Association between response to ovarian stimulation and miscarriage following IVF: an analysis of 124,351 IVF pregnancies

Sesh Kamal Sunkara1,*, Yacoub Khalaf1, Abha Maheshwari2, Paul Seed1, and Arri Coomarasamy3

1King’s College London, London, UK 2University of Aberdeen, Aberdeen, UK 3University of Birmingham, Birmingham, UK

*Correspondence address. E-mail: sksunkara@hotmail.com

Submitted on September 23, 2013; resubmitted on January 24, 2014; accepted on February 13, 2014

STUDY QUESTION: Is there a relationship between ovarian reserve, quantified as ovarian response to stimulation, and miscarriage rate following IVF treatment?

SUMMARY ANSWER: There is a strong association between the number of oocytes retrieved and miscarriage rate following IVF treatment, with the miscarriage rate decreasing with an increasing number of oocytes and then levelling off: poor responders have a higher miscarriage rate across all age groups.

WHAT IS ALREADY KNOWN: Poor ovarian response is a manifestation of a decline in the quantity of the primordial follicle pool. Whether poor ovarian response is associated with a decline in oocyte quality contributing to miscarriage is however debated.

STUDY DESIGN, SIZE, DURATION: Anonymous data were obtained from the Human Fertilization and Embryology Authority (HFEA), the statutory regulator of assisted reproduction treatment (ART) in the UK. The HFEA has collected data on all ART performed in the UK since 1991. Data from 1991 to June 2008 involving 402,185 stimulated fresh IVF cycles and 124,351 pregnancy outcomes were analysed.

PARTICIPANTS/ MATERIALS, SETTING, METHODS: Data on all women undergoing a stimulated fresh IVF treatment cycle with at least one oocyte retrieved during the period from 1991 to June 2008 were analysed for their early pregnancy outcomes.

MAIN RESULTS AND THE ROLE OF CHANCE: There was a strong association between the number of oocytes retrieved and the clinical miscarriage rate. The miscarriage rate fell from 20 to 13% with an increasing number of oocytes before levelling off. Stepwise logistic regression identified three cut-off points (4, 10 and 15 oocytes) at or beyond which the probability of clinical miscarriage fell. There was no increase in miscarriage rate with very high oocyte numbers (>20 oocytes). The lowest risk of miscarriage (9.9%) was for women under 38 years of age, with primary infertility without a female cause and producing more than three oocytes.

LIMITATIONS, REASONS FOR CAUTION: Although the analysis was performed only on stimulated IVF cycles (excluding unstimulated cycles), the data had the limitation that there was no information on the total gonadotrophin consumption. The model was adjusted for age and type of infertility, but the dataset contained no information on other confounders such as body mass index (BMI) of the women to allow adjustment.

WIDER IMPLICATIONS FOR THE FINDINGS: Analysis of this extensive dataset suggests that poor responders have a higher risk of clinical miscarriage, indicating that poor ovarian response is associated with a parallel decline in both oocyte quantity and quality. The miscarriage rate is also higher with advanced age, secondary infertility and a female cause of infertility compared with a younger age, male factor infertility and unexplained cause.

Key words: IVF / ovarian response / oocyte quality / miscarriage
Introduction

The miscarriage rate among pregnancies following assisted reproduction is estimated to be ~15% and is comparable to the miscarriage rate following spontaneous conception (Schieve et al., 2003). The incidence of miscarriage is higher among older women compared with younger women and is attributed to ovarian ageing. Advanced maternal age is the most common aetiological factor for oocyte aneuploidy (Griffin, 1996) and the resultant fetal aneuploidy is the most common cause of miscarriage. The most common cause of oocyte aneuploidy is non-disjunction of chromosomes in the first meiotic division. There is an increased likelihood of non-disjunction with ovarian ageing and aneuploidy reflects poor oocyte quality. It has been a matter of interest whether the ovarian reserve and, indirectly, ovarian response to stimulation quantified by the number of oocytes retrieved could be associated with the risk of miscarriage following IVF treatment. Poor ovarian response to stimulation is a reflection of ovarian ageing which is associated with a decline in the quantity of the primordial follicle pool (Broekmans et al., 2006). It is uncertain whether poor ovarian response is also associated with a decline in oocyte quality (an increased incidence of aneuploidy) and thereby an increased risk of miscarriage.

