Testing for Ovarian Reserve in Assisted Reproduction programs: the current point of view

Frank J. Broekmans

Department of Reproductive Medicine and Gynecology, Division Mother and Child, University Medical Center Utrecht, Utrecht, The Netherlands.

Correspondence at: Heidelberglaan 100, 3584 CX Utrecht, The Netherlands. E-mail: f.broekmans@umcutrecht.nl

Abstract

Age related fertility decline varies considerably among women. Therefore, chronological female age, though inform-ative on pregnancy prospects in assisted reproduction, will often not correctly express a woman's reproductive potential. The value of quantitative ovarian reserve tests prior to IVF/ICSI treatment is still subject of debate. From a series of systematic reviews it has become clear that the added value of these tests upon knowing female age has not been clearly established. Still, several tests, like the AFC and AMH are considered adequate in predicting the response to ovarian stimulation. This claim seems to be truer for poor response prediction, compared to hyper response. Prediction of the outcome pregnancy has repeatedly shown to be cumbersome. As management options for predicted poor or hyper responders are not fully investigated to date, routine ovarian reserve testing is not to be recommended. A first cycle poor response to adequate stimulation in cases with otherwise no signs of advanced ovarian ageing (based on female age and ovarian reserve tests) may offer a tool to identify cases with sufficient prospects for continuation of ART treatment.

Key words: Assisted reproduction, IVF, ovarian ageing, ovarian reserve.

Ovarian ageing

The human species can be considered as relatively infertile (Viudes-de-Castro and Vicente, 1997; Moce *et al*., 2005). The average monthly fecundity rate of about 20% implies that among human couples trying to conceive many exposure months may be needed to achieve their goal (Evers, 2002). It has also been long known that with increasing chronologic age, female fecundity – the ability to produce offspring – decreases. This knowledge is based on studies in-volving both natural historical (Spira, 1988; Wood, 1989) and contemporary populations (Abma *et al*., 1997), as well as on studies of age dependent success rates in Assisted Reproduction Technology (ART) (Nyboe *et al*., 2009; Templeton *et al*., 1996; Centers for Disease Control and Prevention, 2007).

The age related female infertility (Stephen and Chandra, 2006; Noord-Zaadstra *et al*., 1991) is mainly based on changes in ovarian reserve. Ovarian reserve can be defined as the number and quality of the remaining follicles and oocytes in both ovaries at a given age. Decline in follicle numbers dictates the occurrence of irregular cycles and menopause, while quality decay of the oocytes results in decreas-ing fertility, defined as the capacity to conceive and give birth to a child (te Velde and Pearson, 2002) (Figure 1).

There is substantial individual variation in the onset of menopause, varying roughly between 40 and 60 years, with a mean age of 51 (Morabia and Costanza, 1998; Thomas *et al*., 2001). Along the same pattern, the rate of decline in fertility may vary considerably between women of the same age. This implies that a woman at the age of 35 either may be close to natural sterility or have a fertility compara-ble to a 25 year old woman (Broekmans *et al*., 2004; te Velde and Pearson, 2002; Eijkemans *et al*., 2005) (Figure 2).

The insights into the process of ovarian ageing imply that for ovarian reserve testing prior to ART, female age remains the predictor of first choice. The availability of a test capable of providing reliable in-formation regarding a woman's individual ovarian
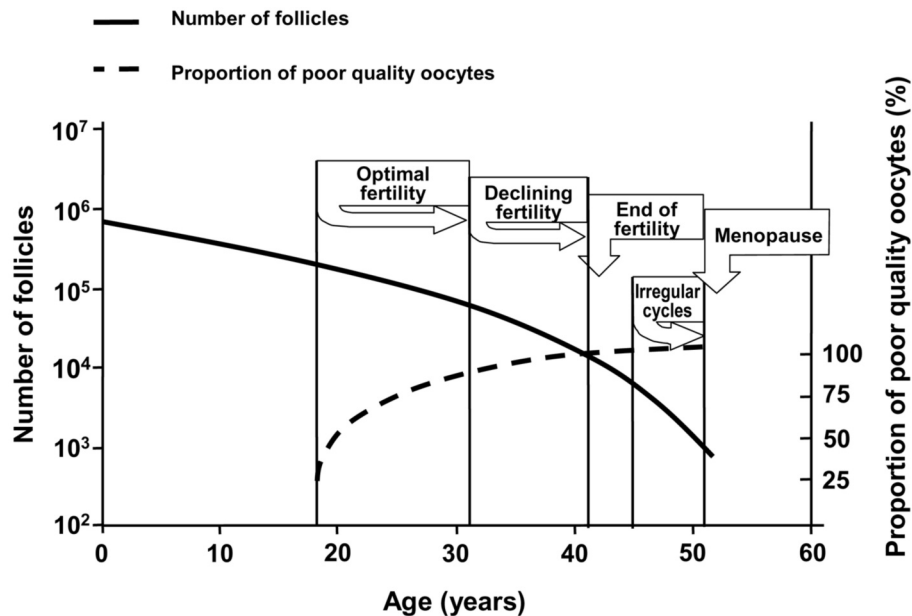
Fig. 1. — Quantitative (solid line) and qualitative (dotted line) decline of the ovarian follicle pool, which is assumed to dictate the onset of the important reproductive events (Graph was drawn after Hansen and de Bruin (Hansen *et al*., 2008; de Bruin *et al*., 2001).
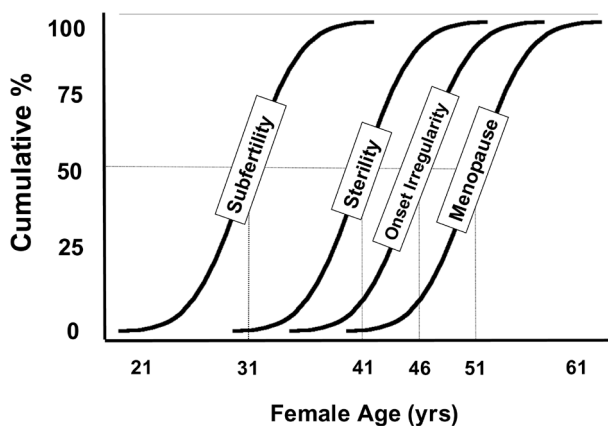

Fig. 2. — Variations in age at the occurrence of specific stages of ovarian ageing. For explanation of the background of data, see te Velde and Pearson (te Velde and Pearson, 2002). Adapted from te Velde and Pearson (2002).

reserve within a certain age category would enable the clinician to provide an individually tailored treatment plan. For instance, in older women the finding of a high ovarian reserve may justify the decision to allow ART treatment, while in young women with exhausted reserve either early application, refusal of ART or choosing for egg donation could be the consequence.

