The predictive value of ovarian reserve tests for miscarriage in a population of subfertile ovulatory women

M.L. Haadsma1,2,6, H. Groen3, V. Fidler3, L.H.M. Seinen1, F.J.M. Broekmans4, M.J. Heineman1,5, and A. Hoek1

1Department of Obstetrics and Gynaecology, University Medical Center Groningen, University of Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands 2Department of Genetics, University Medical Center Groningen, University of Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands 3Department of Epidemiology, University Medical Center Groningen, University of Groningen, PO Box 30.001, 9700 RB Groningen, The Netherlands 4Department of Reproductive Medicine and Gynaecology, University Medical Center Utrecht, PO Box 85500, 3508 GA Utrecht, The Netherlands 5Department of Obstetrics and Gynaecology, Academic Medical Center, PO Box 22660, 1100 DD, Amsterdam, The Netherlands

6Correspondence address. E-mail: m.l.haadsma@medgen.umcg.nl

BACKGROUND: The increase in miscarriage rate with female age is attributed to a decline in oocyte quality. This age-related decrease of oocyte quality is accompanied by a decrease in oocyte quantity. Assessment of the number of oocytes by ovarian reserve tests (ORTs) may therefore also represent their quality. The objective of our study was to assess the predictive value of ORTs for miscarriage in subfertile women.

METHODS: This study was a subanalysis within a prospective cohort study of 474 subfertile ovulatory couples in two hospitals in Groningen, The Netherlands. The ORTs performed were: antral follicle count (AFC), basal and stimulated levels of follicle-stimulating hormone (FSH) and inhibin B, and the clomiphene citrate challenge test (CCCT). Women who achieved an ongoing pregnancy (n = 233) were compared with women experiencing miscarriage (n = 72) on the results of their ORTs and patient characteristics.

RESULTS: In univariate analysis, the outcome of the ORTs did not differ between the groups. Logistic regression analysis including patient characteristics such as female age did not reveal an association between the ORT results and miscarriage either.

CONCLUSIONS: Neither AFC, basal and stimulated levels of FSH and inhibin B, nor the CCCT have a statistically significant predictive value for miscarriage in subfertile ovulatory women.

Key words: ovarian reserve / ovarian reserve test / miscarriage / pregnancy loss / predictive value

Introduction

The chance for a woman to have a live birth declines as her age increases. This is partly due to a diminished chance to conceive, but mostly caused by an increased chance of (very early) pregnancy loss (O’Connor et al., 1998). This process of female reproductive ageing is attributed to a decrease in both oocyte quantity, eventually resulting in menopause, and oocyte quality (Te Velde and Pearson, 2002). The main acknowledged manifestation of decreased quality is the occurrence of chromosomal abnormalities in the oocyte, leading to aneuploidy in the conceptus, which has been established as the reason of pregnancy loss in 35–75% of all cases (Baird et al., 2005; Ljunger et al., 2005; Rai and Regan, 2006).

The quality of a woman’s oocytes cannot be assessed clinically, but the quantity of the remaining follicle pool can be estimated by so-called ovarian reserve tests (ORTs). Various studies have suggested that the quantitative ovarian reserve is predictive for the chance of miscarriage. An elevated level of basal follicle-stimulating hormone (FSH), low level of anti-Müllerian hormone (AMH) and low antral follicle count (AFC) have been described to be related to increased miscarriage rates (Levi et al., 2001; Elter et al., 2005; Lekamge et al., 2007). Also, a high incidence of decreased ovarian reserve has been found among women with unexplained recurrent pregnancy loss (Trout and Seifer, 2000; Gurbuz et al., 2004). In line with these findings are publications suggesting a relationship between decreased ovarian reserve and chromosomal abnormalities in the conceptus (Nasseri et al., 1999; Van Montfrans et al., 1999; Freeman et al., 2000; Kline et al., 2000).