Studies addressing the association between ovarian response and the miscarriage rate following IVF treatment have reported varying results. A small study involving 84 women aged >35 years undergoing their first IVF cycle, with preimplantation genetic screening, showed no significant difference in the number of aneuploid embryos and in miscarriage rates among women with a poor ovarian response (≥4 oocytes) compared with normal responders (≥5 oocytes) (Setti et al., 2011). An earlier study which analysed 3179 pregnancies following IVF found little association between the number of oocytes retrieved and the risk of miscarriage following IVF treatment (De Sutter and Dhont, 2003). A recent study which compared pregnancy outcomes of 137 women with poor response and 1668 women with a normal response in their first IVF cycle found no significant difference in the miscarriage rate between poor responders and normal responders among women aged <36 years, but found a significantly higher risk of miscarriage in poor responders compared with normal responders among women aged ≥36 years (Haadsma et al., 2010a).

Given the discrepancies in the current literature on the miscarriage risk for poor responders, which could be attributed to study weaknesses such as a small sample size or the definition of poor response, we set out to answer the question by examining a large national database of IVF cycles. Furthermore, it is also a matter of interest to study the miscarriage rate in women with hyperresponse following ovarian stimulation. A recent study demonstrated a non-linear association between oocyte numbers and live birth following IVF treatment, with live birth rates decreasing with high oocyte numbers >20 (Sunkara et al., 2011). We therefore explored whether an increase in the miscarriage rate contributed to a decreasing live birth rate in women with a hyperresponse. The aim of this study was to determine whether there is an association between ovarian response to stimulation (in terms of the number of oocytes retrieved) and the miscarriage rate following IVF treatment. We used a large national (UK) database involving 402 185 stimulated fresh IVF cycles and 124 351 pregnancies to address this question.

Materials and Methods

Anonymous data were obtained from the Human Fertilization and Embryology Authority (HFEA), the statutory regulator of assisted reproduction treatment (ART) in the UK on all ART cycles carried out in the UK since 1991. A total of 787 030 ART cycles were recorded during the period from April 1991 to June 2008. The data were prospectively collected over the years; however, our research question was formulated and answered after the data collection, and thus the study is considered a retrospective cohort study. For the purpose of this study only fresh stimulated IVF + ICSI treatment cycles which had one or more oocytes retrieved were analysed. Data were obtained for the age group of the women (18–34, 35–37, 38–39, 40–42, 43–44, and 45 years and over), treatment period (1991–2008), type of infertility, cause of infertility, number of oocytes retrieved following stimulation and early pregnancy outcomes. The early pregnancy outcomes were the combined miscarriage rate (including biochemical and clinical miscarriages), the clinical miscarriage rate and the ectopic pregnancy rate. A clinical miscarriage is defined as a pregnancy loss (resulting in a non-pregnancy) after ultrasound detection of an intraterine pregnancy with a gestational sac (Schieve et al., 2003). A pregnancy loss before ultrasound detection of an intrauterine pregnancy is defined as a biochemical miscarriage.

Statistical analysis

A technical limitation of the database, and therefore the study, is that due to the process of anonymisation it is not possible to link the records of women who had two or more IVF cycles during the study period. Without this information no correction for clustering is possible and confidence intervals may be over-tight as a consequence. To allow for this wider 99.5% confidence intervals (CIs) are used instead of the more usual 95% where necessary.