Ovarian reserve testing prior to IVF

The first notice on ovarian reserve assessment prior to starting IVF was published in 1988. FSH levels in the early follicular phase appeared related to stimu-

lation response and outcome of IVF treatment (Muasher *et al*., 1988). Soon thereafter, the predictive role for basal FSH in IVF treatment was further confirmed by Scott *et al*. (Scott *et al*., 1989), who stated that cycle day three FSH levels predicted pregnancy outcome and stimulation characteristics in IVF, and might be useful in counselling patients. In the two decades thereafter, a large body of additional work was published, showing that several other cycle day 3 parameters, such as Inhibin B, the antral follicle count (AFC) and antimullerian hormone (AMH) were capable of predicting ovarian responsiveness, and, to a much lesser extent, the outcome of IVF in terms of pregnancy (Seifer *et al*., 1997; Seifer *et al*., 2002; Tomas *et al*., 1997). Tests that challenged the cohort of FSH sensitive follicles in various ways had equal predictive capacity to predict response and outcome compared to basal tests, and thus failed to achieve wide application (Loumaye *et al*., 1990; Winslow *et al*., 1991; Fanchin *et al*., 1994; Hendriks *et al*., 2005a; Hendriks *et al*., 2005b; Hendriks *et al*., 2006).

Ovarian reserve can be considered normal in conditions where stimulation by exogenous gonadotropins results in the retrieval of some 6-14 healthy oocytes at follicle puncture (Fasouliotis *et al*., 2000; Popovic-Todorovic *et al*., 2003b; Nargund *et al*., 2007). With such a yield the chances of producing a live birth through IVF are considered optimal (van der Gaast *et al*., 2006). In addition to the number of recruitable follicles (a reflection of the ovarian reserve status), follicle sensitivity to FSH as well as the pharmacokinetics of FSH determine a woman's

ovarian response to stimulation (Karlsson *et al.*, 1998; Karlsson *et al.*, 1997). The dose of FSH used may therefore be a factor of importance, although the therapeutic range of this compound seems quite narrow. Higher doses of FSH may lead to higher numbers of oocytes retrieved in younger patients (Out *et al.*, 2001), but not in all studies (Harrison *et al.*, 2001). Such an approach will certainly fail in older women (Yong *et al.*, 2003) or in women expected to have a poor response to stimulation based on an abnormal ovarian reserve test (Klinkert, 2005). With currently applied dose levels of exogenous FSH (150-450 IU) stimulation of the ovaries will be maximal in virtually 100% of cases.

The preferred outcome of OR test prediction studies would be live birth after a series of ART cycles in order to express of a couple's fertility potential. Other outcome measures (especially oocyte yield or follicle number and pregnancy after one IVF/ICSI cycle) are in fact the most common. However, ovarian reserve tests mainly relate to the size of the follicle cohort that is at any time responsive to FSH. The antral follicle count (AFC) assessed by transvaginal ultrasonography provides direct visual assessment of the cohort (Hendriks *et al.*, 2005), while the endocrine markers anti mullerian hormone (AMH) and inhibin B are released products from the antral follicle pool (Broekmans *et al.*, 2006; Seifer *et al.*, 1997; Seifer, MacLaughlin *et al.*, 2002). Basal FSH provides a more indirect marker, as it reflects a reduced feedback from the antral follicle pool as it becomes smaller in size. It goes without saying that the focus on quantity prohibits high expectations on the relation to oocyte quality and pregnancy as outcome.

Ovarian reserve test evaluation should imply the assessment of predictive accuracy and clinical value of the test. Predictive accuracy refers to the degree by which the outcome condition (pregnancy or poor response) is predicted correctly. Summary statistics of accuracy include sensitivity (rate of correct identification of cases with e.g. poor response) and specificity (rate of correct identification of cases without poor response) (Deeks, 2001; Grimes and Schulz, 2005). Using the sensitivity and specificity for a range of cut off levels a Receiver Operating Characteristic curve can be drawn and area under this curve calculated to represent the overall predictive accuracy of the test. Values of 1.0 imply perfect and 0.5 indicate completely absent discrimination.

Assessment of the clinical value is a complex process through which the applicability in daily practice should become clear. The overall accuracy represented by the ROC curve, the choice of a cut off for abnormality, the rate of abnormal tests at that cut off, the post test probability of disease (i.e., poor response or non-pregnancy), the valuation of false positive and false negative test results and the consequence for patient management of an abnormal test will all contribute to the process of deciding whether a test is useful or not. An overall estimate of test quality is the positive likelihood ratio, which describes the chance of an abnormal test over a normal test in the case of non pregnancy or poor response. The cost of carrying out the test as a routine measure and the burden to the patient, balanced against the reduction in costs by excluding cases with low pregnancy prospects need also to be incorporated in the decision process. Finally, clinical value may also be influenced by valuation from patients and health insurance preference regarding the consequences that should be drawn from abnormal tests (Mol *et al.*, 2006).