The studies describing an association between quantitative ovarian reserve and miscarriage or chromosomal abnormalities are challenged...
by a number of comparable studies that do not demonstrate such a relation (Hofmann et al., 2000; Abdalla and Thurn, 2004; Kline et al., 2004; Van Montfrans et al., 2004; Havryliuk et al., 2006; Luna et al., 2007; Massie et al., 2007). Furthermore, most of the studies on this topic are of retrospective design or include small numbers of subjects. We decided to address the issue in a large prospective cohort study of subfertile ovulatory women to answer the question whether ORTs have independent predictive value for the chance of miscarriage.

Materials and Methods

The present study is part of a prospective cohort study, originally designed to assess the predictive value of ORTs for spontaneous pregnancy in an ovulatory subfertile population (Haadsma et al., 2008).

Study population

From December 1999 to July 2003, patients were recruited at the tertiary fertility centre of the University Medical Center Groningen and the fertility centre of the Martini Hospital, a teaching hospital, in Groningen, the Netherlands. Patients were asked to participate after a routine subfertility evaluation had been completed. This evaluation included assessment of ovulation by sonographic cycle analysis and measurement of midluteal progesterone, semen analysis, post-coital test and hysterosalpingography. The inclusion criteria for study participation were: (i) subfertility for at least 12 months, (ii) a regular ovulatory cycle with midluteal progesterone of >30 nmol/l and a mean cycle length between 21 and 42 days, (iii) at least one result of hysterosalpingography or before starting treatment with intrauterine (e.g. a history of pelvic inflammatory disease), after an abnormal diagnostic laparoscopy was performed if tubal pathology was suspected. Patients were excluded secondarily. Informed consent was obtained from all participants. The study protocol was reviewed and approved by the Institutional Research Boards of both participating centres. Ovarian reserve tests and miscarriage was 3.7% and the sensitivity was <0.05 IU/l. The lower limit of detection was 0.03 IU/l. The standard of the FSH assay was calibrated against the WHO Second International Reference Preparation for human FSH (78/549). Inhibin B concentrations were assayed with Enzyme Immunoassay (ELISA) from Serotec (Kidlington, Oxford, UK). The inter-assay coefficient of variation for inhibin B was 11% and the sensitivity was <10 pg/ml. The lower limit of detection was 5 pg/ml.

Follow-up

After completion of the basal subfertility evaluation, expectant management or treatment was proposed to each couple. The advice was based on duration of subfertility and the estimated chances for spontaneous conception according to the prediction model routinely used in both clinics at the time of study (Eimers et al., 1994). Couples were advised to start treatment if the estimated chance to conceive was below 30% or duration of subfertility exceeded 3 years (2 years in case of female age ≥37 years). Couples who were primarily advised expectant management and did not conceive spontaneously were also offered treatment as soon as they met these criteria. The kind of treatment offered to the couples with a low chance of spontaneous conception depended on the type of subfertility. Patients were offered treatment by IUI with stimulation, in case of unexplained subfertility and mild male factor, or IUI without stimulation, in case of cervical factor, up to a maximum of six cycles. If IUI treatment failed, couples were offered in vitro fertilization (IVF) treatment, generally up to a maximum of three cycles because of the reimbursement policy of the medical insurance companies. If semen quality was insufficient for IUI, couples were advised treatment with IVF or intracytoplasmic sperm injection (ICSI). Treatment was not generally available for women over the age of 40.

Follow-up started on the day of the first ORT. Data on pregnancies and treatment were recorded. Information was obtained from medical files and from questionnaires completed by the participating couples. Couples were followed until a pregnancy leading to a live birth was achieved, either after spontaneous conception or after treatment. For the non-pregnant couples, follow-up ended after fertility treatment was fully completed. Couples who never started treatment or withdrew from treatment received a questionnaire to complete data. Follow-up also ended when couples started contraceptives or ended their relationship. The last date of follow-up was 1 November 2006. All couples who achieved a pregnancy during the follow-up period were identified. For each couple, only the first pregnancy during follow-up was taken into account. Ongoing pregnancies and miscarriages were selected for analysis. An ongoing pregnancy was defined as a viable intrauterine pregnancy of at least 16 weeks gestation. A miscarriage was defined as a pregnancy loss between 4 and 16 weeks amenorrhoea, with the exception of confirmed extra-uterine pregnancies or artificially terminated pregnancies. Time to pregnancy was defined as the period between the first ORT and the first day of the last menstrual cycle. Assisted conception was defined as conception after treatment with IVF or ICSI as opposed to spontaneous conception or conception after IUI. Female age and duration of subfertility were scored on the first day of the last menstrual cycle. For the other characteristics, values at the time of the first ORT were used when applicable.

