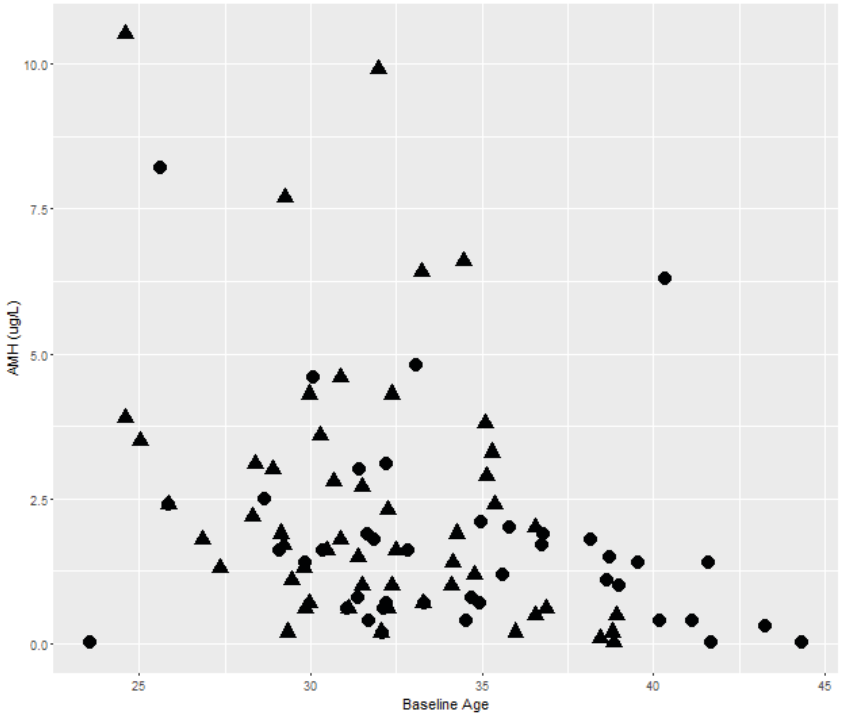
Figure 7.1: Flowchart of inclusions

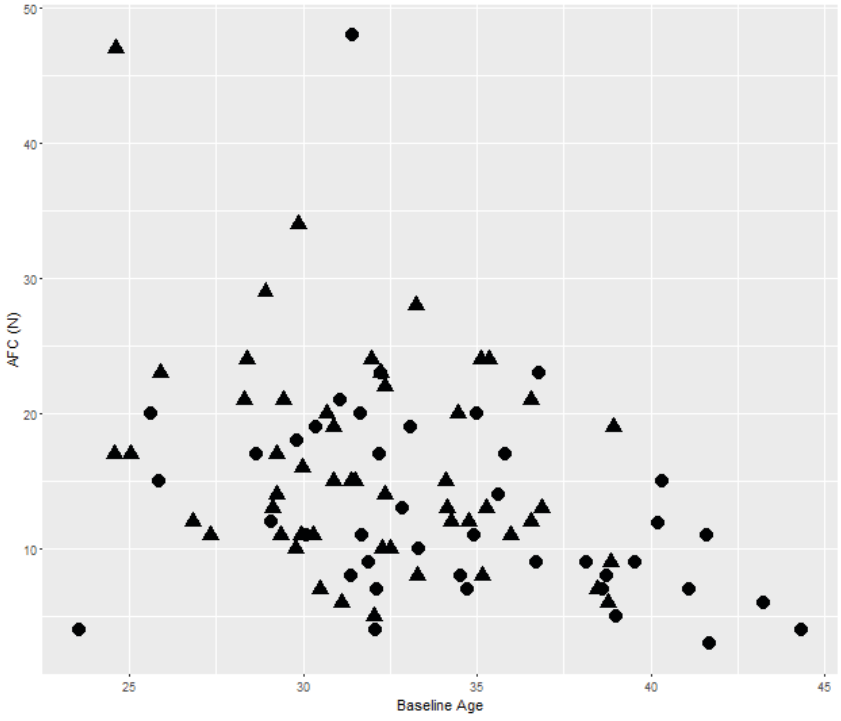
Baseline characteristics (Table 7.1) of couples with an ongoing pregnancy (n=53) were compared to couples that failed to achieve an ongoing pregnancy within the one year period (n=44). Significant differences were present in both female and male age. Couples that experienced an ongoing pregnancy during follow up were younger than those who did not conceive (mean female age 31.9 years vs 34.4 years, p=0.01 and mean male age 33.4 years vs 36.1 years, respectively, p=0.04). There were no significant differences between the presence of a positive chlamydia antibody test or in any of the semen parameter levels.

In Figure 7.2, scatterplots of ORTs versus age are depicted, differentiating between women that did or did not achieve an ongoing pregnancy during follow up.

Logistic regression analysis demonstrated no predictive effect of baseline ORT in the prediction of natural conception or an ongoing pregnancy during the follow up period. The adding of age, smoking status or BMI to a model containing baseline ORT measurements did not alter these results.

Prior to Cox regression modelling, single imputation was performed on a single case with a missing antral follicle count. Moreover, multiple imputation was performed on cases with missing semen analysis. Four cases with an unknown date at cessation of contraceptives, hence an unknown time to event, were discarded from analysis. Therefore, 94 women were available for Cox regression analysis (Table 7.2).





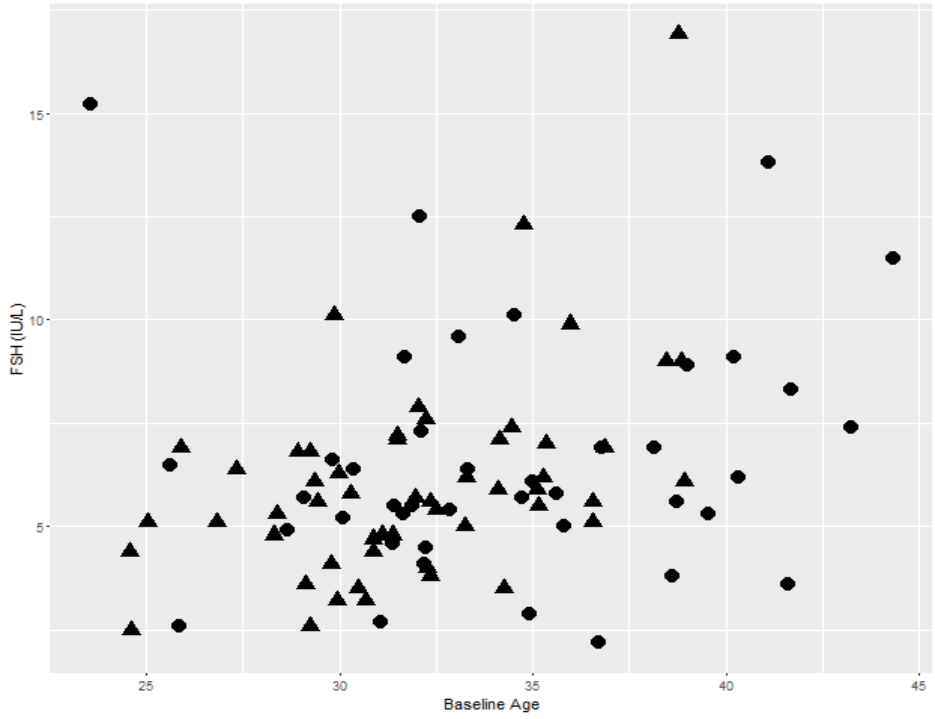


Figure 7.2: scatterplots depicting baseline age, ORTs and ongoing pregnancy
The x-axis displays baseline age, the y-axis baseline ORT level (upper panel AMH, middle panel AFC and lower panel FSH), each dot/triangle represents a women (dot=no ongoing pregnancy, triangle= ongoing pregnancy)

In the univariate analysis both the AFC and female age were significantly capable of predicting time to ongoing pregnancy (p=0.02 and p=0.01 respectively). However the C-statistic for both variables was poor (0.54 and 0.56 respectively).
AMH was not significantly capable of predicting time to ongoing pregnancy (HR 1.66, 95%, CI 0.97-2.85, p value 0.18, C-statistic 0.55). In the multivariate Cox regression analysis, where a correction for female age was performed, none of the variables analyzed was significantly correlated with time to pregnancy, nor did they reach a predictive accuracy level of any importance.

	Hazard Ratio	95% CI	P-value	C- Statistic
Univariate regression				
Female age	**0.92**	**0.87-0.98**	**0.01**	**0.56**
FSH	0.94	0.84-1.05	0.30	0.53
AMH[a]	1.66	0.97-2.85	0.18	0.55
AFC	**1.04**	**1.01-1.07**	**0.02**	**0.54**
CAT+	1.59	0.83-3.03	0.16	0.58
VCM[b]	1.00	1.00-1.00	0.91	0.58
Volume[b]	0.94	0.77-1.15	0.58	0.55
Concentration[b]	1.00	1.00-1.01	0.61	0.59
Motility[b]	1.01	0.99-1.03	0.52	0.53
Morfology[b]	1.02	0.97-1.07	0.51	0.55
IgGMAR[b]	1.01	0.96-1.07	0.58	0.52
Multivariate regression –corrected for age at baseline-				
FSH	0.96	0.86-1.06	0.42	0.57
MH[a]	1.43	0.84-2.46	0.36	0.58
AFC	1.03	1.00-1.07	0.08	0.56
CAT	1.97	1.02-3.84	0.05	0.62
VCM[b]	1.00	1.00-1.00	0.91	0.56
Volume[b]	0.95	0.78-1.15	0.60	0.58
Concentration[b]	1.00	1.00-1.00	0.63	0.57
Motility[b]	1.01	0.99-1.03	0.32	0.57
Morfology[b]	1.02	0.97-1.07	0.42	0.58
IgGMAR[b]	1.01	0.96-1.06	0.68	0.57

Table 7.2: Results of univariate and multivariate Cox regression analysis for time to ongoing pregnancy
FSH= follicle stimulating hormone, AMH[a]= anti-Müllerian hormone, added to the model using a restricted cubic spline, AFC= antral follicle count, VCM= volume x concentration x motility, [b]=imputed data (number of imputed cases in VCM N=31, volume N=31, concentration N=31, total motility N=31, IgGmar N=39, morfology N=34), CAT= chlamydia antibody positive in serum

Discussion

In this study we demonstrated that neither anti-Müllerian hormone nor the AFC nor basal FSH is capable of correctly predicting time to ongoing pregnancy within a 1 year period after cessation of contraceptives in a general population based prospective cohort. Female age was found to be the only significant predictor of time to pregnancy. This indicates that there is no solid basis for the use of ORTs in the prediction of time to pregnancy in both clinical practice and in self-testing. The present results regarding AMH based time to pregnancy prediction are in some contradiction to results published by Steiner et al.[122]. In that study, day-specific probability of pregnancy was calculated using ovarian reserve tests measured in 78 women. Women with lower peripheral AMH levels had a significantly decreased chance of conceiving when compared to women with higher AMH levels in a model next to female age. For these calculations, an arbitrarily set AMH cut-off value of 0.7ng/mL was chosen and age was dichotomized at age 35. Moreover, follow up time did not exceed 6 months. Since it is common use to express chances of conceiving within a one-year window, this short follow up limits the applicability of the findings. Lastly, calculations of results using dichotomized values for female age and hormone levels may lead to loss of information compared to data analyzed using the continuous variables.

Much in the same line, Hagen et al [123] performed a time to pregnancy analysis on 186 first pregnancy planners. Follow up time was again a modest 6 months. Women were divided into quintiles according to their AMH levels. Next, pregnancy rates in the lowest and the highest AMH quintiles were compared to pregnancy rates in the three middle quintiles. After adjusting for female age, BMI, reproductive organ disease, smoking and oligozoospermia, the high AMH group showed a reduced probability of conceiving (fecundability rate 0.62, 95% CI 0.39-0.99). Further adjustment for irregular cycles in this group did not alter this effect (fecundability rate 0.55, 95% CI 0.30-0.98). There was no significant difference in time to pregnancy in the low AMH quintile compared to the normal AMH levels. Interestingly, applying the cut-off value for AMH set in the study by Steiner et al [122], did not demonstrate AMH to predict time to pregnancy. The short follow up of 6 months, and the use of AMH categories instead of continuous data may have made the analyses less reliable.

Streuli et al. [124] performed analyses on 87 postpartum women, where AMH was measured in frozen serum, samples were obtained during the first trimester of pregnancy. No relation could be demonstrated between AMH and effective time

to pregnancy. However, one could postulate that including postpartum women might have resulted in recall bias regarding time to pregnancy and inclusion bias due to the exclusion of infertile women making these results less solid [129]. Zarek et al. [128] performed AMH based time to pregnancy analysis on 1202 women with a history of one or more miscarriages. In order to correct for a U-shaped, non- linear effect of AMH, 3 AMH categories were constructed (low, normal, high). With the normal AMH category as reference value, low or high AMH levels were not associated with altered fecundability rates. The major limitation of this study, however, is the lack of generalizability of the findings to the general population since only women with a history of pregnancy loss were included.

FSH based pregnancy prediction was reviewed by van Montfrans et al. [121]. In that study 129 women aged ≥ 30 years were followed during a 12 month period, or until pregnancy occurred. Mean baseline FSH, measured on cycle day 2-4 in the first three months after inclusion, was used to predict the occurrence of pregnancy. In line with the present findings, no significant relationship between basal FSH and the occurrence of pregnancy could be demonstrated. The previously discussed study by Steiner et al. [122] also reviewed the predictive capacity of FSH and the AFC and noted no predictive effect of these ORTs in a time to pregnancy prediction.

In the present study, no predictive effect for any of the semen parameters measured regarding time to pregnancy prediction could be observed. This is in some contradiction to observations made by Bonde et al. [130]. In this study the probability of conception increased with increasing sperm concentration and with the proportion of normal morphology amongst 430 first pregnancy planners. One can only speculate as to why this was not observed in our cohort. It most likely stems from the fact that in our cohort, a semen analysis was present in only 31 of the 97 couples. Consequently a lack of power to detect a statistical significant difference might be present. This is especially the case since missing data did not occur at random, 75.5% of couples that achieved a pregnancy provided a semen analysis compared to 53% of couples that failed to conceive.

The strength of the present study lies in the fact that this is a prospective cohort study with a follow up time of one year. As previously stated it is common use to express chances of conceiving within a one-year window, therefore the results can be easily translated to the day-to-day clinical practice.

A potential weakness of our data is the fact that inclusion of women took place between 1999 and 2001. Since preliminary analysis at that time did not show

any relation between AMH and time to pregnancy, the data was not previously published. However, with an increasing amount of literature being published suggesting a relation between AMH and current fertility, it was felt appropriate to share our results. Furthermore, an in-house conversion factor is available translating the AMH assay applied to the common used Gen II assay, making our data usable in the current setting [110,131].