One aspect of clinical value deserves special attention. Ovarian reserve tests are mostly used as a diagnostic test, indicating that in case of an abnormal test result the diagnosis diminished ovarian reserve is made (Levi *et al.*, 2001; Scott Jr. and Hofmann, 1995). In fact, ovarian reserve tests may better be considered as screening tests, where an abnormal test necessitates confirmation by another test. This other test may for instance be a first IVF attempt where ovarian response to maximal stimulation is the additional test. Alternatively, combinations of independently predictive tests or repeating of the initial test could improve the diagnostic performance of the single test (Bancsi *et al.*, 2002; Bancsi *et al.*, 2004a; Bancsi *et al.*, 2004b; van Rooij *et al.*, 2002; Ng *et al.*, 2000; Popovic-Todorovic *et al.*, 2003b).

In several systematic reviews in the last decade the true value of ovarian reserve tests for clinical practice has been debated (Bancsi *et al.*, 2003; Broekmans *et al.*, 2006; Broer *et al.*, 2009; Hendriks *et al.*, 2007; Verhagen *et al.*, 2008). Especially, the limited capacity of OR tests to discriminate between pregnant and non pregnant women and the lack of knowledge on the added value of OR tests upon knowing the female's age have been reason not to advocate OR testing as routine test prior to IVF.

Ovarian reserve research has mainly focussed on the explanation and prediction of poor responses and low pregnancy outcome in assisted reproduction technology (ART) treatment. A hyper response to ovarian stimulation, however, also represents an continuous challenge for the clinician. Cycle cancellation at any stage in hyperresponders is often necessary to eliminate the risk of developing the ovarian hyperstimulation syndrome (OHSS), a potentially life threatening condition. Also, the interest in milder stimulation protocols, that lead to lower

Accuracy Poor Response prediction
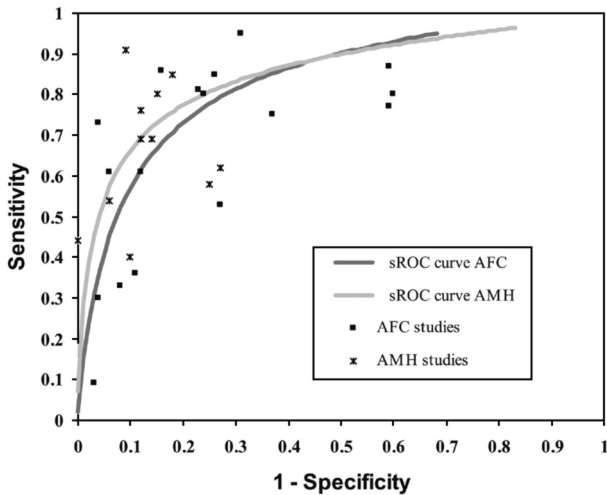
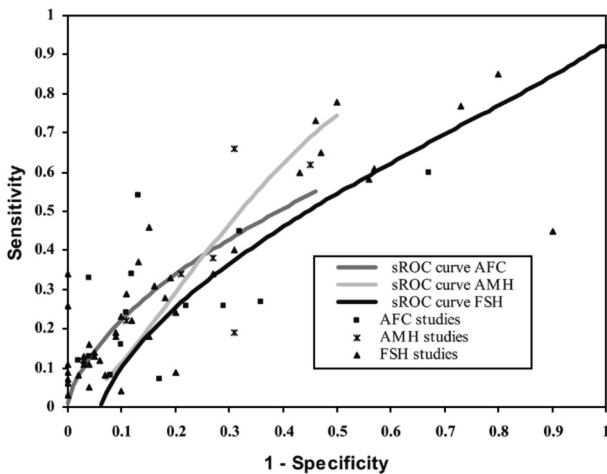## Accuracy Non Pregnancy prediction

Fig. 3. — Example of ovarian reserve test performance (AFC, AMH and FSH) showing receiver operator characteristic (ROC) curves for the prediction of poor response (upper panel) and non pregnancy (lower panel) in IVF. Data were based on a series of meta-analyses on ovarian reserve tests (Broekmans *et al*., 2006; Broer *et al*., 2009).

costs, patient burden and complications (Heijnen *et al*., 2007; Polinder *et al*., 2007), urges for the availability of reliable markers of hyper response. Finally, hyper response to ovarian stimulation is more and more considered as a condition in which low quality or immature oocyes are added to a basal number of best quality oocytes (Kok *et al*., 2006).

Factors that are classically associated with hyper-response and OHSS are lean habitus, young age, presence of multiple antral follicles, and the presence of the polycystic ovary syndrome (PCOS). The prediction of hyper response from prior tests like basal FSH, AMH or the AFC has shown to be quite inaccurate or at least inconsistent (Seifer *et al*., 2002; van Rooij *et al*., 2002; Nelson *et al*., 2007; Ng *et al*.,

2000; Lee *et al*., 2008; Tremellen *et al*., 2005). Currently, no definite strategies on management in case of a predicted hyper response based on such prior test are known, although FSH dose reduction would be the logical step (Olivennes *et al*., 2009). Prevention of hyper response therefore is based on patient profiles, like very young age and the presence of the PCO Syndrome, as well as the general use of modest dosing schemes not exceeding 200 IU in first cycles.

ORT accuracy and clinical value

The findings in a series of systematic reviews of the existing literature (Bancsi *et al*., 2003; Broekmans *et al*., 2006; Broer *et al*., 2008; Hendriks *et al*., 2005; Hendriks *et al*., 2007; Verhagen *et al*., 2008) have demonstrated that several tests have good capacity to predict poor responders to ovarian hyper stimulation for IVF. The areas under the receiver operator characteristic curve (AUC-ROC) for baseline FSH, and especially the AFC and AMH, indicate that the overall accuracy is sufficient (Figure 3, (AUC-ROC: > 0.70).

For instance, if the prevalence of poor response was set at 20% and the cut off chosen at a positive likelihood ratio (ie. the chance of an abnormal test over a normal test in poor responders) level of at least 6 (indicating an overall good test), an abnormal AFC would indicate a post-test probability of poor ovarian response of around 67%. This would make the AFC test a clinically valuable test, especially as an abnormal test result would be found in 12% of patients. The same has shown to be true for AMH, where comparable levels for post test probability and abnormal test rate were observed as for the AFC. The choice for either of these two tests is mainly directed by practical issues, like availability and stability of the AMH assay and the possibilities for standardised use of ultrasound based follicle counting (Broekmans *et al*., 2009). In contrast, for FSH, a positive LR of 6 and over would imply a post-test likelihood of poor response of about 67%, but at such high cut off levels that abnormal tests would occur in only 3% of patients.