Statistical analysis

To analyse the relationship between ORTs and pregnancy outcome, we performed univariate and multivariate analyses. The univariate analyses were scored on the first day of the last menstrual cycle. Assisted conception was defined as conception after treatment with IVF or ICSI as opposed to spontaneous conception or conception after IUI. Female age and duration of subfertility were scored on the first day of the last menstrual cycle. For the other characteristics, values at the time of the first ORT were used when applicable.
with miscarriage status as the outcome and the ORTs and patient characteristics as predictors. To explore the nature of the relationship with miscarriage of each of the continuous predictors, we used basic spline functions (Harrell, 2001). In this way, the possibility of a linear and non-linear relationship of each of the variables and miscarriage was assessed, including the presence of threshold points above or below which miscarriage rates changed. The effect of the variables was described as odds ratios (OR) with 95% confidence intervals (CI). A P-value \( < 0.05 \) was considered statistically significant. Data were analysed with SPSS 14.0 (SPSS Inc., Chicago, IL, USA) and S-plus 7.0 (MathSoft Inc., Seattle, WA, USA).

**Results**

From the 474 couples included in the prospective cohort study, 320 (67.5%) conceived during follow-up (Fig. 1). Of these, 305 of the 320 pregnant couples were selected for analysis. Of these, 233 had an ongoing pregnancy and 72 experienced miscarriage. Fig. 1 also presents the reasons for exclusion. The grounds for artificial termination were once, congenital abnormalities of a fetus with normal karyotype and once, personal reasons. The group of 15 excluded couples did not differ from the selected 305 couples with respect to patient characteristics and results of the ORTs (data not shown). Of the 305 selected couples, 132 (43%) conceived spontaneously, 94 (31%) after IUI, 36 (12%) after IVF and 43 (14%) after ICSI. Of the 132 couples that conceived spontaneously, 37 (28%) had already started therapy and conceived in between treatment cycles or after discontinuing treatment. Of the 154 non-pregnant couples, 19 were lost to follow-up, 14 ended their relationship and 13 started contraceptives during the study period. Median follow-up for the 154 non-pregnant couples was 24.8 months (10th–90th percentiles 2.7–46.9 months).

Table I presents a univariate comparison of patient characteristics and results of the ORTs of the groups with ongoing pregnancy and miscarriage. The miscarriage group was older and had more often conceived through ART. Logistic regression with splines revealed a statistically significant non-linear association of BMI with miscarriage which could be simplified to a piecewise linear relation: up to a BMI of 32 no change in miscarriage rate was observed, but from a BMI of 32 onwards the probability of miscarriage increased. For none of the ORTs, a relation with miscarriage could be demonstrated, both in analyses with and without correction for possible confounders. The best fitting model for miscarriage included female age (OR per year 1.06; 95% CI 0.99–1.14), assisted conception (OR yes to no 1.95; 95% CI 1.02–3.75) and BMI above 32 (OR per kg/m\(^2\) 1.57; 95% CI 1.12–2.21). The c-statistic (area under the receiver–operator curve) for this model was 0.65.

**Discussion**

The present study shows that basal and stimulated FSH, the CCCT, basal and stimulated inhibin B and AFC all have no statistically significant predictive value for the chance of miscarriage in a population of subfertile ovulatory women.