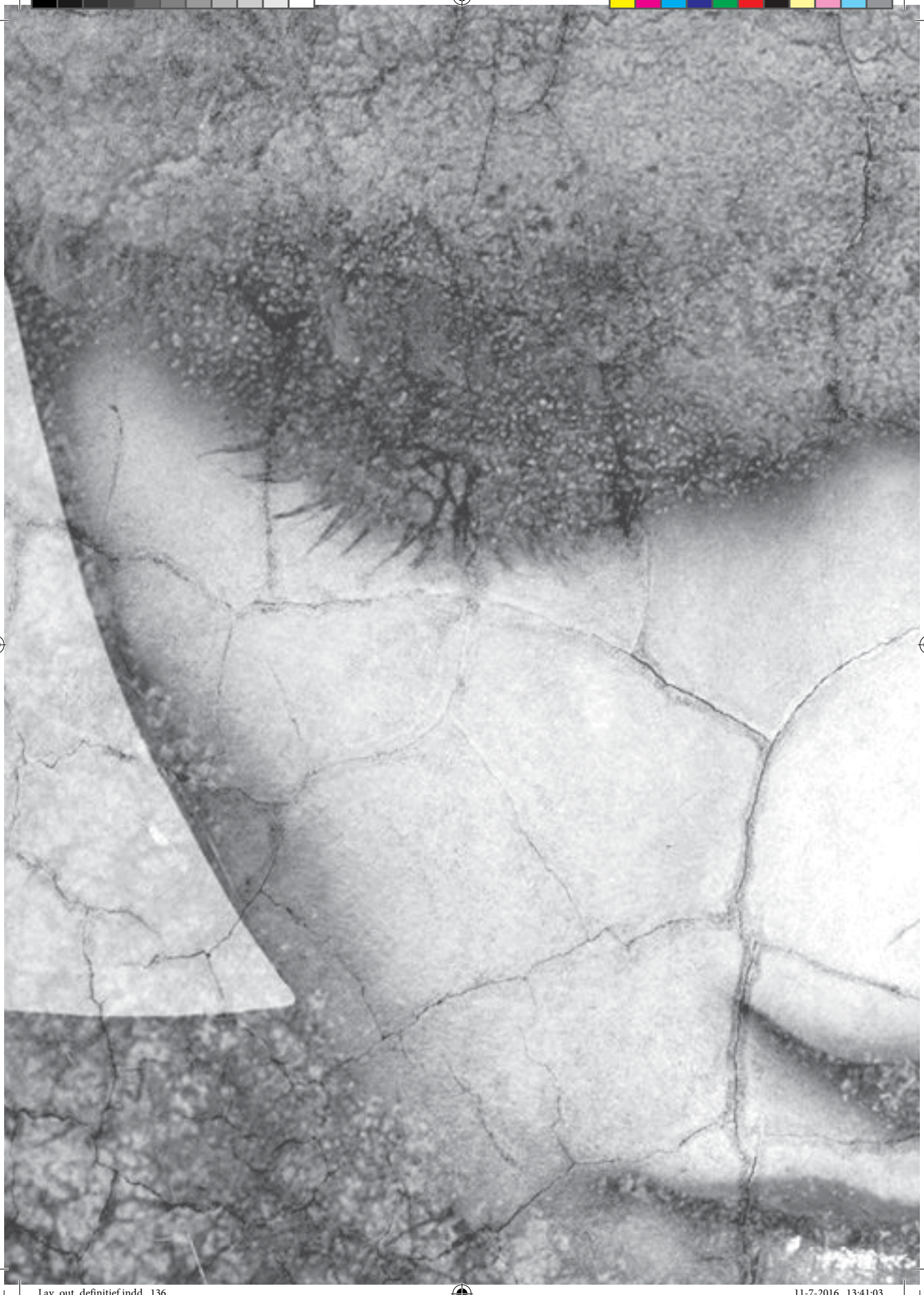
Secondly, despite the fact that, compared to other cohorts available, calculations were made in a large cohort, study power might have had some influence on our results. When looking at the hazard ratio (HR 1.43) of AMH based time to pregnancy prediction corrected for female age and the corresponding confidence interval (0.84-2.46), some AMH predictive effect might be present. The HR of 1.43 reflects a moderate effect, but the current study was not able to demonstrate this effect to be significant, which potentially results from the number of observed events (ongoing pregnancy) in our cohort.

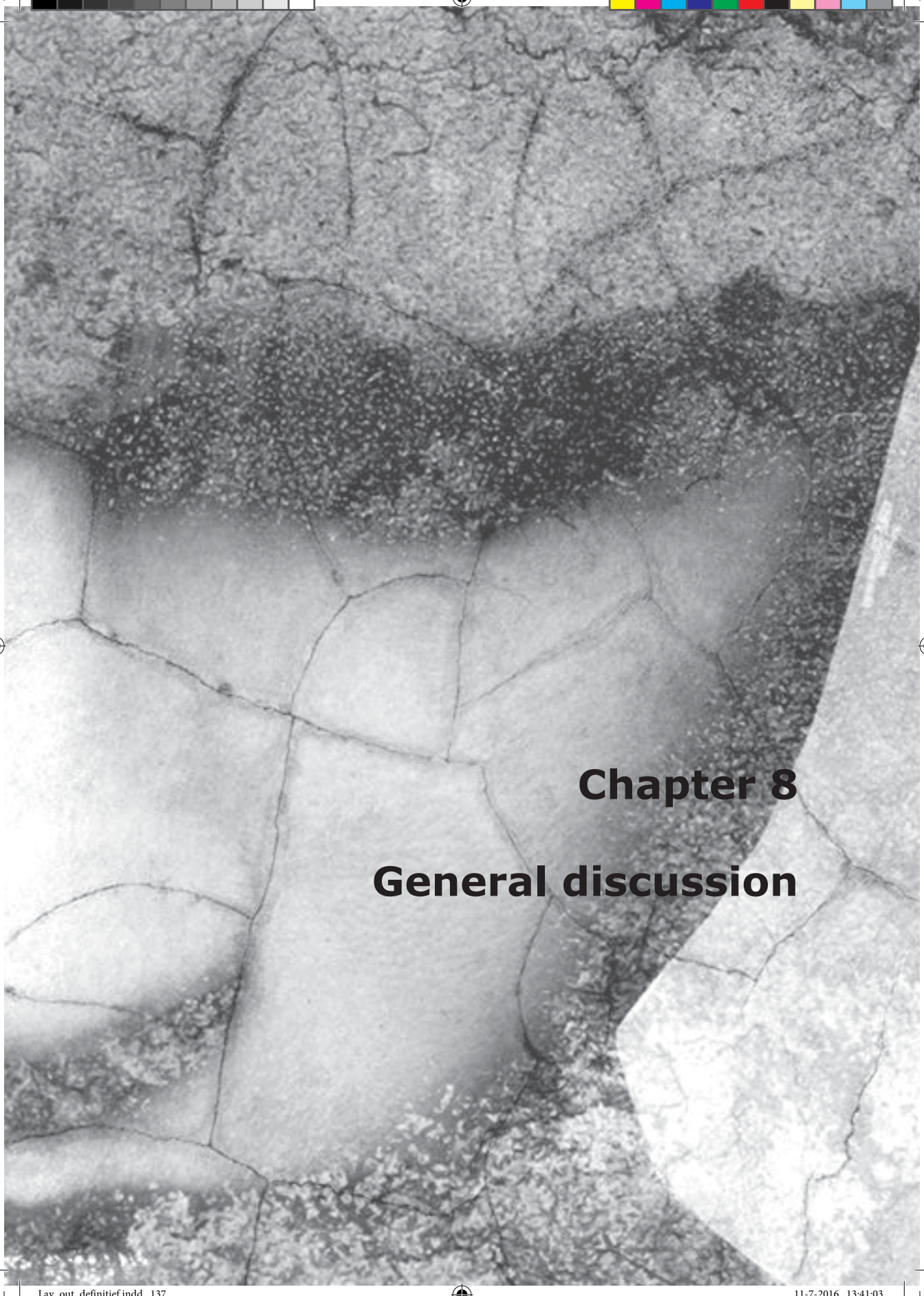
Another potential weakness of our study is the fact that women that had been trying to conceive for some months were included in our study. This might result in a less fertile cohort, where women conceiving immediately after cessation of contraceptives might be underrepresented. This could be reflected by our overall pregnancy rate (ongoing and spontaneous abortion) of 74.5%, which is lower than the 91% pregnancy rate observed in the general population [132,133]. We do however feel that, by using left truncation, we have adequately corrected for the difference in time between cessation of contraceptives and baseline that occurred between women. Another explanation for the our relatively low pregnancy rate may be the mean age in this cohort, which, at 32.9 years, is relatively high.

In contrary to recent publications [122-124], our results clearly state that AMH does not predict time to pregnancy in the general population. We therefore conclude that there is no place for anti-Müllerian hormone measurement in the assessment of time to pregnancy in the general population either in the clinical practice or using home fertility kits.

Conclusion

In this prospective follow up study we demonstrated that baseline ovarian reserve tests (i.e. AMH, FSH and AFC) failed to show any relation to the time to ongoing pregnancy in a cohort of healthy pregnancy planners. These results imply that AMH does not have a clear place in current fertility assessment in the normal population.





Chapter 8

General discussion

Introduction

The end of natural fertility occurs within a broad window of 20 years and, in the absence of a marker, is an event that goes by mainly unnoticed. Menopause is a more clearly marked event, and occurs at a fixed ten years interval after the end of natural fertility. Considering the broad timeframe for the occurrence of both reproductive events, one can postulate that age alone provides only limited knowledge to inform women on their individual reproductive lifespan. There is no marker available reflective of the end of natural fertility or even of the actual individual monthly fecundity rate. Therefore and based on the putative fixed ten year interval between the end of natural fertility and the final menstrual period, age at menopause may be used as a proxy variable in the prediction of the end of natural fertility.

Menopause is believed to occur when the number of primordial follicles in the human ovary falls below a critical threshold [25,39-41]. Moreover, a gradual decline in the number of follicles is hypothesized to be present from conception to menopause [41]. Therefore, markers reflective of the number of primordial follicles present at a given time could be used in order to predict menopause, and thus the end of natural fertility. Aside from a gradual decline in the number of ovarian follicles, a gradual decline in the quality of the oocytes held within the ovarian follicles is also present. This decline in oocyte quality is responsible for the decline in conception rates with increasing age and for the increase in spontaneous pregnancy loss at higher ages [134]. Ovarian reserve tests have therefore gained territory in the field of fertility testing. After all, if the decline in the number of primordial follicles and the decline in oocyte quality go hand in hand, ORTs reflective of the number of primordial follicles could potentially reflect the oocyte quality.

In this thesis we have assessed the origin of fluctuation in AMH levels and investigated the role of the true ovarian reserve and ovarian reserve tests as to their capacity to predict menopause. Furthermore, we have assessed the capacity of ovarian reserve tests in the prediction of spontaneous ongoing pregnancy.

The aims of this thesis were listed as follows:

1. To investigate some basic concepts of ovarian reserve testing.
2. To determine the role for ovarian reserve tests in the prediction of the reproductive lifespan.
3. To examine the role for ovarian reserve tests in the assessment of current fertility.

Using ovarian reserve tests to predict age at natural menopause

The first two aims, covered in chapters 2-6, focused on some basic concepts of ovarian reserve testing and the potential of ovarian reserve tests in the prediction of age at natural menopause. The concept of using ovarian reserve tests for menopause prediction stems from the theory that ORTs are reflective of the true ovarian reserve, defined as the non- growing follicle pool. Taking into account that it was demonstrated that the number of ovarian antral follicles proportionally reflects the number of primordial follicles [29] and ovarian reserve tests represent the number of larger (ie. ultrasonically visible) ovarian antral follicles, it is understandable that ORTs reflect the true ovarian reserve. Before one can use ORTs as derivatives of the true ovarian reserve in order to predict age at natural menopause, one must investigate if the true ovarian reserve itself indeed relates to the timing of natural menopause.
The relation between age at natural menopause and the true ovarian reserve was therefore researched in chapter 2 of this thesis. In this chapter it became clear that the distributions of age at menopause calculated according to the decline rate in primordial follicle number paralleled those observed in population based data. Therefore it was stated that the endowed follicle pool at birth may be an important determining factor of the timing of natural menopause. This finding therefore supports the use of ORTs in in clinical conditions, such as menopause prediction.
It is however important to realize that research investigating menopause prediction using either the true ovarian reserve or derivatives of this true ovarian reserve is based on a major assumption. This major assumption is that the decline pattern in the number of primordial follicles (or any other ovarian reserve marker) from conception to menopause is fixed and similar for all women. Although theoretical research [41] demonstrated models allowing for different decline rates between women to have a lesser fit to the available data than models with a constant and uniform decline, and studies regarding

ORT based menopause prediction, either prospective or cross sectional, found promising results based upon this assumption [14,15,17-19,22,42,117-119,131], prospective research on the true ovarian reserve decline is not available, for obvious reasons. If, for instance, women with a low number of primordial follicles at birth have a relatively low decline rate and women with higher numbers of endowed non- growing follicles a higher decline rate, their age at menopause may differ less than the endowed follicle pool alone (and their corresponding ORT levels) would suggest (Figure 8.1).

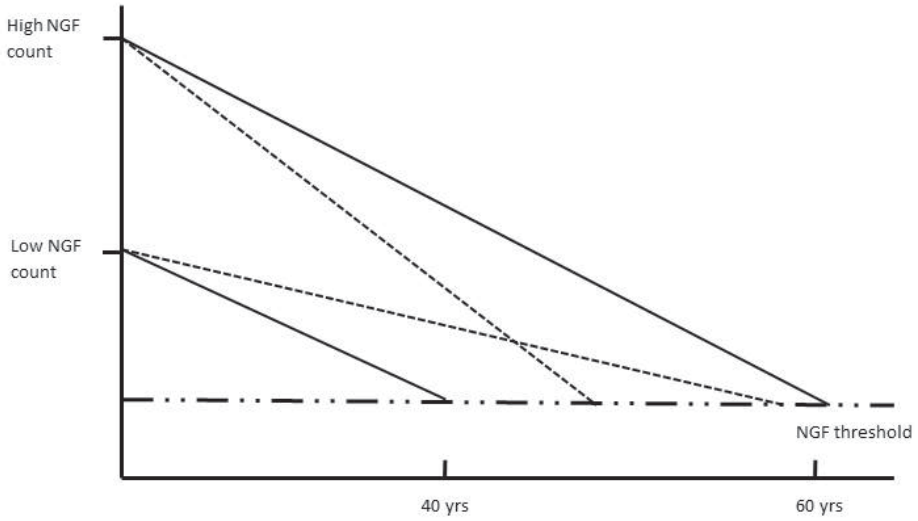


Figure 8.1: a schematic display of the result of different decline rates on age a menopause.
The solid lines represent the effect of non-variable decline rates between women. A woman with a high number of non- growing follicles (NGF) present at birth will reach menopause late in life, while a woman with a low NGF count a birth will experience the final menstrual period early in life.
The dashed lines represent the effect of a variable decline rate between women. Note that a low NGF count at birth combined with a low rate of decline results in a normal-high age at menopause, whereas a high NGF count at birth combined with a high decline rate results in an early-normal age at menopause. The dashed- dotted line represents the NGF threshold where menopause is predicted to occur [41,103].

The possibility of a different decline rate in the number of primordial follicles could be tested by repeated measurements analysis of ovarian reserve tests over a longer period of time. No such research is available at the present time. At present, therefore, all studies addressing the relation between ovarian reserve status and later occurrence of menopause implicitly assume that for any individual, decline rates in primordial follicle number follow the same

pattern.