From the reviews, the predictive ability towards pregnancy after one IVF cycle appeared only marginal for all the tests, as the area under the ROC curve remained very close to the non discriminative value of 0.50. Only with extreme cut offs for an abnormal test some non pregnant cases were predicted correctly, without too many false positives. At such cut off levels the rate of abnormal tests is very low (≤ 2%).

Recent literature has focused on the added value of ovarian reserve tests to the information of female

age, although reports on the univariate prediction from OR tests seem quite hard to eliminate (Maseelall *et al.*, 2009). The relation between age and live birth in ART programs is strong, although it remains difficult to decide at which age level the prospects for pregnancy have become poor enough to advice against or refuse treatment. Adding information from OR tests would help the chance prediction to be individualised. Recent work by Scott (Scott Jr *et al.*, 2008) has attempted to define age specific cut off levels for basal FSH to predict failure to achieve live birth. It appeared that useful cut offs per age class could only be identified at threshold values where the live birth rates were under 2%. At such, high (15-18 IU/l), cut off levels the percentage of abnormal test results appeared to be only 1.6%. Also, in the range of FSH results under 12 IU/l, the added predictive value upon female age has demonstrated to be only marginal, so that lower than extreme cut offs will make the test useless, in spite of more abnormal tests obtained (Henne *et al.*, 2008; Sun *et al.*, 2008).

In general, therefore, ovarian reserve testing prior to starting ART treatment should be regarded useful only if the occurrence of poor response to ovarian stimulation is to be predicted and with the assumption that this foreseen poor response can be effectively prevented with improvement of pregnancy chances (Nelson *et al.*, 2009). However, even a normalized response to ovarian stimulation may not alter the prognosis regarding the chances of pregnancy (Land *et al.*, 1996). Several studies have shown that in observed poor responders in a first IVF cycle no clear benefit can be expected from various changes in management like increasing the dosage (Hoveyda *et al.*, 2002), applying co-medication, or changing the approach of the GnRH agonist administration (Tarlatzis *et al.*, 2003; Klinkert, 2005; Shanbhag *et al.*, 2007; Kyrou *et al.*, 2009; Kolibianakis *et al.*, 2009; Mochtar *et al.*, 2007). This implies that a prior prediction of poor response is to be considered useless, unless this prediction would identify cases with a poor response due to FSH under dosing related to obesity or FSH receptor polymorphisms.

In cases without signs of ovarian ageing the use of a prediction model for ovarian response to FSH, containing the AFC, ovarian volume, power Doppler score, female age and smoking habit, was developed for individualization of the FSH dose from the first cycle onwards (Popovic-Todorovic *et al.*, 2003b). To test whether this FSH dosage score performs well in predicting ovarian response, a randomized trial compared ovarian response in women assigned either to an individual dose of FSH based on her score, or a 'standard' dose of 150 IU/day (Popovic-Todorovic *et al.*, 2003a; Popovic-Todorovic *et al.*, 2004). Women in the individual dose group had a higher proportion of appropriate ovarian response than women in the standard dose group. Even ongoing pregnancy rates were higher in the individualized compared to the standard dose group and dose adjustments were less frequently necessary than in the standard dose group. Whether the increased occurrence of pregnancies had been obtained from dose reduced or dose increased (predicted poor responder) cases was not made clear. Further research on the issue of patient tailored dosing and its possible beneficial effects upon pregnancy rates needs to be awaited to see whether poor responders due to other factors than ovarian ageing indeed will benefit from adapted treatment schedules (Olivennes *et al.*, 2009).

First cycle poor response

Testing for ovarian reserve may also be possible by using the quantity of the ovarian response to maximal ovarian stimulation in the first ART cycle. The assumption would be that the antral follicles visible at transvaginal ultrasound will all grow into dominance with the use of dosages of exogenous FSH of 150 IU daily or over. Support for this comes from studies where the number of oocytes appeared correlated to the number of antral follicles (Hansen *et al.*, 2003; Hsieh *et al.*, 2001; Lorusso *et al.*, 2007). A poor response to stimulation, defined as a low number of mature follicles developed or oocytes obtained after a conventional long GnRH agonist suppression protocol, will generally be interpreted as a proof of diminished ovarian reserve and reduced prognosis for pregnancy. Cycle cancellation to a standard IVF stimulation will predict a poor response in a subsequent cycle more accurately than classical ovarian reserve tests (Penarrubia *et al.*, 2005). Also, poor responders experience an earlier transition into menopause compared to normal responders, confirming the relation between response and subsequent fertility potential.

Still, a poor response may also be caused by conditions like sub maximal stimulation in obese women, the presence of a FSH receptor polymorphism or simply by chance. In such poor responders, the prospects in the actual and subsequent cycles are not so unfavourable that refusal of treatment is justified (Popovic-Todorovic *et al.*, 2003a; Popovic-Todorovic *et al.*, 2003b). The same seems to be true for poor responders of younger age as has been shown from several studies (Lashen *et al.*, 1999). Only if a poor response occurs in cases with an unfavourable additional profile (female age over 38, abnormal ovarian reserve test, repeated poor response) does prognosis for subsequent cycles becomes cumbersome enough for further denial of treatment (Vladimirov *et al.*, 2005; Baka *et al.*, 2006; Zhen *et*