This finding is surprising as a decreased quantitative ovarian reserve is considered to be a reflection of advanced ovarian ageing and ovarian ageing is clearly associated with an increased rate of fetal aneuploidy and miscarriage. A possible explanation may be that ORTs basically relate to the number of remaining oocytes and that their quantity is unrelated to their quality. Oocyte quality might predominantly be determined by biological damage accumulated over time and would thus not be related to the number of oocytes left, but mainly to female age (Tarín, 1995; Eichenlaub-Ritter, 1998). This hypothesis is compatible with recent publications demonstrating that female age does predict pregnancy chances, both spontaneous and after assisted conception, but ORTs do not (Broekmans et al., 2006; Maheshwari et al., 2006; Van Rooij et al., 2006; Van der Steeg et al., 2007). Also, young women who respond poorly to ovarian stimulation during IVF treatment appear to have clearly better prospects than their older counterparts (Hanoch et al., 1998; Saldeen et al., 2007). Alternatively, it has been hypothesized that a biological relation between quantity and quality of oocytes does exist. The so-called production-line theory states that the germ cells produced earliest during fetal life are the least prone to non-disjunction. These oocytes are selected for ovulation first, leaving the oocytes of lesser quality for later years (Henderson and Edwards, 1968; Eichenlaub-Ritter, 1998). The reason that ORTs do not predict miscarriage could hence be that these tests do not accurately reflect oocyte quantity. For instance, an elevated FSH level may indeed be due to a decreased number of follicles, but also to a range of other causes including the presence of heterophilic antibodies or...
FSH-receptor polymorphisms (Lambalk and De Koning, 1998; De Koning et al., 2000, 2006). Especially when the relation between oocyte quantity and quality is subtle, inaccurate estimation of quantitative ovarian reserve might obscure this relation. Finally, the etiology of miscarriage is known to be diverse. If any, ovarian reserve is only one of many contributing factors.

Another possible explanation for the absence of a relation between ovarian reserve and the chance of miscarriage in our population is that the relation between oocyte quantity and quality does exist, but only at the very end of the reproductive period, when ovarian reserve is severely diminished. For that reason, we also analysed the extremes in our population separately, but in this manner we did not find an indication for a relation between ORTs and miscarriage either. For example, 5% of our population had a basal FSH level of $\geq 12$ IU/l ($n = 15$); 4 of these women (27%) miscarried compared with 23% in the population with FSH $<12$ IU/l ($P = 0.73$). However, women with a severely decreased ovarian reserve were not likely to be included in our study population, since they often have irregular and anovulatory cycles, which was an exclusion criterion in our study. In general, we cannot exclude the possibility that the relation between ovarian reserve and miscarriage does exist, but was not discovered in our subfertile study population, since differences between the various ORT values may have been too small.

Our findings that miscarriage rates are increased with higher BMI and after conception through ART have been described before (Wang et al., 2004; Metwally et al., 2008). Furthermore, we found no effect of male factor on miscarriage rate, which minimizes the possibility that a relation between ovarian reserve and miscarriage was obscured by the inclusion of couples with male factor subfertility. Next to analysing a relation between total motile semen count and miscarriage, we assessed the semen parameters volume, concentration, percentage motile sperm and percentage sperm of normal morphology and found no linear or non-linear relation between these factors and miscarriage either (data not shown). However, since one of the inclusion criteria of our study was a total motile semen count of at least one million, our results do not exclude the possibility that a very low semen quality actually does influence the miscarriage risk.

Limitations

The present study is part of a prospective cohort study, which was originally designed to assess the predictive value of ORTs for the chance of spontaneous pregnancy; no power analysis was performed (Haadsma et al., 2008). However, three acknowledged predictive factors for miscarriage were identified in our study (female age, BMI