After investigating some basic concept of menopause prediction using the true ovarian reserve, our interest shifted towards the stability of the most important ovarian reserve tests [38] throughout the menstrual cycle. Since the stability of AMH and the AFC throughout and between menstrual cycles has been subject to debate [64-69], we sought out to research the origin of potential fluctuation of these ORTs in chapter 3. After all, if ovarian reserve tests are to be used as derivatives of the true ovarian reserve, it is important to assess if their levels are constant over short time periods. Moreover, and as explained in chapter 3, if fluctuation is to be present in ORT levels and all ovarian reserve tests are an equal derivative of the true ovarian reserve, any level of fluctuation by one ORT should be equally present in the level of a different ORT. In this chapter we demonstrated that fluctuations in peripheral levels of AMH parallel fluctuations in the number of antral follicles (2-8mm). This information added to the basic understanding of origin of AMH and stability and parallelism of ORTs throughout the menstrual cycle. The conclusions are also in line with observations made in other studies. For instance in the study by Jeppesen et al. [34] a parallelism was observed between peripheral levels of AMH and numbers of follicles measured 2-8mm. The study by Fanchin et al [36] showing a correlation between small antral follicles and AMH levels further corroborates our observations. Based upon the research outlined in chapter 3, variations in peripheral levels of AMH throughout the cycle are more understandable and one can postulate that timing of ovarian reserve testing in the menstrual cycle is of importance. After all, if circulating levels of AMH reflect the antral follicle count and AMH levels drop when follicles grow beyond 8 mm in size, one can understand that the early follicular phase provides possibly the best window for ovarian reserve testing. Moreover, the reduction in peripheral levels of AMH observed in women using oral contraceptives [73] is understandable since a concurrent reduction in the antral follicle count is observed in these women, affecting output of AMH.

The current status of age at natural menopause (ANM) prediction using derivatives of this true ovarian reserve was investigated in chapter 4. In this extensive literature review ANM prediction based on a baseline measurement of AMH, the AFC or on the mother's ANM was investigated. It became clear that, apart from actual age at assessment, AMH and mother's ANM are the strongest tools available for the prediction of ANM. As promising as this might seem, some major pitfalls were revealed. When looking at ANM predictions it became clear that the models available uniformly lacked the capacity to predict the extreme ages at menopause (meaning either early or late age at menopause)

and provided wide prediction intervals [17-19,21,22,117]. Translating this into a clinical setting would imply that these models cannot correctly identify those women destined to experience early menopause and thus experiencing the end of natural fertility early in life, making them at risk for age- related subfertility. Moreover, when predicting age at menopause, the prediction interval is still wide making these predictions clinically less relevant.

It was hypothesized that predictions of the extreme menopausal ages were lacking due to a two-jointed problem. On the one hand, early and late menopausal ages are relatively rare and the cohorts available performing these analyses were relatively small [17-19,21,22,117]. It was deemed therefore possible that due to a simple fact of small numbers, these women were not represented in the available cohorts. Furthermore, all cohorts available researching menopause prediction included only women with regular cycles. This could have resulted in the situation that women destined for early menopause, and thus already experiencing irregular cycles at a young age, were excluded at baseline. Moreover, in many cohorts the available follow up time was not completed, therefore women experiencing menopause relatively late are also not available in the cohorts since they simply have not experienced menopause yet.

The second problem observed in AMH based menopause prediction is the existence of a large age interval surrounding the predicted age at menopause. We hypothesized that this broad interval may stem from imprecision originating from AMH itself or from the fact that not all factors contributing to individual age at menopause were incorporated in the models, such as genetic- and lifestyle factors. Regarding the imprecision resulting from AMH, it is well known that the large variety of assays used over time and assay stability itself has resulted in discussion on the use of AMH [28]. When keeping the results presented in chapter 3 in mind, it is understandable that some variation in AMH levels exists throughout the menstrual cycle. The imprecision resulting from this within cycle fluctuation could be subdued by performing AMH based menopause predictions using repeatedly measured levels of AMH. The imprecision arising from the AMH assay used is anticipated to be straightened out with the use of rigid assays which are available soon [95]. The broad age interval resulting from the fact that not all factors contributing to age at menopause are incorporated in models used is however more difficult to resolve. This problem can only be solved by identifying these currently unknown factors contributing to age at menopause. This is a search that is ongoing, factors of interest are of course, and as emphasized in this chapter, mother's age at natural menopause and also genetic factors and lifestyle factors such as body mass index.

The above mentioned pitfalls and the hypothesis of their origin were investigated in chapters 5 and 6.

In chapter 5 we have provided the results of an ongoing follow up study investigating ORT based age at natural menopause prediction. While it was expected that a prolonged follow up would affirm our previous conclusion that AMH predicts menopause [18] and would most likely expand ANM predictions to higher ages at menopause due to this prolonged follow up, this was not entirely the case. With an increase in follow up time, a decrease in the capacity of AMH in the prediction of age at menopause was observed. This unexpected observation most likely resulted from the fact that when performing analyses using age-specific AMH percentiles, women within each percentile category vary in their individual risk of reaching menopause. We hypothesized that women with the highest hazard of becoming menopausal would have experienced menopause at the first follow up round, whilst women with lower risks of menopause remained available for follow up. This observation clearly reduces the potential of AMH in menopause prediction. Interestingly the reduced potential of AMH based menopause prediction was also revealed in a recent study [24]. In line with our results, a non–proportional hazard for the association of AMH based menopause prediction and time was observed. Unfortunately, no possible explanation for the reducing hazard was provided by the author's.

All above mentioned pitfalls and our hypothesis of their origin are depicted in Table 8.1.

In spite of our expectation to resolve the ever present obstacles in menopause prediction described above (i.e. lack of prediction of extreme menopausal ages and a broad prediction interval) we only added the problem of a reduced AMH effect over time to the equation. We therefore felt it was time to combine all known prospective cohorts investigating AMH based menopause prediction available in an individual patient data meta- analysis. This was performed in chapter 6 of this thesis.

Problem	Hypothesis of origin	Possible solution	Addressed in
Lack of prediction of early age at menopause	-women experiencing early age at menopause excluded due to inclusion criterion "having a regular cycle" -relative rareness of event	-including young women -including women irrespective of cycle state -increase numbers	Chapter 4-6
Lack of prediction of late age at menopause	-incomplete follow up -relative rareness of event	-long follow up -increase numbers	Chapter 4-6
Wide prediction interval	-imprecision of AMH assay -variation by cycle sample moment -failure to include all factors contributing to age at menopause	-repeated AMH measurement -rigid AMH assay -incorporating other (unknown) factors	Chapter 4-6
Non-proportionality of AMH	-heterogeneity in individual risk of reaching menopause		Chapter 5, 6

Table 8.1: Current problems accompanying AMH based age at menopause predictions

In this chapter, five of the seven known cohorts that contain data on AMH based menopause prediction where combined to form a single large cohort. A Cox regression analysis was performed assessing AMH based time to menopause prediction. We demonstrated AMH to be a significant predictor of time to menopause, but revealed that the predictive capacity on top of female age is poor. Moreover, an AMH based age at menopause prediction was performed again using age-specific AMH percentile categories. It was revealed that despite a larger volume of data, the known pitfalls in the AMH based age at menopause predictions remained present. Regarding the lack of the prediction of the extreme ages at menopause, the IPD dataset used "having a regular cycle" as inclusion criterion in order to harmonize the data. As previously stated, this inclusion criterion most likely resulted in the exclusion of younger women experiencing an irregular cycle due to imminent ovarian depletion resulting in menopause. Consequently, our predictions did not contain early menopausal

ages. We were able however to expand age at natural menopause to the higher ages due to this larger dataset. This matches our hypothesis that prolonging follow up will result in the ability to predict the later menopausal ages.
Unfortunately, prediction interval remained wide and thus clinically not likely to become applicable. This was expected since we hypothesized that the wide prediction interval resulted from imprecisions originating from AMH assays and short term variation and from the lack of incorporating other factors contributing to age at menopause in the prediction model. Using five different datasets and henceforward an even higher number of different AMH assays naturally did not improve imprecision arising from AMH in spite of our attempt to bypass this problem by constructing age- specific AMH percentile categories within the study of origin. Furthermore, and as explained in this chapter, the largest datasets dictated the available variables to be incorporated in the prediction model and these where unfortunately lacking additional prediction factors.
We have previously explained that chapter 5 revealed a remarkable finding regarding a reducing predictive capacity of AMH with increasing age due to non- proportionality of the effect. Interestingly, this non-proportionality was present to an even larger extend in the individual patient data meta-analysis. When looking at the predictive capacity of AMH in younger women, a quite strong AMH effect was present, but this effect was virtually lost in older women. As explained in chapter 6, is not surprising that for women aged older than 50 years the occurrence of menopause in the next ten years is an inevitable event and their AMH levels will not contribute to an age at menopause prediction. Based on the current literature and on chapter six, and in spite of a significant predictive effect of AMH in the prediction of age at menopause, AMH based age at natural menopause predictions have no place in clinical practice due to the unacceptable inaccuracy surrounding the predictions.

Using ovarian reserve tests to predict time to pregnancy

The third and final aim of this thesis was addressed in chapter 7. In this chapter the aforementioned theory of a simultaneous decline in oocyte quantity and quality was tested by using ovarian reserve tests to predict time to pregnancy. In contrast to some evidence available [122,123] but in line with other studies [121,124,128] we demonstrated that none of the ORTs researched (AMH, the AFC, FSH) could predict time to ongoing pregnancy. Our analyses were corrected for female age and semen parameters, chlamydia antibody titers amongst other characteristics were taken into account. We established that female age

alone is the dominant predictor for time to ongoing pregnancy, and that the addition of ovarian reserve tests provides no added information for the general population.

This information is important when counseling women regarding their current fertility. In spite of the ever increasing amount of expensive "home fertility kits", there is no solid evidence for the use of these kits. More importantly, when faced with low levels of ovarian reserve tests some fertility clinics either immediately start fertility treatment, or refuse a woman treatment based on "a reduced ovarian reserve". Based on the present research, there is no sufficient evidence available for this practice.

The use of ovarian reserve tests in current clinical practice

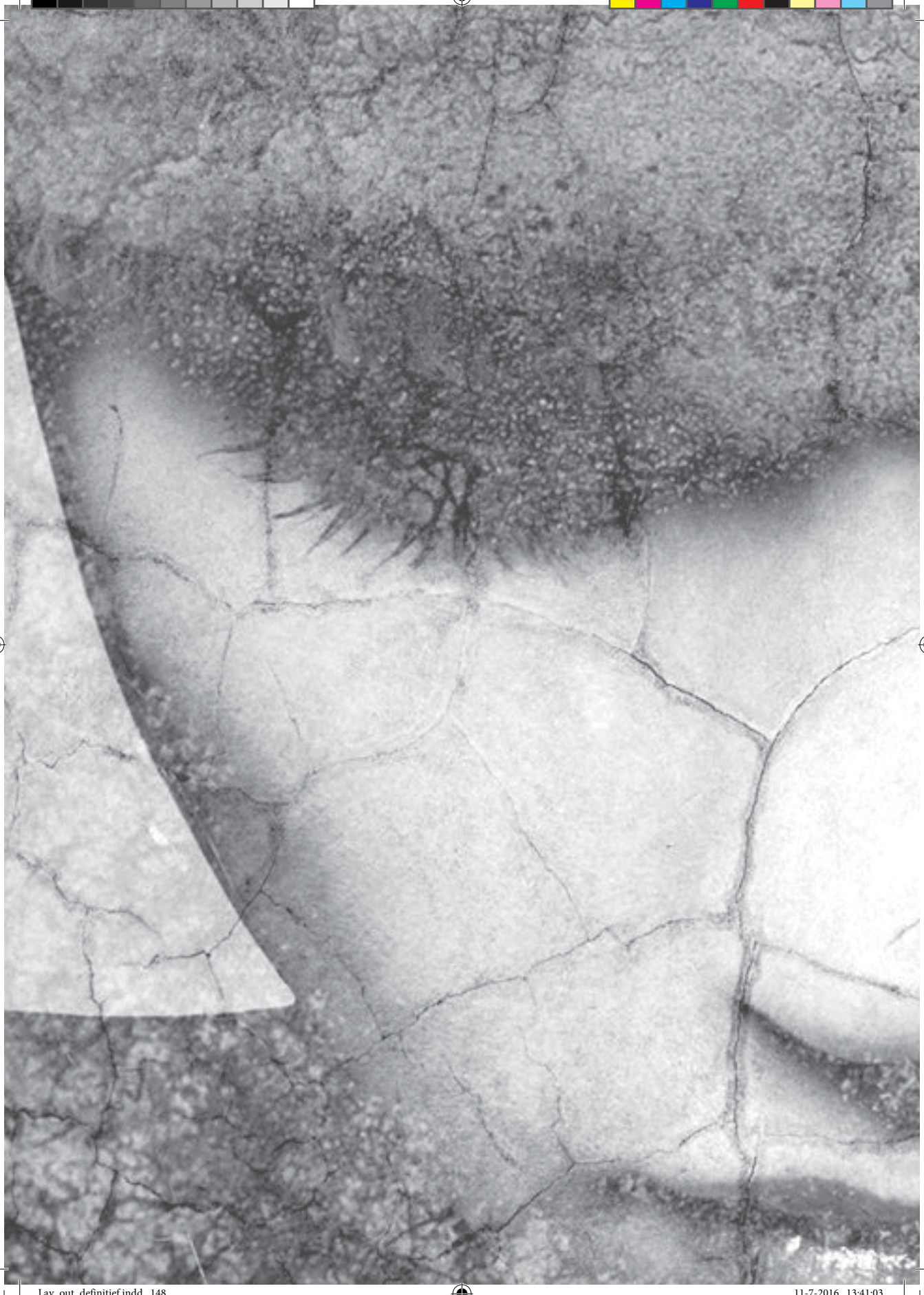
This thesis sought out to research the capacity of ovarian reserve tests in the prediction of current and future fertility. Despite the fact that there clearly is a relation between the true ovarian reserve (i.e. the non- growing follicle pool) and age at natural menopause, ovarian reserve tests currently fail to predict individual age at natural menopause or current fertility with a clinically relevant precision. This thesis clearly revealed that based on current knowledge, there is no place for AMH based fertility assessment or menopause prediction in clinical practice.