Probability Ongoing Pregnancy in IVF cycle 1, 2 or 3
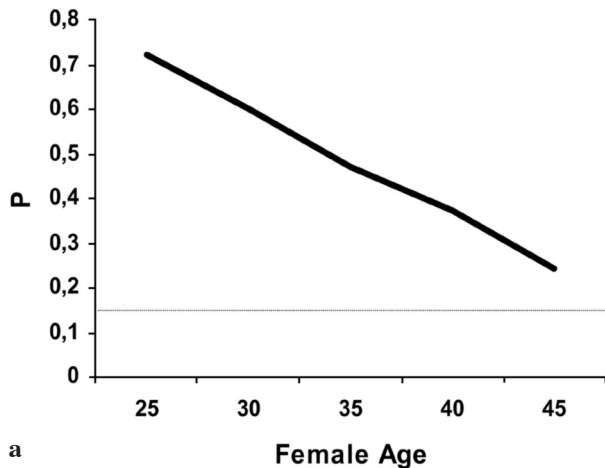
a

Probability Ongoing Pregnancy in cycle 2 or 3
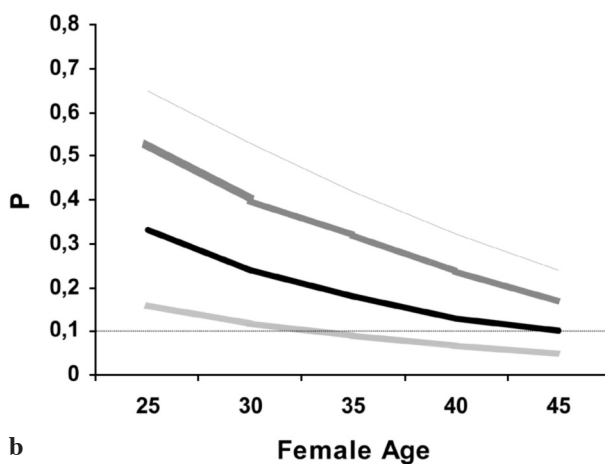
b

Fig. 4. — Predicted cumulative ongoing pregnancy rates for several conditions.
(a) Predicted probabilities using female age as predictor. Taking a 15% cumulative pregnancy rate in three cycles as minimal level of success, such a poor prospect cannot be predicted by age for any single woman.
(b) Predicted probabilities using female age, poor response, and poor response type as predictors. Taking a 10% cumulative pregnancy rate in cycles 2 and 3 as minimum success level, only in cycle 1 poor responders, who were expected based on the ovarian reserve test combination (FSH, antral follicle count, inhibin B), can such poor prognosis situations be identified.
PR = cycle 1 poor responder; NR = cycle 1 normal responder; Exp PR = cycle 1 poor responder with abnormal ovarian reserve testing; Unexp PR = cycle 1 poor responder with normal ovarian reserve testing.
Reprint from Hendriks *et al*. (Hendriks *et al*., 2008).

al., 2008; Lawson *et al*., 2003; de Boer *et al*., 2003; Klinkert *et al*., 2004).

The best policy for IVF cases therefore could be the unrestricted entry into IVF treatment, without prior testing. The occurrence of a poor response to

stimulation would then urge for a further classification: can the response be classified as expected or unexpected in view of female age or the result of an ovarian reserve test. Expected poor responders could then be counselled for further refraining from treatment and egg donation, unexpected poor responders may still have reasonable prospects in subsequent cycles and benefit from the use of a higher FSH stimulation dosage (Popovic-Todorovic *et al*., 2003a; Popovic-Todorovic *et al*., 2003b) (Figure 4).

The true challenge for ovarian reserve tests lies in the possibility of identifying women with a reduced reproductive lifespan at such a stage in their lives that adequate action can be taken. In such test the preferable outcome variable to judge the test upon is the age at which a woman will become menopausal. The relation between menopausal age and the end of natural fertility has been hypothesized to be fixed (te Velde and Pearson, 2002). If a test existed that adequately predicts age at menopause, then adequate prevention of at least age related infertility would become reality.

Summary

From the overview on ovarian reserve testing provided, two main points of attention can be deduced. First, the routine use of ORTs prior to starting ART can not be justified, as clear therapeutic options in cases with anticipated low response are lacking. Second, a first IVF attempt poor responder without signs of advanced ovarian ageing does not bear a poor prognosis and may benefit from adapted treatment schedules in subsequent cycles.

References

Abma JC, Chandra A, Mosher WD, Peterson LS, Piccinino LJ. Fertility, family planning, and women's health: new data from the 1995 National Survey of Family Growth. Vital Health Stat 1997;23:1-114.

Baka S, Makrakis E, Tzanakaki D, Konidaris S, Hassiakos D, Moustakarias T, Creatsas G. Poor responders in IVF: cancellation of a first cycle is not predictive of a subsequent failure. Ann N Y Acad Sci 2006;1092:418-425.

Bancsi LF, Broekmans FJ, Eijkemans MJ, de Jong FH, Habbema JD, te Velde ER. Predictors of poor ovarian response in in vitro fertilization: a prospective study comparing basal markers of ovarian reserve. Fertil Steril 2002;77:328-336.

Bancsi LF, Broekmans FJ, Looman CW, Habbema JD, te Velde ER. Impact of repeated antral follicle counts on the prediction of poor ovarian response in women undergoing in vitro fertilization. Fertil Steril. 2004a;81:35-41.

Bancsi LF, Broekmans FJ, Looman CW, Habbema JD, te Velde ER. Predicting poor ovarian response in IVF: use of repeat basal FSH measurement. J Reprod Med. 2004b;49:187-194.

Bancsi LF, Broekmans FJ, Mol BW, Habbema JD, te Velde ER. Performance of basal follicle-stimulating hormone in the prediction of poor ovarian response and failure to become pregnant after in vitro fertilization: a meta-analysis. Fertil Steril. 2003;79:1091-1100.

Broekmans FJ, de ZD, Howles CM, Gougeon A, Trew G, Olivennes F. The antral follicle count: practical recommendations for better standardization. Fertil Steril. 2009.

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;11:607-614.

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;12:685-718.

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. 2008.

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;91:705-714.

Centers for Disease Control and Prevention (2007) 2005 Assisted Reproductive Technology Success Rates: National Summary and Fertility Clinic Reports. In pp. http://www.cdc.gov/ART/ART2006-508PDF/2006ART.pdf.