### Table I Patient characteristics and results of ovarian reserve tests according to pregnancy outcome

<table>
<thead>
<tr>
<th>Ongoing pregnancy ($n = 233$)</th>
<th>Miscarriage ($n = 72$)</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>32.4 27.0–38.1</td>
<td>34.2 28.0–39.6</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>22.8 19.4–30.1</td>
<td>22.9 19.5–33.2</td>
</tr>
<tr>
<td>Smoking habit*</td>
<td>61 28.4%</td>
<td>16 27.1%</td>
</tr>
<tr>
<td>Duration of subfertility (years)</td>
<td>3.3 1.9–5.8</td>
<td>3.4 1.9–5.4</td>
</tr>
<tr>
<td>Primary subfertility*</td>
<td>162 69.5%</td>
<td>52 72.2%</td>
</tr>
<tr>
<td>Previous miscarriage*</td>
<td>35 15.1%</td>
<td>14 19.4%</td>
</tr>
<tr>
<td>Mean cycle length (days)</td>
<td>28 26–33</td>
<td>28 25–32</td>
</tr>
<tr>
<td>Semen analysis (TMC, $\times 10^6$)</td>
<td>38.6 4.1–179.8</td>
<td>33.6 4.1–205.0</td>
</tr>
<tr>
<td><strong>Diagnostic category of subfertility</strong>*</td>
<td>125 53.6%</td>
<td>35 48.6%</td>
</tr>
<tr>
<td>Unexplained</td>
<td>99 42.2%</td>
<td>34 47.2%</td>
</tr>
<tr>
<td>Male factor</td>
<td>9 3.9%</td>
<td>3 4.2%</td>
</tr>
<tr>
<td>Cervical factor</td>
<td>8.8 1.2–28.4</td>
<td>10.6 0.8–34.4</td>
</tr>
<tr>
<td>Time to pregnancy (months)</td>
<td>52 22.3%</td>
<td>27 37.5%</td>
</tr>
<tr>
<td><strong>Results of ovarian reserve tests</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antral follicle count (n)</td>
<td>11 5–23</td>
<td>11 5–22</td>
</tr>
<tr>
<td>Basal FSH (IU/l) (bFSH)</td>
<td>6.3 4.5–9.6</td>
<td>6.3 4.6–10.1</td>
</tr>
<tr>
<td>Stimulated FSH (IU/l) (sFSH)</td>
<td>6.2 3.9–10.7</td>
<td>6.5 4.2–9.7</td>
</tr>
<tr>
<td>CCCT (bFSH+sFSH)(IU/l)</td>
<td>12.9 9.0–19.6</td>
<td>12.8 9.1–19.6</td>
</tr>
<tr>
<td>Basal inhibin B (ng/l)</td>
<td>89.0 40.1–144.9</td>
<td>79.0 30.0–132.0</td>
</tr>
<tr>
<td>Stimulated inhibin B (ng/l)</td>
<td>230.0 98.0–144.9</td>
<td>238.5 113.2–416.8</td>
</tr>
</tbody>
</table>

TMC, total motile count; ART, assisted reproductive technology (including in vitro fertilization and intracytoplasmic sperm injection); FSH, follicle-stimulating hormone; CCCT, clomiphene citrate challenge test.
and assisted conception), which minimizes the chance that a major relation between ORTs and miscarriage was missed. Yet, the size of our study population might have been too small to detect a more subtle relationship between ORTs and miscarriage. To indicate the size of the effect of the ORTs on miscarriage risk, we chose the example of basal FSH. In order to be able to provide CI, we dichotomized basal FSH and included it in the best fitting model (consisting of age, BMI and conception with or without ART). For example, using FSH = 8 IU/l as threshold point, the effect was an OR 1.03 with a 95% CI of 0.48–2.2 and P = 0.95. Please note that our analysis of the relation between the ORTs and miscarriage was more refined than shown in this example as we explored the possible linear and non-linear nature of the relationship as well.

The results of most ORTs may vary per cycle in the same woman, especially basal FSH and the CCCT (Scott et al., 1990; Kwee et al., 2004; Hendriks et al., 2005; De Koning et al., 2008). In our study, all ORTs were only performed once per participant. It is not known whether repeating these tests would enhance their predictive value for miscarriage in subfertile populations. However, in a prospective study among fertile women, Van Montfrans et al. (2004) could not show a relation between repeatedly measured basal FSH and the chance of miscarriage.

In our study, we did not measure AMH, which is nowadays a promising ORT (Van Rooij et al., 2005; Visser et al., 2006). Since AMH was not such an acknowledged ORT at the start of the study in 1999, we did not perform this test and unfortunately no serum was stored.