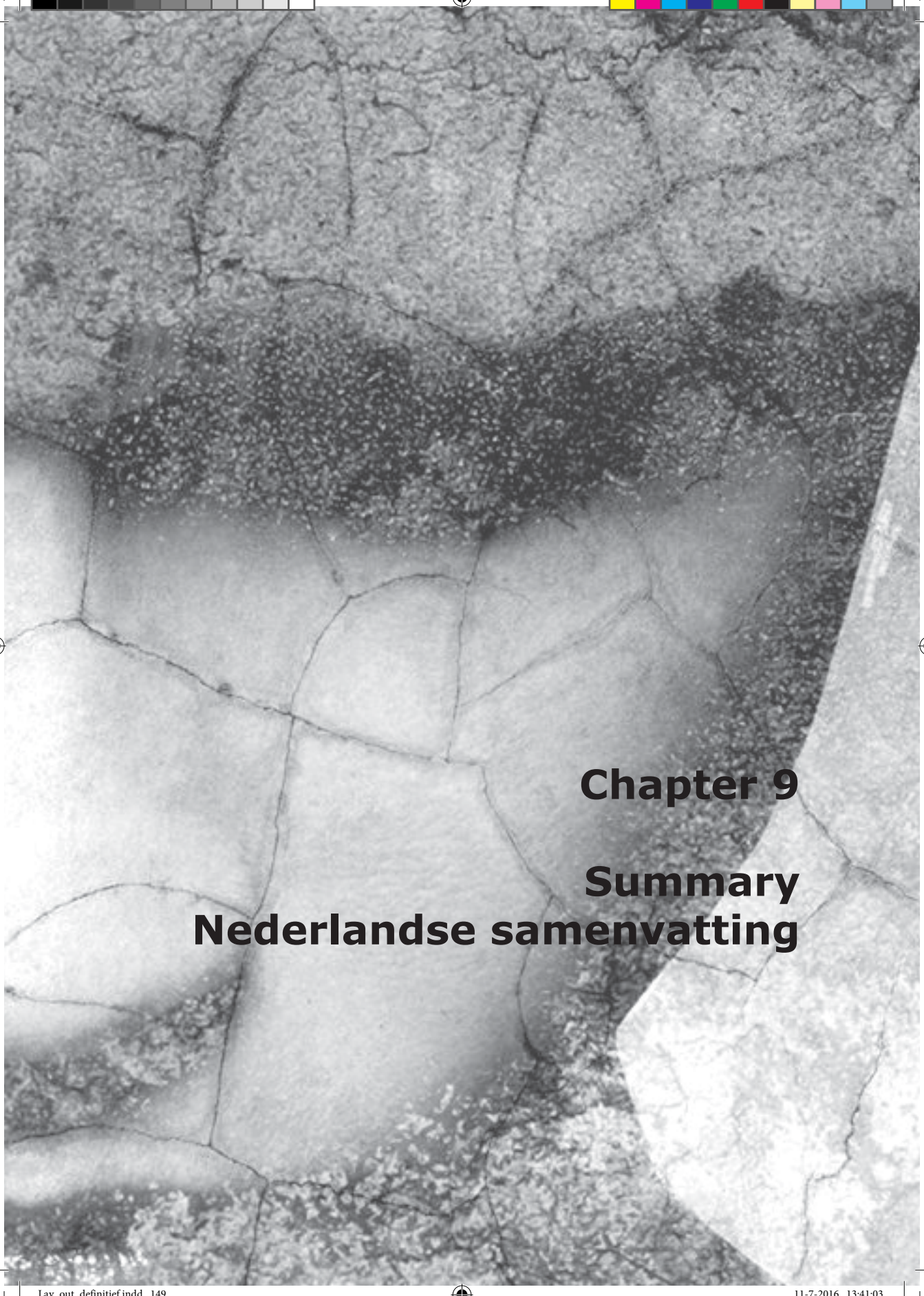
Directions for future research

This thesis identified directions for future research. The strong relation between the true ovarian reserve and age at natural menopause as portrayed in chapter two left an important question unanswered. This question arises from the fact that predictions were based on the assumption that a constant and uniform decline in the number of follicles from conception to menopause is present between women. Since it is not possible to perform an in vivo analysis of this assumption, and based on the relation between the true ovarian reserve and ORTs, decline patterns should be researched using repeatedly measured ORT levels. These measurements should take place over a time span of numerous years.

The pitfalls that became apparent in AMH based time to menopause analysis, as discussed in chapters four-six, would benefit from new research. As explained, a large cohort including women of all ages, irrespective of cycle state and offering a long follow up could undo the fact that menopause predictions do

not cover the full age ranges known for this event. Furthermore, performing analysis using repeatedly measured AMH with a rigid assay methodology at a fixed time in relation to the menstrual cycle, as explained in chapter three, could narrow prediction intervals. Narrowing of the prediction interval could also be achieved by incorporating currently unknown factors contributing to age at menopause in prediction models. Research should therefore be aimed at identifying these factors. Lastly, the observation of a non- proportional AMH effect, as revealed in chapter five and six, with a strong AMH effect present in young women leaves the door for AMH slightly ajar, although the current imprecisions surrounding AMH based predictions need to be resolved before this marker could be used in the day to day clinical practice.





Chapter 9

Summary
Nederlandse samenvatting

Summary

The objective of this thesis was to assess the value of ovarian reserve tests in the prediction of the reproductive life span and current fertility prospects.

Chapter 1
In the introduction of this thesis the relation between ovarian reserve tests (ORTs) and the primordial (or non- growing) follicle pool was determined. Moreover, it was addressed that in order to provide any information on current fertility or on the length of the individual reproductive life span, age at natural menopause must be used as a proxy variable. The hypothesized fixed 10-year interval between the end of natural fertility and menopause opened avenues for the use of this proxy variable. Since decline in follicle number and follicle quality go hand in hand, ORTs, reflective of the quantitative ovarian reserve, can be used to predict current fertility and to demarcate the fertile life span.

Chapter 2
In this chapter age at natural menopause was assessed based on the non-growing follicle (NGF) pool, also known as the true ovarian reserve. NGF counts were obtained from 8 different databases containing histologically derived NGF counts in human ovaries. Age at natural menopause was derived from the Prospect-EPIC database. This Dutch population based cohort contains data on 17357 women aged 50-70 years. After applying selection criteria, 4037 women provided data on age at natural menopause. Using a robust regression analysis, a close conformity was observed between NGF based predicted age at menopause and actual observed age at menopause. Moreover, a critical threshold in the number of NGFs present within a single ovary could be calculated. At this threshold the ovary can no longer maintain a menstrual cycle and menopause consequently occurs. The close conformity observed between NGF based predicted age at menopause and observed age at menopause supports the theory that the endowed size of the primordial follicle is an important determinant for age at menopause. This study provided a solid base for the use of derivatives of the primordial follicle pool as markers of ovarian reserve (i.e. ORTs) in the prediction of menopause.

Chapter 3
This chapter describes the relation between fluctuations in peripheral levels of anti-Müllerian hormone (AMH) and the antral follicle count (AFC). This study

aimed to improve the understanding of the origin of AMH and thereby aiding interpretation of these AMH levels. The number of antral follicles and peripheral levels of AMH were measured repeatedly during a full menstrual cycle. A mixed model analysis demonstrated that fluctuations in circulating levels of AMH were paralleled by a fluctuation in the number of antral follicles (2-8 mm). These results identify antral follicles sized 2-8mm as the dominant source for circulating AMH. Furthermore, these findings supported the use of AMH in clinical practice. For instance, a high level of peripheral AMH in women with polycystic ovarian syndrome was made understandable since they express a high antral follicle count. And on the other side of the spectrum low AMH levels reflective of a diminished ovarian reserve in women with premature ovarian insufficiency were made comprehensible since the diminished ovarian reserve in these women is reflected by a low AFC.

Chapter 4

Here an extensive literature review is presented regarding the current status of menopause prediction based on either ORTs or mother's age at natural menopause (ANM). Furthermore, directions for further research were identified. This review focused on AMH, the AFC and mother's ANM in the prediction of individual age at menopause (or time to menopause) and yielded 7 useful studies. Together these studies identified AMH and mother's ANM as the strongest predictors of age at menopause. However, prediction intervals remained unsatisfactory wide and predictions lacked the prediction of the extreme ages at menopause (i.e. the early or late menopausal ages). It was hypothesized that these problems originated from small numbers of women included in the studies, and thereby causing an underrepresentation of women with an extreme age at menopause, as these are rare events. Moreover, by applying the inclusion criteria "having a regular cycle", women with irregular cycles were excluded. Exactly these women could have irregular cycles due to a diminished ovarian reserve and would therefore be more likely to experience menopause at a young age. Also not all factors contributing to age at menopause were identified or incorporated in analysis. These pitfalls in the studies and analyses result in the fact that there is currently not enough evidence for the use of AMH and mother's ANM in the prediction of age at menopause in the day-to-day clinical practice. Future research should address these problems and therefore the prediction models need to be studied in in a large cohort including young women irrespective of cycle state and providing a long follow up time.

Chapter 5

This chapter provides the results of an ongoing prospective cohort study researching the predictive capacity of ORTs in the prediction of age at natural menopause. The results of the first follow up round (2009) were previously published. 265 normo-ovulatory were included between 1992 and 2001. The mean time to follow up was 14 years. At this follow up round 37.5% of women had reached menopause. A Cox regression analysis assessing time to menopause was performed using ovarian reserve tests (AMH, AFC and follicle stimulating hormone (FSH)). Furthermore, a Weibull survival model was composed assessing individual age at menopause and a non-proportionality analysis of AMH over time was performed. All analyses were performed correcting for female age. We demonstrated AMH to be the only significant predictor of time to menopause in a model next to female age. In line with previous publications and the review presented in chapter 4, ANM predictions did not cover the extreme ages at menopause (i.e. very early or late menopause) and prediction intervals remained wide. As stated above, this could be caused by underrepresentation of women experiencing extreme ages at menopause due to the combination of the rareness of the event and a small sample size. Also contributing to the wide intervals could be imprecision of AMH itself due to assay instability, together with the fact that not all factors contributing to age at menopause were taking into consideration. Moreover, the non- proportionality analysis demonstrated the predictive capacity of AMH to decline with increasing age. As women with the highest hazard of reaching menopause became menopausal in the previous follow up round, only women with a relative low hazard of menopause are represented in the current follow up round. The average hazard of menopause at the most recent follow up round will thus be lower than the hazard of the whole group at the previous rounds, resulting in a reduced AMH effect. This could explain why the predictions of age at menopause are concentrated around the general mean age at menopause, since a reduced predictive effect results in the fact that AMH does not shift the general age at menopause to a large extend. Our results clearly state that there is currently no place for AMH in the prediction of age at menopause, or in the identification of the reproductive life span in a clinical setting.

Chapter 6

This chapter describes the results of an individual patient data (IPD) meta-analysis researching AMH based predictions of age at menopause. In the previous chapters it became apparent that a small study size was a potential

problem for AMH based menopause prediction models. An IPD meta-analysis would overcome this problem as it combines all known cohorts and thereby creates a large database for analysis.

An extensive literature search was performed in order to identify all cohorts regarding AMH based menopause predictions. Identified cohorts were invited to participate in this IPD meta-analysis. Of the seven cohorts eligible for inclusion, 5 were willing to participate. The data of these five cohorts were merged into a large summary database containing 2518 women. Based on a check for heterogeneity, AMH was assessed in age-specific percentiles per study to overcome study associated heterogeneity. A significant predictive capacity of AMH in the prediction of age at menopause was described. Also, after correction for female age this predictive capacity remained significantly present, but the added value to predictions based on female age alone was poor. Moreover, the previously identified pitfalls in AMH based menopause predictions remained present in these analyses. Although this large study population did result in identification of women experiencing menopause at a relatively late age, the prediction model still lacks the capacity to identify women with an early age at menopause. Moreover, this combined larger cohort, did not narrow the prediction intervals of age at menopause. It was hypothesized that the wide prediction interval originated from the lack of incorporating all factors that contribute to menopause prediction in the model. Due to differences between the available variables between the 5 cohorts, some variables potentially contributing to age at menopause had to be excluded as not all cohorts could provide these data. Lastly, and in accordance with chapter 5 a strong non-proportional effect of AMH was observed resulting in a decrease of predictive capacity with increasing age. As the pitfalls in AMH based prediction models were not overcome by performing an IPD meta-analysis, we still believe that future research regarding menopause prediction should focus on a large cohort of women included at a young age and irrespective of cycle state. Moreover, research should focus on identifying unknown factors contributing to age at menopause in order to incorporate these factors into prediction models.

Chapter 7

Here we describe the results of a prospective cohort study investigating the predictive capacity of ORTs in the prediction of current fertility prospects. Between 1999 and 2001, 102 healthy women planning a pregnancy were included. The capacity of ORTs to predict an ongoing pregnancy was studied. Follow up time was completed one year after cessation of contraceptives or

when an ongoing pregnancy was achieved. After completion of follow up, a semen analysis and a test for chlamydia antibody titers was performed. A time to pregnancy analysis based on a baseline ORT measurement (AMH, AFC, FSH), semen parameters and chlamydia antibody titers was performed using a Cox regression analysis. Corrections were made for female age. We demonstrated that none of these parameters researched could significantly predict time to ongoing pregnancy. The only significant predictor was female age. Our results clearly demonstrate that there is no evidence for the use of ovarian reserve tests in the assessment of current fertility. These results make the use of ovarian reserve tests obsolete in either the clinical setting or in home fertility tests.

Chapter 8

In the final chapter of this thesis we reflected on our findings and report recommendations for future research. In this thesis we clearly demonstrated that based on current knowledge, there is no place for ovarian reserve tests in the identification of the reproductive lifespan and current fertility prospects. We therefore strongly discourage the use of AMH and other ovarian reserve tests in such clinical practice.

Future research should first be aimed at assessing individual decline rates since ORT based predictions of reproductive lifespan and current fertility are based on the assumption that a fixed and uniform decline rate is present for all women.

Next, in order to finally determine if there is any place for AMH in the prediction of the reproductive lifespan, AMH based age a menopause prediction should be researched in a cohort capable of overcoming current AMH pitfalls. These pitfalls are the lack of prediction of the full age range of menopause, broad prediction intervals and a non- proportional AMH effect. This cohort should include a large group of women, the follow up time should be extensive and women need to be included irrespective of cycle state. Moreover repeated AMH measurement are required using a rigid assay.

In conclusion, nowadays there is no evidence supporting the use of AMH in clinical practice for the prediction of the reproductive life span and current fertility prospects.

Samenvatting

Het doel van dit proefschrift was het evalueren van de rol van ovariële reserve testen in het voorspellen van de lengte van de fertiele levensfase en in het voorspellen van de huidige vruchtbaarheid.

Hoofdstuk 1
In de introductie van dit proefschrift werd de relatie tussen ovariële reserve testen (ORTs) en de rustende follikel poel (NGF) geschetst. Voorts werd uiteengezet dat om informatie te verschaffen over de huidige vruchtbaarheid of de lengte van de fertiele levensfase, gebruik kan worden gemaakt van de leeftijd ten tijde van menopauze als afgeleide variabele. Dit kan gedaan worden op basis van het veronderstelde interval van 10 jaar tussen het einde van de natuurlijke vruchtbaarheid en de menopauze. Gezien het feit dat de afname in het aantal follikels en in de kwaliteit van eicellen met elkaar verbonden zijn, kunnen ORTs gebruikt worden om de huidige vruchtbaarheid en de lengte van de fertiele levensfase te bepalen.

Hoofdstuk 2
In dit hoofdstuk werd de leeftijd ten tijde van de natuurlijke menopauze bepaald op basis van de rustende follikel poel (NGF). NGF telling werden verkregen uit 8 verschillende datasets waarin histologische tellingen werden verricht in humaan ovarieel weefsel. Data betreffende de leeftijd ten tijde van menopauze werd verkregen uit de Prospect-EPIC dataset. Dit is een Nederlands cohort welk data bevat van 17357 vrouwen in leeftijd variërend van 50-70 jaar. Na het toepassen van enkele selectiecriteria kon data van 4037 vrouwen worden gebruikt.
Gebruik makend van een robuuste regressie analyse werd een sterke overeenkomst gezien tussen de voorspelde leeftijd ten tijde van de menopauze op basis van NGF tellingen en de geobserveerde menopauze leeftijden in het Prospect-EPIC cohort. Buiten dit werd een kritieke drempelwaarde in het aantal NGFs berekend in een ovarium waarbij menopauze optreedt.
De sterke overeenkomst tussen de voorspelde en de geobserveerde leeftijd ten tijde van menopauze ondersteunt de theorie dat de bij de geboorte verkregen hoeveelheid NGFs een belangrijke determinant is voor de leeftijd ten tijde van menopauze. Deze studie verschaft een solide basis voor het gebruik van ORTs in het voorspellen van menopauze.