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

de Bruin JP, Bovenhuis H, Van Noord PA, Pearson PL, van Arendonk JA, te Velde ER, Kuurman WW, Dorland M. The role of genetic factors in age at natural menopause. Hum Reprod. 2001;16:2014-2018.

Deeks JJ. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. BMJ. 2001;323:157-162.

Eijkemans MJ, Habbema J.D.F., te Velde ER. Age at last childbirth and fertility at young age. In In: Fertility In Populations and In Patients; M.J. Eijkemans; Academic Thesis, Erasmus Medical Center Rotterdam, The Netherlands, 2005, pp. 23-34.

Evers JL. Female subfertility. Lancet. 2002;360:151-159.

Fanchin R, de Ziegler D, Olivennes F, Taieb J, Dzik A, Frydman R. Exogenous follicle stimulating hormone ovarian reserve test (EFORT): a simple and reliable screening test for detecting 'poor responders' in in-vitro fertilization. Hum Reprod. 1994;9:1607-1611.

Fasouliotis SJ, Simon A, Laufer N. Evaluation and treatment of low responders in assisted reproductive technology: a challenge to meet. J Assist Reprod Genet. 2000;17:357-373.

Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. Lancet. 2005;365:1500-1505.

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;23:699-708.

Hansen KR, Morris JL, Thyer AC, Soules MR. Reproductive aging and variability in the ovarian antral follicle count: application in the clinical setting. Fertil Steril. 2003;80:577-583.

Harrison RF, Jacob S, Spillane H, Mallon E, Hennelly B. A prospective randomized clinical trial of differing starter doses of recombinant follicle-stimulating hormone (follitropin-beta) for first time in vitro fertilization and intracytoplasmic sperm injection treatment cycles. Fertil Steril. 2001;75:23-31.

Heijnen EM, Eijkemans MJ, De Klerk C, Polinder S, Beckers NG, Klinkert ER, Broekmans FJ, Passchier J, te Velde ER, Macklon NS, et al. A mild treatment strategy for in-vitro fertilisation: a randomised non-inferiority trial. Lancet. 2007;369:743-749.

Hendriks DJ, Broekmans FJ, Bancsi LF, de Jong FH, Looman CW, te Velde ER. Repeated clomiphene citrate challenge testing in the prediction of outcome in IVF: a comparison with basal markers for ovarian reserve. Hum Reprod. 2005a;20:163-169.

Hendriks DJ, Broekmans FJ, Bancsi LF, Looman CW, de Jong FH, te Velde ER. Single and repeated GnRH agonist stimulation tests compared with basal markers of ovarian reserve in the prediction of outcome in IVF. J Assist Reprod Genet. 2005b;22:65-73.

Hendriks DJ, Kwee J, Mol BW, te Velde ER, Broekmans FJ. Ultrasonography as a tool for the prediction of outcome in IVF patients: a comparative meta-analysis of ovarian volume and antral follicle count. Fertil Steril. 2007;87:764-775.

Hendriks DJ, Mol BW, Bancsi LF, te Velde ER, Broekmans FJ. Antral follicle count in the prediction of poor ovarian response and pregnancy after in vitro fertilization: a meta-analysis and comparison with basal follicle-stimulating hormone level. Fertil Steril. 2005;83:291-301.

Hendriks DJ, Mol BW, Bancsi LF, te Velde ER, Broekmans FJ. The clomiphene citrate challenge test for the prediction of poor ovarian response and nonpregnancy in patients undergoing in vitro fertilization: a systematic review. Fertil Steril. 2006.

Hendriks DJ, te Velde ER, Looman CW, Bancsi LF, Broekmans FJ. Expected poor ovarian response in predicting cumulative pregnancy rates: a powerful tool. Reprod Biomed Online. 2008;17:727-736.

Henne MB, Stegmann BJ, Neithardt AB, Catherino WH, Armstrong AY, Kao TC, Segars JH. The combined effect of age and basal follicle-stimulating hormone on the cost of a live birth at assisted reproductive technology. Fertil Steril. 2008;89:104-110.

Hoveyda F, Engmann L, Steele J, Lopez BA, Barlow DH. Ovarian response in three consecutive in vitro fertilization cycles. Fertil Steril. 2002;77:706-710.

Hsieh YY, Chang CC, Tsai HD. Antral follicle counting in predicting the retrieved oocyte number after ovarian hyperstimulation. J Assist Reprod Genet. 2001;18:320-324.

Karlsson MO, Wade JR, Loumaye E, Munafo A. A population model for the follicular growth in women treated with follicle stimulating hormone. Clin Pharmacol Ther. 1997;62:665-674.

Karlsson MO, Wade JR, Loumaye E, Munafo A. The population pharmacokinetics of recombinant- and urinary-human follicle stimulating hormone in women. Br J Clin Pharmacol. 1998;45:13-20.

Klinkert ER. Clinical significance and management of poor response in IVF. 2005.

Klinkert ER, Broekmans FJ, Looman CW, te Velde ER. A poor response in the first in vitro fertilization cycle is not necessarily related to a poor prognosis in subsequent cycles. Fertil Steril. 2004;81:1247-1253.

Kok JD, Looman CW, Weima SM, te Velde ER. A high number of oocytes obtained after ovarian hyperstimulation for in vitro fertilization or intracytoplasmic sperm injection is not associated with decreased pregnancy outcome. Fertil Steril. 2006;85:918-924.

Kolibianakis EM, Venetis CA, Diedrich K, Tarlatzis BC, Griesinger G. Addition of growth hormone to gonadotrophins in ovarian stimulation of poor responders treated by in-vitro fertilization: a systematic review and meta-analysis. Hum Reprod Update. 2009.

Kyrou D, Kolibianakis EM, Venetis CA, Papanikolaou EG, Bontis J, Tarlatzis BC. How to improve the probability of pregnancy in poor responders undergoing in vitro fertilization: a systematic review and meta-analysis. Fertil Steril. 2009;91:749-766.