Finally, it has been demonstrated that the shorter the duration of pregnancy before miscarriage, the higher the probability that the loss is caused by aneuploidy (Boue et al., 1985; Hassold et al., 1996; Hassold and Hunt, 2001). We defined miscarriage as pregnancy loss between 4 and 16 weeks, which is a wide definition, including late pregnancy loss. In addition, we cannot exclude that we missed several very early pregnancy losses since couples did not routinely perform a pregnancy test every month. Early biochemical pregnancy loss was most likely to be detected in patients receiving ART, since it is usual (though not obligatory) in our IVF clinic to perform a pregnancy test after each ART cycle. This might partly explain the predictive value of assisted conception for miscarriage in our population.

Comparison with other studies

Three small retrospective studies suggest a relation between ORTs and miscarriage. Levi et al. found 20 miscarriages among 28 pregnant women with highly elevated basal FSH levels, which is a significantly increased miscarriage rate (71%) compared with a large control group with normal FSH levels (Levi et al., 2001). Levi et al. included all subfertile women who visited their clinic, all women with irregular cycles who had possibly already entered menopausal transition. Elter et al. compared 28 women who m miscarried with 34 women who delivered a healthy baby after ICSI treatment. AFC proved to have predictive value for miscarriage, while female age, basal FSH and estradiol values did not (Elter et al., 2005). Most interestingly, Lekamge et al. demonstrated a significantly higher miscarriage rate amongst IVF-treated women with low levels of AMH (5/17, 29.4%) compared with women with high AMH levels (6/36, 16.7%) (Lekamge et al., 2007). These results are of special interest since their study population was not at risk for severely decreased ovarian reserve: among their inclusion criteria were a basal FSH level < 10 IU/l and a proven ovulatory cycle. The findings of Lekamge et al. have not yet been confirmed or rejected in other studies.

Among the publications supporting our findings is a prospective study of Van Montfrans et al. performed in women who were over 30 years of age without a history of subfertility, and pursuing a spontaneous pregnancy (Van Montfrans et al., 2004). Of the 86 pregnant women, 41 (48%) had a miscarriage, including very early pregnancy loss. No relationship between miscarriage and basal FSH level was found. Luna et al. performed a retrospective study within an IVF population. The miscarriage rate in the group with elevated basal FSH levels was 22.9% (11/48), similar to the controls with normal FSH (19.3%, 226/1169) (Luna et al., 2007). Abdalla and Thum also retrospectively evaluated miscarriage rates in an IVF population, selecting the outcome of the first IVF cycle, and found no differences between various groups when divided by three different threshold levels for basal FSH (Abdalla and Thum, 2004).

Clinical implications

The relation between oocyte quantity and oocyte quality is widely discussed, but not agreed on (yet). Therefore caution is required when using ORTs and interpreting “abnormal” results. The present study shows that ORTs have no predictive value for miscarriage in a subfertile ovulatory population. AMH may prove an exception but was not evaluated in our study. The results of our study are in line with studies showing no predictive value of ORTs for spontaneous and assisted conception in subfertile populations. Since the ORTs evaluated in this study have no apparent predictive value for both the chance to conceive and pregnancy outcome, we recommend not using them in the general subfertile ovulatory population, not only to avoid unnecessary testing itself, but also to avoid unsupported interpretation of their results.

Authors’ contribution


Acknowledgements

We thank Dr A Bukman and Dr PFJ Donderwinkel for the inclusion of patients and the performance of ovarian reserve tests. Parts of these data were presented at the 23rd Annual Meeting of the European Society of Human Reproduction and Embryology, 2007, Lyons, France.

Funding

University Medical Center Groningen and Organon, the Netherlands.
References


Luna M, Grunfeld L, Mukherjee T, Sandler B, Copperman AB. Moderately elevated levels of basal follicle-stimulating hormone in young patients predict low ovarian response, but should not be used to disqualify patients from attempting in vitro fertilization. Fertil Steril 2007;87:782–787.


Trout SW, Seifer DB. Do women with unexplained recurrent pregnancy loss have higher Day 3 serum FSH and estradiol values? Fertil Steril 2000;74:335–337.


Submitted on June 4, 2008; resubmitted on August 22, 2008; accepted on September 24, 2008.