Hoofdstuk 3

Dit hoofdstuk beschrijft de relatie tussen fluctuaties in perifere anti-Mullers hormoon (AMH) spiegels en het aantal antrale follikels (AFC). De studie had als doel het bevorderen van het begrip over de herkomst van AMH en de interpretatie van AMH waarden.

De AFC en perifere AMH spiegels werden herhaaldelijk gemeten binnen een menstruatie cyclus. Een mixed model analyse toonde aan dat fluctuaties in de AFC (follikels 2-8mm) worden weerspiegeld in fluctuaties in AMH. De resultaten identificeerden antrale follikels van 2-8mm als dominante bron voor circulerend AMH. Daarnaast ondersteunden deze bevindingen het gebruik van AMH in de kliniek. Een hoog AMH bij een vrouw met het polycysteus ovarium syndroom kan bijvoorbeeld worden verklaard door een hoge AFC. Aan de andere kant van het spectrum staan dan vrouwen met een laag AMH, bij een premature ovariële insufficiëntie en een lage AFC.

Hoofdstuk 4

Dit hoofdstuk beschrijft een literatuurstudie waarin de mogelijkheid van het voorspellen van de menopauze leeftijd op basis van AMH, de AFC en moeders leeftijd ten tijde van de menopauze werd onderzocht. Verder werd de richting van toekomstig onderzoek gedefinieerd.

7 studies werden geïdentificeerd, deze studies toonden aan dat AMH en moeders leeftijd ten tijde van menopauze de sterkste voorspellers zijn voor menopauze leeftijd. Echter, predictie intervallen bleven breed en de extreme leeftijden ten tijde van menopauze (zeer vroeg of zeer laat) werden niet voorspeld. Het werd verondersteld dat deze predictieproblemen voortkwamen uit het feit dat de studies een klein aantal vrouwen bevatten waardoor vrouwen met een extreme menopauze leeftijd ondervertegenwoordigd raakten, zeker gezien het feit dat een extreme menopauze leeftijd zeldzaam is. Daarnaast zorgde het toegepaste inclusiecriterium 'een regulaire cyclus' ervoor dat vrouwen die op jonge leeftijd een irregulaire cyclus hebben, als uiting van premature ovariële insufficiëntie, werden uitgesloten van inclusie. Bovendien werden niet alle factoren die kunnen bijdragen aan het optreden van menopauze geïdentificeerd en meegenomen in de analyses.

Deze beperkingen in de predictie van menopauze resulteerden in de conclusie dat er momenteel onvoldoende bewijs is voor het gebruiken van AMH, de AFC of moeders leeftijd ten tijde van menopauze voor menopauze predictie in de dagelijkse praktijk.

Toekomstig onderzoek zal zich moeten richten op het toepassen van

predictiemodellen in grote cohorten waarin vrouwen op jonge leeftijd worden geïncludeerd ongeacht de reproductieve fase en met een langdurige follow up.

Hoofdstuk 5

Dit hoofdstuk beschrijft de resultaten van een prospectieve cohort studie waarin werd onderzocht of ORTs de leeftijd ten tijde van menopauze kunnen voorspellen. De resultaten van de eerste follow up ronde werden eerder gepubliceerd (2009).

265 vrouwen met een regulaire cyclus werden geïncludeerd tussen 1991 en 2001. De follow up tijd bedroeg 14 jaar. Ten tijde van deze follow up ronde bleek 37.5% van de vrouwen postmenopauzaal.

Een Cox regressie analyse werd verricht en analyseerde tijd tot menopauze op basis van ORTs (AMH, AFC en follikel stimulerend hormoon (FSH)). Verder werd een Weibull model opgesteld teneinde individuele leeftijd ten tijde van menopauze te voorspellen op basis van AMH en werd een non- proportionaliteit analyse verricht van AMH in de tijd. In alle analyses werd gecorrigeerd voor leeftijd.

We toonden aan dat alleen AMH een significante voorspeller is van tijd tot menopauze in een model naast leeftijd. Echter, en in overeenstemming met andere publicaties en met hoofdstuk 4, de voorspellingen van menopauze leeftijd bleken niet het volledige spectrum van leeftijden waarin menopauze optreedt te bevatten en predictie intervallen bleken breed. Zoals beschreven in hoofdstuk 4 veronderstelden we dat dit voort komt uit het feit dat vrouwen met een extreme menopauze leeftijd ondervertegenwoordigd zijn door de combinatie van een zeldzame aandoening en een klein cohort. Daarnaast droegen onnauwkeurigheid van AMH en het feit dat niet alle factoren die bijdragen aan leeftijd ten tijde van menopauze bekend zijn en deze derhalve dus niet meegenomen worden in het predictiemodel, bij aan de brede predictie intervallen.

De non- proportionaliteit analyse toonde aan dat de voorspelkracht van AMH afneemt met het voortschrijden van de leeftijd. Daar vrouwen met het hoogste risico op menopauze reeds postmenopauzaal waren ten tijde van de eerste follow up ronde blijven enkel vrouwen met een lager menopauze risico over. Derhalve is de gemiddelde kans op menopauze deze ronde lager dan voorheen, iets wat weerspiegeld wordt in een afgenomen predictief effect van AMH. Dit is ook de reden dat voorspellingen betreffende menopauze leeftijden zich centreren rondom de bekende gemiddelde leeftijd van 51 jaar. Een afgenomen AMH effect verschuift de voorspelde leeftijd ten tijde van de menopauze

immers niet sterk ten opzichte van de gemiddelde leeftijd. Onze resultaten tonen duidelijk aan dat er momenteel geen plaats is voor het gebruik van AMH in het voorspellen van de menopauze, of in het afbakenen van de fertiele levensfase.

Hoofdstuk 6

Dit hoofdstuk beschrijft de resultaten van een individuele patiënt data meta-analyse (IPD) waarin het voorspellen van de menopauze leeftijd op basis van AMH werd onderzocht. In eerdere hoofdstukken werd duidelijk dat studie aantallen een potentieel probleem zijn voor predictiemodellen. Een IPD analyse waarin data wordt gepoold en dus een grote dataset wordt gecreëerd, zal dit probleem mogelijk dan ook oplossen.

De literatuur werd doorzocht teneinde alle cohorten te identificeren die onderzoek doen naar het voorspellen van de menopauze op basis van AMH. Geïdentificeerde cohorten werden gevraagd deel te nemen aan deze studie. Van de 7 beschikbare cohorten, bleken er 5 bereid de data te delen. De data van deze 5 cohorten werd samengevoegd tot een grote dataset van 2518 vrouwen. Gebaseerd op een controle voor heterogeniteit werd besloten AMH te bestuderen middels leeftijd specifieke AMH percentielen, zodoende werd het effect van heterogeniteit tussen de verschillende studies verminderd.

AMH bleek een significante voorspeller van tijd tot menopauze in een model naast leeftijd. Echter, de aanvullende waarde op voorspellingen enkel op basis van leeftijd was matig. Bovendien bleken de eerdere benoemde problemen in menopauze voorspellingen op basis van AMH ook nu aanwezig. Ondanks het feit dat in deze grotere dataset een latere menopauze leeftijd beter voorspeld werd, werden vrouwen met een vroege menopauze leeftijd nog altijd niet geïdentificeerd. Buiten dit bleven de predictie intervallen breed. Het werd verondersteld dat de brede predictie intervallen voortkwamen uit het gebruik van diverse AMH assays en uit het feit dat niet alle factoren die bijdragen aan AMH bekend zijn en derhalve niet konden worden geïncludeerd in het predictiemodel. Bovendien moesten sommige variabelen waarvan bekend is dat ze bijdragen aan de leeftijd ten tijde van menopauze worden geëxcludeerd omdat niet alle deelnemende studies over deze data beschikten.

Tot slot bleek ook hier een sterk non- proportioneel AMH effect aanwezig met een afname van het AMH effect met het voortschrijden van leeftijd.

Gezien het feit dat een IPD analyse de problemen gepaard gaande met het voorspellen van de menopauze op basis van AMH niet konden oplossen pleiten we opnieuw voor nieuw onderzoek. Dit onderzoek zal verricht moeten worden

binnen een groot cohort waarin vrouwen op jonge leeftijd, ongeacht de reproductieve fase, worden geïncludeerd en gedurende een lange tijd worden vervolgd. Daarnaast zal onderzoek zich moeten richten op het identificeren van tot op heden onbekende factoren die bijdragen aan de leeftijd van menopauze, zodoende kunnen deze factoren worden meegenomen in predictie modellen.

Hoofdstuk 7

Hier beschrijven we de resultaten van een prospectieve studie waarin werd gekeken naar de capaciteit van ORTs in het voorspellen van tijd tot doorgaande zwangerschap. Tussen 1999 en 2001 werden 102 gezonde paren met een kinderwens geïncludeerd. De follow up tijd was een jaar, tenzij er eerder een doorgaande zwangerschap optrad. Wanneer de studie voltooid was werd een semen analyse verricht en werd de vrouw onderzocht op de aanwezigheid van chlamydia antistoffen. Tijd tot zwangerschap werd geanalyseerd op basis van een Cox regressie analyse, hierin werd gecorrigeerd voor leeftijd. Geen van de ORTs onderzocht kon tijd tot doorgaande zwangerschap significant voorspellen. De enige voorspeller bleek de leeftijd van de vrouw. Deze resultaten tonen aan dat er geen plek is voor het gebruik van ORTs bij het bepalen van de huidige vruchtbaarheid. Deze studie pleit derhalve tegen het gebruik van ORTs in fertiliteitklinieken of in zogenoemde 'thuistesten'.

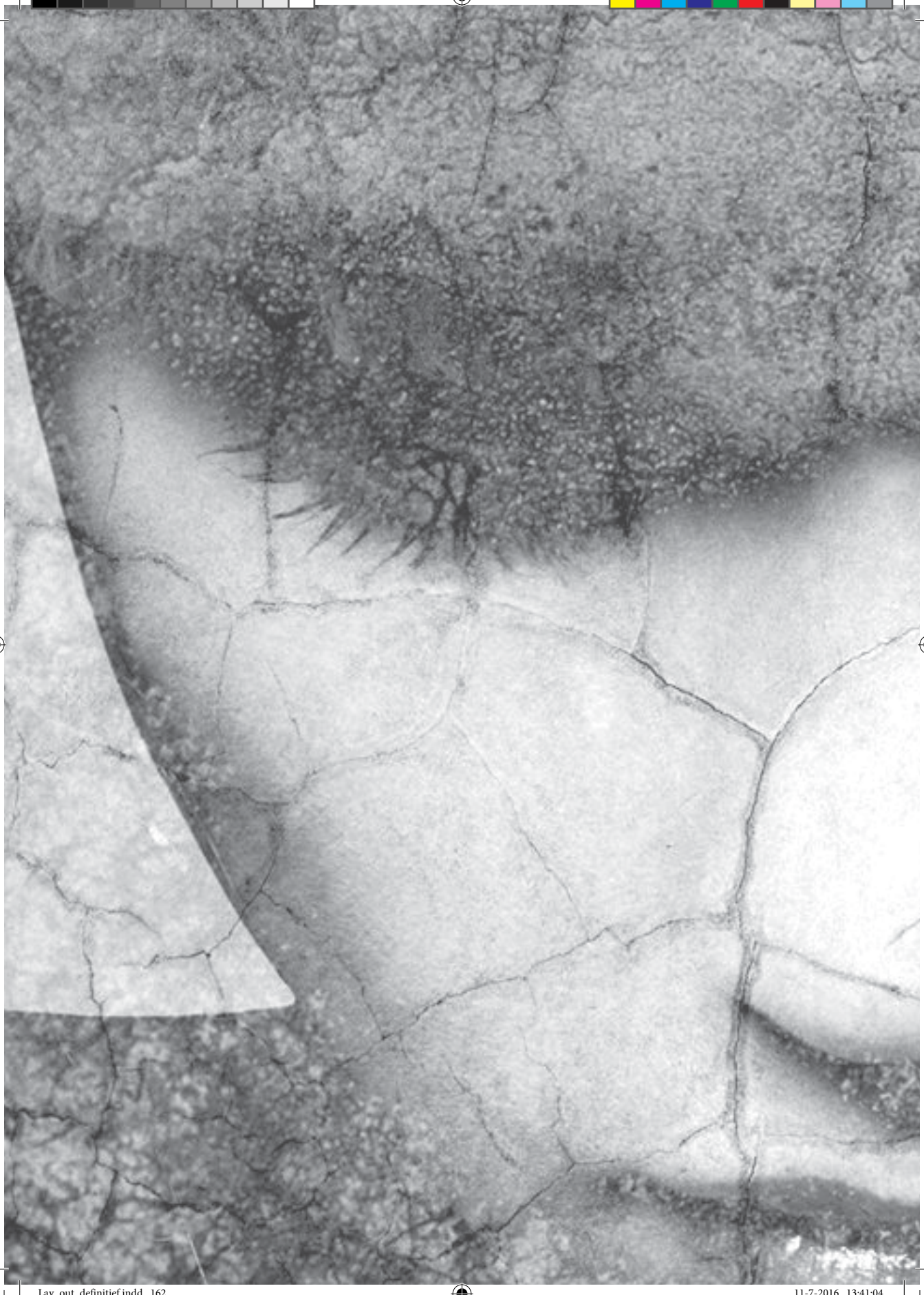
Hoofdstuk 8

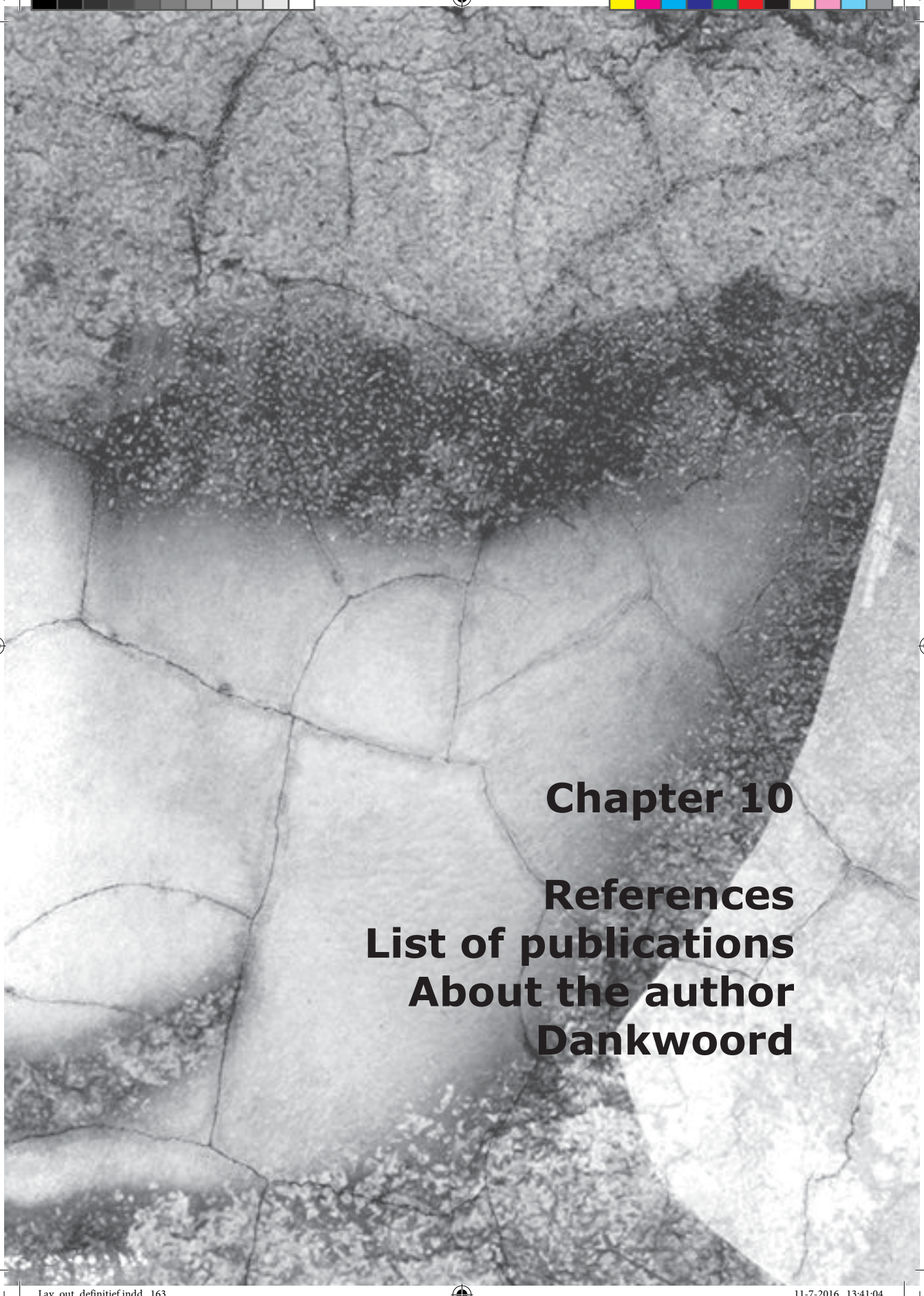
In dit hoofdstuk reflecteren we op onze bevindingen en doen we aanbevelingen voor toekomstig onderzoek.