Land JA, Yarmolinskaya MI, Dumoulin JC, Evers JL. High-dose human menopausal gonadotropin stimulation in poor responders does not improve in vitro fertilization outcome. Fertil Steril. 1996;65:961-965.

Lashen H, Ledger W, Lopez-Bernal A, Barlow D. Poor responders to ovulation induction: is proceeding to in-vitro fertilization worthwhile? Hum Reprod. 1999;14:964-969.

Lawson R, El Toukhy T, Kassab A, Taylor A, Braude P, Parsons J, Seed P. Poor response to ovulation induction is a stronger predictor of early menopause than elevated basal FSH: a life table analysis. Hum Reprod. 2003;18:527-533.

Lee TH, Liu CH, Huang CC, Wu YL, Shih YT, Ho HN, Yang YS, Lee MS. Serum anti-Mullerian hormone and estradiol levels as predictors of ovarian hyperstimulation syndrome in assisted reproduction technology cycles. Hum Reprod. 2008;23:160-167.

Levi AJ, Raynault MF, Bergh PA, Drews MR, Miller BT, Scott RT. Reproductive outcome in patients with diminished ovarian reserve. Fertil Steril. 2001;76:666-669.

Lorusso F, Vicino M, Lamanna G, Trerotoli P, Serio G, Depalo R. Performance of different ovarian reserve markers for predicting the numbers of oocytes retrieved and mature oocytes. Am J Human Biol. 2007;56:429-435.

Loumaye E., Bilion J.M., Mine J.M., Psalti I., Pensis M, Thomas K. Prediction of individual response to controlled ovarian hyperstimulation by means of a clomiphene citrate challenge test. Fertil Steril. 1990;53:295-301.

Maseelall PB, Hernandez-Rey AE, Oh C, Maagdenberg T, McCulloh DH, McGovern PG. Antral follicle count is a significant predictor of livebirth in in vitro fertilization cycles. Fertil Steril. 2009;91:1595-1597.

Moce E, Lavara R, Vicente JS. Influence of the donor male on the fertility of frozen-thawed rabbit sperm after artificial insemination of females of different genotypes. Reprod Domest Anim. 2005;40:516-521.

Mochtar MH, Van der Veen F, Ziech M, van Wely M. Recombinant Luteinizing Hormone (rLH) for controlled ovarian hyperstimulation in assisted reproductive cycles. Cochrane Database Syst Rev. 2007, CD005070.

Mol BW, Verhagen TE, Hendriks DJ, Collins JA, Coomarasamy A, Opmeer BC, Broekmans FJ. Value of ovarian reserve testing before IVF: a clinical decision analysis. Hum Reprod. 2006;21:1816-1823.

Morabia A and Costanza MC. International variability in ages at menarche, first livebirth, and menopause. World Health Organization Collaborative Study of Neoplasia and Steroid Contraceptives. Am J Epidemiol. 1998;148:1195-1205.

Muasher SJ, Oehninger S, Simonetti S, Matta J, Ellis LM, Liu HC, Jones GS, Rosenwaks Z. The value of basal and/or stimulated serum gonadotropin levels in prediction of stimulation response and in vitro fertilization outcome. Fertil Steril. 1988;50:298-307.

Nargund G, Fauser BC, Macklon NS, Ombelet W, Nygren K, Frydman R. The ISMAAR proposal on terminology for ovarian stimulation for IVF. Hum Reprod. 2007;22:2801-2804.

Nelson SM, Yates RW, Fleming R. Serum anti-Mullerian hormone and FSH: prediction of live birth and extremes of response in stimulated cycles – implications for individualization of therapy. Hum Reprod. 2007;22:2414-2421.

Nelson SM, Yates RW, Lyall H, Jamieson M, Traynor I, Gaudoin M, Mitchell P, Ambrose P, Fleming R. Anti-Mullerian hormone-based approach to controlled ovarian stimulation for assisted conception. Hum Reprod. 2009;24:867-875.

Ng EH, Tang OS, Ho PC. The significance of the number of antral follicles prior to stimulation in predicting ovarian responses in an IVF programme. Hum Reprod. 2000;15:1937-1942.

Noord-Zaadstra BM, Looman CW, Alsbach H, Habbema JD, te Velde ER, Karbaat J. Delaying childbearing: effect of age on fecundity and outcome of pregnancy. BMJ. 1991;302:1361-1365.

Nyboe AA, Goossens V, Bhattacharya S, Ferraretti AP, Kupka MS, de MJ, Nygren KG. Assisted reproductive technology and intrauterine inseminations in Europe, 2005: results generated from European registers by ESHRE: ESHRE. The European IVF Monitoring Programme (EIM), for the European Society of Human Reproduction and Embryology (ESHRE). Hum Reprod. 2009.

Olivennes F, Howles CM, Borini A, Germond M, Trew G, Wikland M, Zegers-Hochschild F, Saunders H, Alam V. Individualizing FSH dose for assisted reproduction using a novel algorithm: the CONSORT study. Reprod Biomed Online. 2009;18:195-204.

Out HJ, David I, Ron-El R, Friedler S, Shalev E, Geslevich J, Dor J, Shulman A, Ben Rafael Z, Fisch B, et al. A randomized, double-blind clinical trial using fixed daily doses of 100 or 200 IU of recombinant FSH in ICSI cycles. Hum Reprod. 2001;16:1104-1109.

Penarrubia J, Fabregues F, Manau D, Creus M, Carmona F, Casamitjana R, Vanrell JA, Balasch J. Previous cycle cancellation due to poor follicular development as a predictor of ovarian response in cycles stimulated with gonadotrophin-releasing hormone agonist-gonadotrophin treatment. Hum Reprod. 2005;20:622-628.

Polinder S, Heijnen EM, Macklon NS, Habbema JD, Fauser BJ, Eijkemans MJ. Cost-effectiveness of a mild compared with a standard strategy for IVF: a randomized comparison using cumulative term live birth as the primary endpoint. Hum Reprod. 2007.

Popovic-Todorovic B, Loft A, Bredkjaeer HE, Bangsboll S, Nielsen IK, Andersen AN. A prospective randomized clinical trial comparing an individual dose of recombinant FSH based on predictive factors versus a 'standard' dose of 150 IU/day in 'standard' patients undergoing IVF/ICSI treatment. Hum Reprod. 2003a;18:2275-2282.

Popovic-Todorovic B, Loft A, Lindhard A, Bangsboll S, Andersson AM, Andersen AN. A prospective study of predictive factors of ovarian response in 'standard' IVF/ICSI patients treated with recombinant FSH. A suggestion for a recombinant FSH dosage normogram. Hum Reprod. 2003b;18:781-787.

Popovic-Todorovic B, Loft A, Ziebe S, Andersen AN. Impact of recombinant FSH dose adjustments on ovarian response in the second treatment cycle with IVF or ICSI in "standard" patients treated with 150 IU/day during the first cycle. Acta Obstet Gynecol Scand. 2004;83:842-849.

Scott RT, Jr., Elkind-Hirsch KE, Styne-Gross A, Miller KA, Frattarelli JL. The predictive value for in vitro fertility delivery rates is greatly impacted by the method used to select the threshold between normal and elevated basal follicle-stimulating hormone. Fertil Steril. 2008;89:868-878.

Scott RT, Jr. and Hofmann GE. Prognostic assessment of ovarian reserve [see comments]. Fertil Steril. 1995;63:1-11.

Scott RT, Toner JP, Muasher SJ, Oehninger S, Robinson S, Rosenwaks Z. Follicle-stimulating hormone levels on cycle day 3 are predictive of in vitro fertilization outcome. Fertil Steril. 1989;51:651-654.

Seifer DB, Lambert Messerlian G, Hogan JW, Gardiner AC, Blazar AS, Berk CA. Day 3 serum inhibin-B is predictive of assisted reproductive technologies outcome [see comments]. Fertil Steril. 1997;67:110-114.

Seifer DB, MacLaughlin DT, Christian BP, Feng B, Shelden RM. Early follicular serum mullerian-inhibiting substance levels are associated with ovarian response during assisted reproductive technology cycles. Fertil Steril. 2002;77:468-471.

Shanbhag S, Aucott L, Bhattacharya S, Hamilton MA, McTavish AR. Interventions for 'poor responders' to controlled ovarian hyperstimulation (COH) in in-vitro fertilisation (IVF). Cochrane Database Syst Rev. 2007, CD004379.

Spira A. The decline of fecundity with age. Am J Human Biol. 1988, Suppl 1, 15-22.

Stephen EH and Chandra A. Declining estimates of infertility in the United States: 1982-2002. Fertil Steril. 2006;86:516-523.

Sun W, Stegmann BJ, Henne M, Catherino WH, Segars JH. A new approach to ovarian reserve testing. Fertil Steril. 2008;90:2196-2202.

Tarlatzis BC, Zepiridis L, Grimbizis G, Bontis J. Clinical management of low ovarian response to stimulation for IVF: a systematic review. Hum Reprod Update. 2003;9:61-76.

te Velde ER and Pearson PL. The variability of female reproductive aging. In pp. 141-154, 2002.

Templeton A, Morris JK, Parslow W. Factors that affect outcome of in-vitro fertilisation treatment. Lancet. 1996;348:1402-1406.

Thomas F, Renaud F, Benefice E, de Meeus T, Guegan JF. International variability of ages at menarche and menopause: patterns and main determinants. Hum Biol. 2001;73:271-290.

Tomas C, Nuojua Huttunen S, Martikainen H. Pretreatment transvaginal ultrasound examination predicts ovarian responsiveness to gonadotrophins in in-vitro fertilization. Hum Reprod. 1997;12:220-223.

Tremellen KP, Kolo M, Gilmore A, Lekamge DN. Antimullerian hormone as a marker of ovarian reserve. Aust N Z J Obstet Gynaecol. 2005;45:20-24.

van der Gaast MH, Eijkemans MJ, van der Net JB, de Boer EJ, Burger CW, van Leeuwen FE, Fauser BC, Macklon NS. Optimum number of oocytes for a successful first IVF treatment cycle. Reprod Biomed Online. 2006;13:476-480.

van Rooij IA, Broekmans FJ, te Velde ER, Fauser BC, Bancsi LF, Jong FH, Themmen AP. Serum anti-Mullerian hormone levels: a novel measure of ovarian reserve. Hum Reprod. 2002;17:3065-3071.

Verhagen TE, Hendriks DJ, Bancsi LF, Mol BW, Broekmans FJ. The accuracy of multivariate models predicting ovarian reserve and pregnancy after in vitro fertilization: a meta-analysis. Hum Reprod Update. 2008;14:95-100.

Viudes-de-Castro MP, Vicente JS. Effect of sperm count on the fertility and prolificity rates of meat rabbits. Anim Reprod Sci. 1997;46:313-319.

Vladimirov IK, Tacheva DM, Kalinov KB, Ivanova AV, Blagoeva VD. Prognostic value of some ovarian reserve tests in poor responders. Arch Gynecol Obstet. 2005;272:74-79.

Winslow KL, Toner JP, Brzyski RG, Oehninger SC, Acosta AA, Muasher SJ. The gonadotropin-releasing hormone agonist stimulation test – a sensitive predictor of performance in the flare-up in vitro fertilization cycle. Fertil Steril. 1991;56:711-717.

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

Yong PY, Brett S, Baird DT, Thong KJ. A prospective randomized clinical trial comparing 150 IU and 225 IU of recombinant follicle-stimulating hormone (Gonal-F*) in a fixed-dose regimen for controlled ovarian stimulation in in vitro fertilization treatment. Fertil Steril. 2003;79:308-315.

Zhen XM, Qiao J, Li R, Wang LN, Liu P. The clinical analysis of poor ovarian response in in-vitro-fertilization embryo-transfer among Chinese couples. J Assist Reprod Genet. 2008;25:17-22.