Toekomstig onderzoek moet zich richten op het onderzoeken van individuele afname snelheden in het aantal rustende follikels. Dit gezien het feit dat op ORT gebaseerde voorspellingen aangaande de reproductieve levensfase, uitgaan van een vaststaande en uniforme afname snelheid in follikelaantallen voor alle vrouwen en deze aanname onvoldoende is onderzocht.

Daarnaast en teneinde definitief te kunnen vaststellen of er plek is voor AMH in het voorspellen van de reproductieve levensfase, zal menopauze predictie op basis van AMH moeten worden onderzocht in een cohort dat de huidige problemen kan oplossen. Deze problemen zijn het onvermogen de extreme leeftijden ten tijde van menopauze te voorspellen, de brede predictie intervallen en de non-proportionaliteit van het AMH effect. Dit cohort zal een grote groep vrouwen moeten includeren, ongeacht de reproductieve fase en moeten zorgen voor een langdurige follow up. Verder adviseren wij herhaalde AMH metingen te verrichten met een stabiel assay.

In dit proefschrift werd aangetoond dat, gebaseerd op de huidige kennis, er geen plek is voor ORTs in het identificeren van de reproductieve levensfase of het voorspellen van de huidige vruchtbaarheid. We pleiten we dan ook tegen het gebruik van AMH en andere ORTs voor het afbakenen van de reproductieve levensfase in de kliniek.





Chapter 10

References
List of publications
About the author
Dankwoord

References

1. Habbema JD, Eijkemans MJ, Leridon H, te Velde ER. Realizing a desired family size: When should couples start? *Hum Reprod*. 2015;30(9):2215-2221. doi: 10.1093/humrep/dev148 [doi].

2. Eijkemans MJ, van Poppel F, Habbema DF, Smith KR, Leridon H, te Velde ER. Too old to have children? lessons from natural fertility populations. *Hum Reprod*. 2014;29(6):1304-1312. doi: 10.1093/humrep/deu056 [doi].

3. Treloar AE. Menstrual cyclicity and the pre-menopause. *Maturitas*. 1981;3(3-4):249-264.

4. te Velde ER, Pearson PL. The variability of female reproductive ageing. *Hum Reprod Update*. 2002;8(2):141-154.

5. Eijkemans M, Habbema J, te Velde E. Age at last childbirth and fertility at young age. In: *Eijkemans MJ, academic thesis; Fertility in populations and in patients.* ; 2004:23.

6. de Boer EJ, den Tonkelaar I, te Velde ER, Burger CW, van Leeuwen FE, OMEGA-project group. Increased risk of early menopausal transition and natural menopause after poor response at first IVF treatment. *Hum Reprod*. 2003;18(7):1544-1552.

7. Daan NM, Hoek A, Corpeleijn E, et al. Reproductive characteristics of women diagnosed with premature ovarian insufficiency. *Reprod Biomed Online*. 2016;32(2):225-232. doi: 10.1016/j.rbmo.2015.11.009 [doi].

8. Nybo Andersen AM, Wohlfahrt J, Christens P, Olsen J, Melbye M. Maternal age and fetal loss: Population based register linkage study. *BMJ*. 2000;320(7251):1708-1712.

9. Wallace WH, Kelsey TW. Human ovarian reserve from conception to the menopause. *PLoS One*. 2010;5(1):e8772. doi: 10.1371/journal.pone.0008772; 10.1371/journal.pone.0008772.

10. Bjelland EK, Wilkosz P, Tanbo TG, Eskild A. Is unilateral oophorectomy associated with age at menopause? A population study (the HUNT2 survey). *Hum Reprod*. 2014;29(4):835-841. doi: 10.1093/humrep/deu026 [doi].

11. den Tonkelaar I. Validity and reproducibility of self-reported age at menopause in women participating in the DOM-project. *Maturitas*. 1997;27(2):117-123. doi: S0378-5122(97)01122-5 [pii].

12. Colditz GA, Stampfer MJ, Willett WC, et al. Reproducibility and validity of self-reported menopausal status in a prospective cohort study. *Am J Epidemiol*. 1987;126(2):319-325.

13. Leridon H. Demographic effects of the introduction of steroid contraception in developed countries. *Hum Reprod Update*. 2006;12(5):603-616. doi: dml025 [pii].

14. Broekmans FJ, Faddy MJ, Scheffer G, te Velde ER. Antral follicle counts are related to age at natural fertility loss and age at menopause. *Menopause*. 2004;11(6 Pt 1):607-614. doi: 00042192-200411060-00007 [pii].

15. van Disseldorp J, Faddy MJ, Themmen AP, et al. Relationship of serum antimullerian hormone concentration to age at menopause. *J Clin Endocrinol Metab*. 2008;93(6):2129-2134. doi: 10.1210/jc.2007-2093; 10.1210/jc.2007-2093.

16. Sowers MR, McConnell D, Yosef M, Jannausch ML, Harlow SD, Randolph JF,Jr. Relating smoking, obesity, insulin resistance, and ovarian biomarker changes to the final menstrual period. *Ann N Y Acad Sci*. 2010;1204:95-103. doi: 10.1111/j.1749-6632.2010.05523.x; 10.1111/j.1749-6632.2010.05523.x.

17. Tehrani FR, Shakeri N, Solaymani-Dodaran M, Azizi F. Predicting age at menopause from serum antimullerian hormone concentration. *Menopause*. 2011;18(7):766-770. doi: 10.1097/gme.0b013e318205e2ac [doi].

18. 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;96(8):2532-2539. doi: 10.1210/jc.2010-2776; 10.1210/jc.2010-2776.

19. 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;97(5):1673-1680. doi: 10.1210/jc.2011-3032 [doi].

20. Wellons MF, Bates GW, Schreiner PJ, Siscovick DS, Sternfeld B, Lewis

CE. Antral follicle count predicts natural menopause in a population-based sample: The coronary artery risk development in young adults women's study. *Menopause*. 2013;20(8):825-830. doi: 10.1097/GME.0b013e31827f06c2 [doi].

21. 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;98(2):729-735. doi: 10.1210/jc.2012-3176; 10.1210/jc.2012-3176.

22. Dolleman M, Depmann M, Eijkemans MJ, et al. Anti-mullerian hormone is a more accurate predictor of individual time to menopause than mother's age at menopause. *Hum Reprod*. 2014;29(3):584-591. doi: 10.1093/humrep/det446 [doi].

23. Dolleman M, Verschuren WM, Eijkemans MJ, Broekmans FJ, van der Schouw YT. Added value of anti-mullerian hormone in prediction of menopause: Results from a large prospective cohort study. *Hum Reprod*. 2015. doi: dev145 [pii].

24. Nair S, Slaughter JC, Terry JG, et al. Anti-mullerian hormone (AMH) is associated with natural menopause in a population-based sample: The CARDIA women's study. *Maturitas*. 2015;81(4):493-498. doi: 10.1016/j.maturitas.2015.06.026 [doi].

25. Faddy MJ, Gosden RG, Gougeon A, Richardson SJ, Nelson JF. Accelerated disappearance of ovarian follicles in mid-life: Implications for forecasting menopause. *Hum Reprod*. 1992;7(10):1342-1346.

26. Guerif F, Lemseffer M, Couet M-, Gervereau O, Ract V, Royere D. Serum antimullerian hormone is not predictive of oocyte quality in vitro fertilization. *Ann Endocrinol (Paris)*. 2009;70(4):230-234. doi: 10.1016/j.ando.2009.03.007 [doi].

27. Riggs R, Kimble T, Oehninger S, et al. Anti-mullerian hormone serum levels predict response to controlled ovarian hyperstimulation but not embryo quality or pregnancy outcome in oocyte donation. *Fertil Steril*. 2011;95(1):410-412. doi: 10.1016/j.fertnstert.2010.07.1042 [doi].

28. Broer SL, Broekmans FJ, Laven JS, Fauser BC. Anti-mullerian hormone: Ovarian reserve testing and its potential clinical implications. *Hum Reprod*

Update. 2014. doi: dmu020 [pii].

29. Gougeon A, Ecochard R, Thalabard JC. Age-related changes of the population of human ovarian follicles: Increase in the disappearance rate of non-growing and early-growing follicles in aging women. *Biol Reprod*. 1994;50(3):653-663.

30. Hansen KR, Hodnett GM, Knowlton N, Craig LB. Correlation of ovarian reserve tests with histologically determined primordial follicle number. *Fertil Steril*. 2011;95(1):170-175. doi: 10.1016/j.fertnstert.2010.04.006 [doi].

31. Haadsma ML, Bukman A, Groen H, et al. The number of small antral follicles (2-6 mm) determines the outcome of endocrine ovarian reserve tests in a subfertile population. *Hum Reprod*. 2007;22(7):1925-1931. doi: dem081 [pii].

32. Broekmans FJ, de Ziegler D, Howles CM, Gougeon A, Trew G, Olivennes F. The antral follicle count: Practical recommendations for better standardization. *Fertil Steril*. 2010;94(3):1044-1051. doi: 10.1016/j.fertnstert.2009.04.040 [doi].

33. Scheffer GJ, Broekmans FJ, Bancsi LF, Habbema JD, Looman CW, Te Velde ER. Quantitative transvaginal two- and three-dimensional sonography of the ovaries: Reproducibility of antral follicle counts. *Ultrasound Obstet Gynecol*. 2002;20(3):270-275. doi: 10.1046/j.1469-0705.2002.00787.x [pii].

34. Jeppesen JV, Anderson RA, Kelsey TW, et al. Which follicles make the most anti-mullerian hormone in humans? evidence for an abrupt decline in AMH production at the time of follicle selection. *Mol Hum Reprod*. 2013;19(8):519-527. doi: 10.1093/molehr/gat024 [doi].

35. Weenen C, Laven JS, Von Bergh AR, et al. Anti-mullerian hormone expression pattern in the human ovary: Potential implications for initial and cyclic follicle recruitment. *Mol Hum Reprod*. 2004;10(2):77-83.

36. Fanchin R, Schonauer LM, Righini C, Frydman N, Frydman R, Taieb J. Serum anti-mullerian hormone dynamics during controlled ovarian hyperstimulation. *Hum Reprod*. 2003;18(2):328-332.

37. Harlow SD, Gass M, Hall JE, et al. Executive summary of the stages of reproductive aging workshop + 10: Addressing the unfinished agenda of

staging reproductive aging. *J Clin Endocrinol Metab*. 2012;97(4):1159-1168. doi: 10.1210/jc.2011-3362 [doi].

38. Broekmans FJ, Kwee J, Hendriks DJ, Mol BW, Lambalk CB. A systematic review of tests predicting ovarian reserve and IVF outcome. *Hum Reprod Update*. 2006;12(6):685-718. doi: dml034 [pii].

39. Richardson SJ, Senikas V, Nelson JF. Follicular depletion during the menopausal transition: Evidence for accelerated loss and ultimate exhaustion. *J Clin Endocrinol Metab*. 1987;65(6):1231-1237. doi: 10.1210/jcem-65-6-1231.

40. Wallace WH, Kelsey TW. Ovarian reserve and reproductive age may be determined from measurement of ovarian volume by transvaginal sonography. *Hum Reprod*. 2004;19(7):1612-1617. doi: 10.1093/humrep/deh285.

41. Hansen KR, Knowlton NS, Thyer AC, Charleston JS, Soules MR, Klein NA. A new model of reproductive aging: The decline in ovarian non-growing follicle number from birth to menopause. *Hum Reprod*. 2008;23(3):699-708. doi: 10.1093/humrep/dem408; 10.1093/humrep/dem408.

42. Tehrani FR, Solaymani-Dodaran M, Azizi F. A single test of antimullerian hormone in late reproductive-aged women is a good predictor of menopause. *Menopause*. 2009;16(4):797-802. doi: 10.1097/gme.0b013e318193e95d; 10.1097/gme.0b013e318193e95d.

43. Wood JW. Fecundity and natural fertility in humans. *Oxf Rev Reprod Biol*. 1989;11:61-109.

44. Nelson SM, Telfer EE, Anderson RA. The ageing ovary and uterus: New biological insights. *Hum Reprod Update*. 2013;19(1):67-83. doi: 10.1093/humupd/dms043 [doi].

45. Shuster LT, Rhodes DJ, Gostout BS, Grossardt BR, Rocca WA. Premature menopause or early menopause: Long-term health consequences. *Maturitas*. 2010;65(2):161-166. doi: 10.1016/j.maturitas.2009.08.003; 10.1016/j.maturitas.2009.08.003.

46. Rocca WA, Shuster LT, Grossardt BR, et al. Long-term effects of bilateral oophorectomy on brain aging: Unanswered questions from

the mayo clinic cohort study of oophorectomy and aging. *Womens Health (Lond Engl)*. 2009;5(1):39-48. doi: 10.2217/17455057.5.1.39; 10.2217/17455057.5.1.39.

47. Gallagher JC. Effect of early menopause on bone mineral density and fractures. *Menopause*. 2007;14(3 Pt 2):567-571. doi: 10.1097/gme.0b013e31804c793d.

48. Hartge P. Genetics of reproductive lifespan. *Nat Genet*. 2009;41(6):637-638. doi: 10.1038/ng0609-637 [doi].

49. Block E. Quantitative morphological investigations of the follicular system in women; variations at different ages. *Acta Anat (Basel)*. 1952;14(1-2):108-123.

50. Block E. A quantitative morphological investigation of the follicular system in newborn female infants. *Acta Anat (Basel)*. 1953;17(3):201-206.

51. Baker TG. A quantitative and cytological study of germ cells in human ovaries. *Proc R Soc Lond B Biol Sci*. 1963;158:417-433.

52. Gougeon A, Chainy GB. Morphometric studies of small follicles in ovaries of women at different ages. *J Reprod Fertil*. 1987;81(2):433-442.

53. Bendsen E, Byskov AG, Andersen CY, Westergaard LG. Number of germ cells and somatic cells in human fetal ovaries during the first weeks after sex differentiation. *Hum Reprod*. 2006;21(1):30-35. doi: dei280 [pii].

54. Forabosco A, Sforza C. Establishment of ovarian reserve: A quantitative morphometric study of the developing human ovary. *Fertil Steril*. 2007;88(3):675-683. doi: 10.1016/j.fertnstert.2006.11.191.

55. Boker LK, van Noord PA, van der Schouw YT, et al. Prospect-EPIC utrecht: Study design and characteristics of the cohort population. european prospective investigation into cancer and nutrition. *Eur J Epidemiol*. 2001;17(11):1047-1053.

56. Riboli E, Hunt KJ, Slimani N, et al. European prospective investigation into cancer and nutrition (EPIC): Study populations and data collection. *Public Health Nutr*. 2002;5(6B):1113-1124. doi: 10.1079/PHN2002394.

57. Jones MC, Faddy MJ. A skew extension of the t-distribution, with

applications. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2003;65(1):159-174. doi: 10.1111/1467-9868.00378.

58. Dolleman M, Faddy MJ, van Disseldorp J, et al. The relationship between anti-mullerian hormone in women receiving fertility assessments and age at menopause in subfertile women: Evidence from large population studies. *J Clin Endocrinol Metab*. 2013;98(5):1946-1953. doi: 10.1210/jc.2012-4228; 10.1210/jc.2013-3105.

59. 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;98(5):1254-9.e1-2. doi: 10.1016/j.fertnstert.2012.07.1139 [doi].

60. La Marca A, Sighinolfi G, Papaleo E, Cagnacci A, Volpe A, Faddy MJ. Prediction of age at menopause from assessment of ovarian reserve may be improved by using body mass index and smoking status. *PLoS One*. 2013;8(3):e57005. doi: 10.1371/journal.pone.0057005 [doi].

61. La Marca A, Dondi G, Sighinolfi G, et al. The ovarian response to controlled stimulation in IVF cycles may be predictive of the age at menopause. *Hum Reprod*. 2014;29(11):2530-2535. doi: 10.1093/humrep/deu234 [doi].

62. Broer SL, Mol B, Dolleman M, Fauser BC, Broekmans FJ. The role of anti-mullerian hormone assessment in assisted reproductive technology outcome. *Curr Opin Obstet Gynecol*. 2010;22(3):193-201. doi: 10.1097/GCO.0b013e3283384911 [doi].

63. Iliodromiti S, Kelsey TW, Anderson RA, Nelson SM. Can anti-mullerian hormone predict the diagnosis of polycystic ovary syndrome? A systematic review and meta-analysis of extracted data. *J Clin Endocrinol Metab*. 2013;98(8):3332-3340. doi: 10.1210/jc.2013-1393 [doi].

64. 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(10):4057-4063. doi: jc.2006-0331 [pii].

65. Streuli I, Fraisse T, Pillet C, Ibecheole V, Bischof P, de Ziegler D. Serum antimullerian hormone levels remain stable throughout the menstrual cycle

and after oral or vaginal administration of synthetic sex steroids. *Fertil Steril*. 2008;90(2):395-400. doi: S0015-0282(07)01366-0 [pii].

66. van Disseldorp J, Lambalk CB, Kwee J, et al. Comparison of inter- and intra-cycle variability of anti-mullerian hormone and antral follicle counts. *Hum Reprod*. 2010;25(1):221-227. doi: 10.1093/humrep/dep366 [doi].

67. Robertson DM, Hale GE, Fraser IS, Hughes CL, Burger HG. Changes in serum antimullerian hormone levels across the ovulatory menstrual cycle in late reproductive age. *Menopause*. 2011;18(5):521-524. doi: 10.1097/gme.0b013e3181f8d9e0 [doi].

68. Overbeek A, Broekmans FJ, Hehenkamp WJ, et al. Intra-cycle fluctuations of anti-mullerian hormone in normal women with a regular cycle: A re-analysis. *Reprod Biomed Online*. 2012;24(6):664-669. doi: 10.1016/j.rbmo.2012.02.023 [doi].

69. Hadlow N, Longhurst K, McClements A, Natalwala J, Brown SJ, Matson PL. Variation in antimullerian hormone concentration during the menstrual cycle may change the clinical classification of the ovarian response. *Fertil Steril*. 2013;99(6):1791-1797. doi: 10.1016/j.fertnstert.2013.01.132 [doi].

70. Scheffer GJ, Broekmans FJ, Dorland M, Habbema JD, Looman CW, te Velde ER. Antral follicle counts by transvaginal ultrasonography are related to age in women with proven natural fertility. *Fertil Steril*. 1999;72(5):845-851. doi: S0015-0282(99)00396-9 [pii].

71. van Rooij IA, Broekmans FJ, te Velde ER, et al. Serum anti-mullerian hormone levels: A novel measure of ovarian reserve. *Hum Reprod*. 2002;17(12):3065-3071.

72. Cohen J.
 In: *Statistical power analysis for the behavioral sciences.* 2nd ed. Lawrence Erlbaum, Hillsdale, NJ; 1988:25-26.

73. 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;98(5):2106-2115. doi: 10.1210/jc.2012-3995 [doi].

74. Hemrika DJ, Slaats EH, Kennedy JC, de Vries Robles-Korsen TJ, Schoemaker J. Pulsatile luteinizing hormone patterns in long term oral

contraceptive users. *J Clin Endocrinol Metab*. 1993;77(2):420-426. doi: 10.1210/jcem.77.2.8345046 [doi].

75. Fanchin R, Schonauer LM, Righini C, Guibourdenche J, Frydman R, Taieb J. Serum anti-mullerian hormone is more strongly related to ovarian follicular status than serum inhibin B, estradiol, FSH and LH on day 3. *Hum Reprod*. 2003;18(2):323-327.

76. Eldar-Geva T, Ben-Chetrit A, Spitz IM, et al. Dynamic assays of inhibin B, anti-mullerian hormone and estradiol following FSH stimulation and ovarian ultrasonography as predictors of IVF outcome. *Hum Reprod*. 2005;20(11):3178-3183. doi: dei203 [pii].

77. Freeman EW, Gracia CR, Sammel MD, Lin H, Lim LC, Strauss JF,3rd. Association of anti-mullerian hormone levels with obesity in late reproductive-age women. *Fertil Steril*. 2007;87(1):101-106. doi: S0015-0282(06)03168-2 [pii].

78. van Rooij IA, Broekmans FJ, Scheffer GJ, et al. Serum antimullerian hormone levels best reflect the reproductive decline with age in normal women with proven fertility: A longitudinal study. *Fertil Steril*. 2005;83(4):979-987. doi: S0015-0282(04)03209-1 [pii].

79. Voorhuis M, Onland-Moret NC, van der Schouw YT, Fauser BC, Broekmans FJ. Human studies on genetics of the age at natural menopause: A systematic review. *Hum Reprod Update*. 2010;16(4):364-377. doi: 10.1093/humupd/dmp055 [doi].

80. Wood MA, Rajkovic A. Genomic markers of ovarian reserve. *Semin Reprod Med*. 2013;31(6):399-415. doi: 10.1055/s-0033-1356476 [doi].

81. Kok HS, van Asselt KM, van der Schouw YT, Peeters PH, Wijmenga C. Genetic studies to identify genes underlying menopausal age. *Hum Reprod Update*. 2005;11(5):483-493. doi: dmi024 [pii].

82. Rosen MP, Johnstone EB, Gillham SJ, et al. Is antral follicle count a genetic trait? *Menopause*. 2010;17(1):109-113. doi: 10.1097/gme.0b013e3181b48a88 [doi].

83. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate

healthcare interventions: Explanation and elaboration. *BMJ*. 2009;339:b2700. doi: 10.1136/bmj.b2700 [doi].

84. Azizi F, Rahmani M, Emami H, et al. Cardiovascular risk factors in an iranian urban population: Tehran lipid and glucose study (phase 1). *Soz Praventivmed*. 2002;47(6):408-426.

85. Mohammad K, Sadat Hashemi SM, Farahani FK. Age at natural menopause in iran. *Maturitas*. 2004;49(4):321-326. doi: S0378512204000969 [pii].

86. van der Stroom EM, Konig TE, van Dulmen-den Broeder E, et al. Early menopause in mothers of children with down syndrome? *Fertil Steril*. 2011;96(4):985-990. doi: 10.1016/j.fertnstert.2011.07.1149 [doi].

87. Cutter GR, Burke GL, Dyer AR, et al. Cardiovascular risk factors in young adults. the CARDIA baseline monograph. *Control Clin Trials*. 1991;12(1 Suppl):1S-77S.

88. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: Study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol*. 1988;41(11):1105-1116. doi: 0895-4356(88)90080-7 [pii].

89. Miltenburg GA, Peeters PH, Fracheboud J, Collette HJ. Seventeen-year evaluation of breast cancer screening: The DOM project, the netherlands. diagnostisch onderzoek (investigation) mammacarcinoom. *Br J Cancer*. 1998;78(7):962-965.

90. Cramer DW, Xu H, Harlow BL. Family history as a predictor of early menopause. *Fertil Steril*. 1995;64(4):740-745.

91. Torgerson DJ, Avenell A, Russell IT, Reid DM. Factors associated with onset of menopause in women aged 45-49. *Maturitas*. 1994;19(2):83-92.

92. Torgerson DJ, Thomas RE, Campbell MK, Reid DM. Alcohol consumption and age of maternal menopause are associated with menopause onset. *Maturitas*. 1997;26(1):21-25. doi: S0378512296010754 [pii].

93. Torgerson DJ, Thomas RE, Reid DM. Mothers and daughters menopausal ages: Is there a link? *Eur J Obstet Gynecol Reprod Biol*. 1997;74(1):63-66. doi: S0301211597000857 [pii].

94. Bentzen JG, Forman JL, Larsen EC, et al. Maternal menopause as a predictor of anti-mullerian hormone level and antral follicle count in daughters during reproductive age. *Hum Reprod*. 2013;28(1):247-255. doi: 10.1093/humrep/des356 [doi].

95. Gassner D, Jung R. First fully automated immunoassay for anti-mullerian hormone. *Clin Chem Lab Med*. 2014;52(8):1143-1152. doi: 10.1515/cclm-2014-0022 [doi].

96. Hayatbakhsh MR, Clavarino A, Williams GM, Sina M, Najman JM. Cigarette smoking and age of menopause: A large prospective study. *Maturitas*. 2012;72(4):346-352. doi: 10.1016/j.maturitas.2012.05.004 [doi].

97. Leader B, Baker VL. Maximizing the clinical utility of antimullerian hormone testing in women's health. *Curr Opin Obstet Gynecol*. 2014;26(4):226-236. doi: 10.1097/GCO.0000000000000087 [doi].

98. Bhide P, Gudi A, Shah A, Homburg R. Serum anti-mullerian hormone levels across different ethnic groups: A cross-sectional study. *BJOG*. 2014. doi: 10.1111/1471-0528.13103 [doi].

99. Hoek A, Schoemaker J, Drexhage HA. Premature ovarian failure and ovarian autoimmunity. *Endocr Rev*. 1997;18(1):107-134. doi: 10.1210/edrv.18.1.0291 [doi].

100. Stoop D, Nekkebroeck J, Devroey P. A survey on the intentions and attitudes towards oocyte cryopreservation for non-medical reasons among women of reproductive age. *Hum Reprod*. 2011;26(3):655-661. doi: 10.1093/humrep/deq367 [doi].

101. Mac Dougall K, Beyene Y, Nachtigall RD. Age shock: Misperceptions of the impact of age on fertility before and after IVF in women who conceived after age 40. *Hum Reprod*. 2013;28(2):350-356. doi: 10.1093/humrep/des409 [doi].

102. Treloar AE, Boynton RE, Behn BG, Brown BW. Variation of the human menstrual cycle through reproductive life. *Int J Fertil*. 1967;12(1 Pt 2):77-126.

103. Depmann M, Faddy MJ, van der Schouw YT, et al. The relation between variation in size of the primordial follicle pool and age at natural menopause.

J Clin Endocrinol Metab. 2015:jc20151298. doi: 10.1210/jc.2015-1298 [doi].

104. Scheffer GJ, Broekmans FJ, Looman CW, et al. The number of antral follicles in normal women with proven fertility is the best reflection of reproductive age. *Hum Reprod*. 2003;18(4):700-706.

105. van Rooij IA, Tonkelaar I, Broekmans FJ, et al. Anti-mullerian hormone is a promising predictor for the occurrence of the menopausal transition. *Menopause*. 2004;11(6 Pt 1):601-606.

106. van Santbrink EJ, Hop WC, van Dessel TJ, de Jong FH, Fauser BC. Decremental follicle-stimulating hormone and dominant follicle development during the normal menstrual cycle. *Fertil Steril*. 1995;64(1):37-43.

107. Schipper I, de Jong FH, Fauser BC. Lack of correlation between maximum early follicular phase serum follicle stimulating hormone concentrations and menstrual cycle characteristics in women under the age of 35 years. *Hum Reprod*. 1998;13(6):1442-1448.

108. Hohmann FP, Laven JS, de Jong FH, Eijkemans MJ, Fauser BC. Low-dose exogenous FSH initiated during the early, mid or late follicular phase can induce multiple dominant follicle development. *Hum Reprod*. 2001;16(5):846-854.

109. de Vet A, Laven JS, de Jong FH, Themmen AP, Fauser BC. Antimullerian hormone serum levels: A putative marker for ovarian aging. *Fertil Steril*. 2002;77(2):357-362.

110. Lie Fong S, Visser JA, Welt CK, et al. Serum anti-mullerian hormone levels in healthy females: A nomogram ranging from infancy to adulthood. *J Clin Endocrinol Metab*. 2012;97(12):4650-4655. doi: 10.1210/jc.2012-1440 [doi].

111. Ramezani Tehrani F, Dolleman M, van Disseldorp J, et al. Predicting menopausal age with anti-mullerian hormone: A cross-validation study of two existing models. *Climacteric*. 2014. doi: 10.3109/13697137.2014.898264.

112. Bonifacio M, Bradley CK, Karia S, Livingstone M, Bowman MC, McArthur SJ. The original beckman coulter generation II assay significantly underestimates AMH levels compared with the revised protocol. *J Assist Reprod Genet*. 2015. doi: 10.1007/s10815-015-0579-y [doi].

113. Caserta D, Bordi G, Di Segni N, D'Ambrosio A, Mallozzi M, Moscarini M. The influence of cigarette smoking on a population of infertile men and women. *Arch Gynecol Obstet*. 2013;287(4):813-818. doi: 10.1007/s00404-012-2643-5 [doi].

114. Pelosi E, Simonsick E, Forabosco A, Garcia-Ortiz JE, Schlessinger D. Dynamics of the ovarian reserve and impact of genetic and epidemiological factors on age of menopause. *Biol Reprod*. 2015. doi: biolreprod.114.127381 [pii].

115. Depmann M, Broer SL, van der Schouw YT, et al. Can we predict age at natural menopause using ovarian reserve tests or mother's age at menopause? A systematic literature review. *Menopause*. 2015. doi: 10.1097/GME.0000000000000509 [doi].

116. Depmann M, Eijkemans MJ, Broer SL, et al. Does anti-mullerian hormone predict menopause in the general population? results of a prospective ongoing cohort study. *Hum Reprod*. 2016. doi: dew112 [pii].

117. Sowers MR, McConnell D, Yosef M, Jannausch ML, Harlow SD, Randolph Jr. JF. Relating smoking, obesity, insulin resistance, and ovarian biomarker changes to the final menstrual period. *Ann New York Acad Sci*. 2010;1204:95-103.

118. Dolleman M, Verschuren WMM, Eijkemans MJC, Broekmans FJM, Van Der Schouw YT. Anti-müllerian hormone in prediction of menopause: Results from a large prospective cohort study. *Hum Reprod*. 2014;29:i13-i14.

119. 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;98(2):729-735. doi: 10.1210/jc.2012-3176 [doi].

120. Ramezani Tehrani F, Mansournia MA, Solaymani-Dodaran M, Steyerberg E, Azizi F. Flexible parametric survival models built on age-specific antimullerian hormone percentiles are better predictors of menopause. *Menopause*. 2016. doi: 10.1097/GME.0000000000000599 [doi].

121. van Montfrans JM, van Hooff MH, Huirne JA, et al. Basal FSH concentrations as a marker of ovarian ageing are not related to pregnancy outcome in a general population of women over 30 years. *Hum Reprod*.

2004;19(2):430-434.

122. Steiner AZ, Herring AH, Kesner JS, et al. Antimullerian hormone as a predictor of natural fecundability in women aged 30-42 years. *Obstet Gynecol*. 2011;117(4):798-804. doi: 10.1097/AOG.0b013e3182116bc8 [doi].

123. Hagen CP, Vestergaard S, Juul A, et al. Low concentration of circulating antimullerian hormone is not predictive of reduced fecundability in young healthy women: A prospective cohort study. *Fertil Steril*. 2012;98(6):1602-8. e2. doi: 10.1016/j.fertnstert.2012.08.008 [doi].

124. Streuli I, de Mouzon J, Paccolat C, et al. AMH concentration is not related to effective time to pregnancy in women who conceive naturally. *Reprod Biomed Online*. 2014;28(2):216-224. doi: 10.1016/j.rbmo.2013.10.007 [doi].

125. Broer SL, Dolleman M, Opmeer BC, Fauser BC, Mol BW, Broekmans FJ. AMH and AFC as predictors of excessive response in controlled ovarian hyperstimulation: A meta-analysis. *Hum Reprod Update*. 2011;17(1):46-54. doi: 10.1093/humupd/dmq034 [doi].

126. Broer SL, Mol BW, Hendriks D, Broekmans FJ. The role of antimullerian hormone in prediction of outcome after IVF: Comparison with the antral follicle count. *Fertil Steril*. 2009;91(3):705-714. doi: 10.1016/j.fertnstert.2007.12.013 [doi].

127. Sopher AB, Grigoriev G, Laura D, et al. Anti-mullerian hormone may be a useful adjunct in the diagnosis of polycystic ovary syndrome in nonobese adolescents. *J Pediatr Endocrinol Metab*. 2014. doi: 10.1515/jpem-2014-0128 [doi].

128. Zarek SM, Mitchell EM, Sjaarda LA, et al. Is anti-mullerian hormone associated with fecundability? findings from the EAGeR trial. *J Clin Endocrinol Metab*. 2015;100(11):4215-4221. doi: 10.1210/jc.2015-2474 [doi].

129. Daan NM, Louwers YV, Koster MP, et al. Cardiovascular and metabolic profiles amongst different polycystic ovary syndrome phenotypes: Who is really at risk? *Fertil Steril*. 2014;102(5):1444-1451.e3. doi: 10.1016/j.fertnstert.2014.08.001 [doi].

130. Bonde JP, Ernst E, Jensen TK, et al. Relation between semen quality and

fertility: A population-based study of 430 first-pregnancy planners. *Lancet*. 1998;352(9135):1172-1177. doi: S0140-6736(97)10514-1 [pii].

131. Tehrani FR, Mansournia MA, Solaymani-Dodaran M, Azizi F. Age-specific serum anti-mullerian hormone levels: Estimates from a large population-based sample. *Climacteric*. 2014. doi: 10.3109/13697137.2014.912262 [doi].

132. van den Boogaard NM, Kersten FA, Goddijn M, et al. Improving the implementation of tailored expectant management in subfertile couples: Protocol for a cluster randomized trial. *Implement Sci*. 2013;8:53-5908-8-53. doi: 10.1186/1748-5908-8-53 [doi].

133. Gnoth C, Godehardt D, Godehardt E, Frank-Herrmann P, Freundl G. Time to pregnancy: Results of the german prospective study and impact on the management of infertility. *Hum Reprod*. 2003;18(9):1959-1966.

134. Battaglia DE, Goodwin P, Klein NA, Soules MR. Influence of maternal age on meiotic spindle assembly in oocytes from naturally cycling women. *Hum Reprod*. 1996;11(10):2217-2222.

List of publications

Does anti-Müllerian hormone predict menopause in the general population? Results of a prospective ongoing cohort study.
Depmann M, Eijkemans MJ, Broer SL, Scheffer GJ, van Rooij IA, Laven JS, Broekmans FJ.; Hum Reprod. 2016 May 13. pii: dew112. [Epub ahead of print]

Fluctuations in anti-Müllerian hormone levels throughout the menstrual cycle parallel fluctuations in the antral follicle count: a cohort study.
Depmann M, van Disseldorp J, Broer SL, Eijkemans MJ, Laven JS, Visser JA, de Rijke YB, Mol BW, Broekmans FJ.;Acta Obstet Gynecol Scand. 2016 Feb 26. doi: 10.1111/aogs.12886. [Epub ahead of print]

Can we predict age at natural menopause using ovarian reserve tests or mother's age at menopause? A systematic literature review.
Depmann M, Broer SL, van der Schouw YT, Tehrani FR, Eijkemans MJ, Mol BW, Broekmans FJ,; Menopause. 2016 Feb;23(2):224-32.

The Relationship Between Variation in Size of the Primordial Follicle Pool and Age at Natural Menopause.
Depmann M, Faddy MJ, van der Schouw YT, Peeters PH, Broer SL, Kelsey TW, Nelson SM, Broekmans FJ.; J Clin Endocrinol Metab. 2015 Jun;100(6):E845-51. doi: 10.1210/jc.2015-1298. Epub 2015 Apr 27.

Anti-Müllerian Hormone is a More Accurate Predictor of Individual Age at Menopause Than Mother's Age at Menopause.
M. Dólleman, **M. Depmann**, M.J.C. Eijkemans, J. Heimensem, S.L. Broer, E.M.van der Stroom, J.S.E.Laven, I.A.J. Van Rooij, G.J. Scheffer, P.H.M. Peeters, Y.T. van der Schouw, C.B. Lambalk, F.J.M. Broekmans; Hum Reprod. 2014 Mar;29(3):584-91. doi: 10.1093/humrep/det446. Epub 2014 Jan 15

Counseling bij het structureel echoscopisch onderzoek. Is groepscounseling een alternatief voor individuele counseling?
H. de Lau, **M. Depmann**, Y.J.M. Laeven, P.H. Stoutenbeek, L.R. Pistorius, E. van Beek, N. W.E. Schuitemaker
Ned Tijdschr Geneeskd. 2013;157(23):A5935. Dutch.

About the author

Martine Depmann was born in Tilburg, the Netherlands, on September 27th in 1984. There, she grew up with her older sister Maaike and younger brother Lars. After graduating from high school in 2003, she started medical school at the University of Utrecht. Obstetrics and Gynaecology caught her special interest during an internship in the Diakonessenhuis in Utrecht. Therefore, after obtaining her degree in 2011, she started working in this hospital as a resident (not in training) and fertility doctor at the department of Obstetrics and Gynaecology.

Between 2012 and 2014 she started her scientific career by combining her clinical work at the Diakonessenhuis with data acquisition of an ongoing follow up study following the work of dr. S.L. Broer in the University Medical Center Utrecht.

In 2015, after the acquisition of data was completed, she continued this research project as a PhD student at the department of 'Woman and Baby' under the supervision of prof. dr. Broekmans and prof. dr. B.W. Mol.

In July 2015 she started her residency Obstetrics and Gynaecology at the Gelre Hospital in Apeldoorn (supervisor dr K.M. Paarlberg).

Dankwoord

Daar gaan we dan, het dankwoord en dat betekent dat het officieel is: het boek is af!

Dit proefschrift zou nooit tot stand gekomen zijn zonder de hulp van een groot aantal mensen. Een aantal van jullie wil ik in het bijzonder bedanken.

Allereerst wil ik de vrouwen bedanken die belangeloos hebben deelgenomen aan de studies in dit proefschrift. Zonder jullie zou dit boek nooit tot stand gekomen zijn.

Geachte prof. dr. Broekmans, beste Frank. Wat begon als een parttime onderzoeksproject is geëindigd in het boek wat je nu in je handen hebt. Ik wil je heel erg bedanken voor de enorme kans die je mij hebt gegeven. Je was als promotor scherp, flamboyant, gaf sturing en stelde kritische vragen. Maar buiten dat hebben jouw betrokkenheid en jouw adviezen, zowel op werkgebied als privé, mij enorm geraakt.

Geachte prof. dr. Mol, beste Ben Willem. Een promotor op afstand, maar daar viel in de snelheid van reactie en in jouw mate van betrokkenheid bij de projecten niets van te merken. Dank voor jouw soms kritische en altijd scherpe blik.

Dr. S.L. Broer, beste Simone, ik ben er vreselijk trots op dat ik jouw eerste promovenda mag zijn. Ik mocht in jouw voetsporen treden en het was een feestje om jou als co-promotor te hebben. Ondanks drukke agenda's, het soms wonen op een ander continent en wisselende diensten vond jij altijd een moment om met me te sparren als ik vastliep of een in mijn ogen vastgelopen stuk met een scherpe blik vlot te trekken. Dank je wel voor alles.

Geachte leden van de beoordelingscommissie, prof dr. R.H.M. Verheijen, prof. dr. W.M. Verschuren, prof. dr. D.M. Braat, prof. dr. C.M. Lambalk, prof. dr. N. Wulffraat. Hartelijk dank voor de tijd en moeite die u heeft gestoken in het beoordelen van mijn manuscript.

Dr. I.A. van Rooij, beste Ilse, dr. G.J. Scheffer, beste Gabrielle en dr. A. de Vet, beste Annemarie. Het is jullie basis waarop ik heb mogen voortborduren, dank jullie wel hiervoor.

Beste co-auteurs, dank voor jullie betrokkenheid en jullie bijdrage aan dit proefschrift. Ik wil twee co-auteurs in het bijzonder bedanken:
Prof. dr. Eijkemans, beste René, zonder jou was dit boek er nooit geweest. Je maakte altijd tijd voor me vrij om na te denken over analyse stappen, me bij te staan tijdens analyses en me te behoeden voor mijn fouten. Daarnaast maakte je ook altijd tijd vrij voor een persoonlijk gesprekje. Dank je wel voor je begeleiding.
Dear Malcolm Faddy, your contribution to 'our paper', as we would like to call it, was not only most valuable but also great fun. Thank you for taking time out of retirement for our project. I am looking forward to showing you all the places we discussed in our emails next time you come to visit the Netherlands.

Lieve Ellis, Ingrid, Tessa en Marieke, dank voor alles!

Beste gynaecologen, A(N)IOS, verloskundigen, verpleegkundigen, poli-assistenten en fertiliteitsverpleegkundigen van het Diakonessenhuis Utrecht. Het was tijdens mijn co-schap bij jullie dat een bijzondere interesse voor dit vak ontstond, maar het was mijn AGNIO/fertiliteit tijd die dit bevestigde. Dank jullie wel voor drie fantastische jaren. De maatschap heeft mij belangeloos onderzoekstijd gegeven en de stappen die ik daarin heb kunnen maken hebben geleid tot dit proefschrift. Ik ben jullie enorm dankbaar voor dat vertrouwen en die kans.

Lieve meiden van kamertje 1, Kèm, de Kat, Daan en lieve boys, Ger en Tobs. Jullie hebben mij door zoveel meer heen gesleept dan alleen deze promotie. Dit was me nooit gelukt zonder jullie. Onze trip naar Kopenhagen was er één om nooit te vergeten, fijn dat de volgende alweer gepland staat.
Vrienden voor het leven!

Lieve onderzoekers op andere kamers maar daarom niet minder dierbaar, Japie, Marlieke (gelukkig werken nu weer elke dag samen!), Smitje, Charine, Kristine, Wendy en natuurlijk 'de overkant'. Dank voor alle fijne gesprekken, de koffie momenten, de overwinningskroketten en natuurlijk de maandaglunch en vrijdagmiddagborrels.

'Nieuwe lichting', Marlise, Chris, Nienke, Simone, Laura, Hans, Leon en Fieke. We hebben maar kort samengewerkt, maar wat een leuke groep is er weer! Ik geef het stokje graag aan jullie door, ik hoop dat jullie er minstens zo van zullen genieten.

Lieve gynaecologen, arts- assistenten, verloskundigen, verpleegkundigen en poli- assistentes in het Gelre ziekenhuis, bij jullie mocht ik mijn eerste stappen als AIOS zetten. Dank voor het warme welkom en de fijne en enorm leerzame tijd.

Lieve Gelre carpool collega's, met jullie vliegen de kilometers voorbij. Ik stap elke dag met plezier uit op de meest troosteloze carpoolstrook van Nederland.

Vrienden, clubgenoten en oud-huisgenoten dank jullie wel voor jullie vriendschap.

Lieve El en Wil, dank voor onze heerlijke weekendjes en vakanties. Dat er nog vele mogen volgen!

Broer en zus, mijn paranimfen en zo hoort het, ook al hebben jullie geen idee waar jullie ja tegen hebben gezegd. Samen groot geworden en nu staan jullie me bij tijdens mijn verdediging. Ook nu we allemaal onze eigen weg gaan, weten we elkaar te vinden. Jullie zijn de beste broer en zus die ik me had kunnen wensen.

Lieve mama & papa, dank jullie wel voor jullie altijd aanwezige betrokkenheid bij wat we ook ondernemen. Jullie hebben me geleerd door te zetten als dingen tegen zitten en te vertrouwen op mijn eigen oordeel. Ik leerde van jullie dat de hele wereld voor me open stond en ik alles kon worden wat ik wilde, maar dat er soms voor gewerkt moet worden. Ik vind het ongelooflijk bijzonder te zien hoe gelukkig jullie elkaar na al die jaren nog steeds maken, het is een prachtig voorbeeld.

Lieve Jochem, wat ben ik blij dat ik jou (weer of nu echt?) heb leren kennen. Naast alle leuke momenten en slechte grappen breng je me rust. Bij jou zijn voelt als thuis komen. Ik kan niet wachten op de rest van ons avontuur.









En dat... is alles wat ik daarover te zeggen heb