The characteristics of the cohort are described using absolute and relative frequencies with 99.5% CIs. The combined miscarriage rate, clinical miscarriage rate and ectopic pregnancy rate were computed for the entire cohort and stratified by age group. To study the association between the number of oocytes and the miscarriage rate, a maximum likelihood logistic regression model was fitted with clinical miscarriage as the dependent variable and oocyte (cumulus oocyte complex) number and age as independent variables. To generate a prediction model for clinical miscarriage, separate training and validation sets were created. The entire cohort was split into two sets to validate the prediction model. The first, comprising data between 1991 and 2002, was used to derive the model while data generated from 2003 onwards were used to validate it. Prediction variables were: cause of infertility (ovulatroy, tubal, male factor, unexplained), type of infertility (primary or secondary), number of oocytes and female age category. Cycles missing information on the cause or type of infertility were not included in the analysis, leaving 105 417 cycles with clinical pregnancies with predictor variables and outcome data. The historical training set consisted of 58 267 cycles from 1991 to 2002; and the validation set consisted of 47 150 cycles from 2003 to 2008. Female age category and number of oocytes were replaced by a series of binary indicator variables; these recorded whether the variable was greater than or equal to a series of cut points (for age: 35, 38, 40, 43, 45 years; for oocytes: 2, 3, 4, 5, 10, 15, 20, 30, 40). A backward stepwise logistic regression model (P < 0.001) was fitted to the training set; and further checks were then carried out for interactions or terms that could be combined; using Akike’s Information Criterion (AIC) and the likelihood ratio test to compare models. The final model fitted to the training set was used to produce prediction scores for all cycles, which were divided into categories of width of 5%. The accuracy of the prediction score was tested by comparing the actual and predicted percentage of events in each category, and by calculating the ROC area.
Results

The process of data selection is detailed in Fig. 1. From the initial cohort of 787,030 ART cycles, 384,845 cycles were excluded from the analysis for the following reasons: cycles with missing data, cycles with no stimulation or where there was no information regarding the use of stimulation, cycles involving embryo donation, surrogacy, preimplantation genetic diagnosis (PGD), oocyte donation or oocyte sharing, frozen embryos, oocyte freezing, gamete intrafallopian transfer (GIFT) or IVF + zygote intrafallopian transfer (ZIFT), cycles where embryos were created for reasons other than treatment, and cycles with no oocytes retrieved and/or no fresh embryo transfer.

Overall 402,185 fresh stimulated IVF + ICSI cycles were eligible for analysis of which 124,368 cycles resulted in a pregnancy, giving an overall pregnancy rate of 30.9% (99.5% CI: 30.7–31.1%) for the entire cohort. However, 17 pregnancies were lost to follow up with no information on the early pregnancy outcomes. Therefore, 124,351 IVF + ICSI cycles/pregnancies with early pregnancy outcomes were analysed. There were 27,682 biochemical and clinical pregnancy losses, with a combined miscarriage rate of 22.3% (99.5% CI: 22.0–22.6%); 1281 ectopic pregnancies, with an ectopic pregnancy rate 1.03% (99.5% CI: 0.95–1.11%) and 18 molar pregnancies reported for the entire cohort. The combined miscarriage rate was calculated as the number of pregnancies (biochemical and clinical) ending in a miscarriage. There were 111,653 clinical pregnancies, with 16,283 subsequent miscarriages; the clinical miscarriage rate was therefore 14.6% (99.5% CI: 14.3–14.9%).

Early pregnancy outcomes stratified by age

The majority (57.2%) of the 124,368 IVF + ICSI cycles that resulted in a pregnancy were in women aged ≤34 years, 24.9% were in women aged...
35–37 years, 11.0% were in women aged 38–39 years, 6.1% were in women aged 40–42 years, 0.7% were in women aged 43–44 years and 0.1% were in women aged ≥45 years. The miscarriage rates (both the combined and clinical miscarriage rates) gradually increased with increasing female age (Fig. 2a). The combined miscarriage rate was 18.4% (99.5% CI: 18.0–18.9%) in women aged ≤34 years, 23.1% (CI: 22.4–23.8%) in women aged 35–37 years, 29.6% (CI: 28.5–30.7%) in women aged 38–39 years, 40.6% (CI: 39.0–42.2%) in women aged 40–42 years, 57.3% (CI: 52.5–62.1%) in women aged 43–44 years and 65.6% (CI: 53–76.9%) in women aged ≥45 years. The clinical miscarriage rate was 11.5% (99.5% CI: 11.2–11.9%) in women aged ≤34 years, 15.2% (CI: 14.6–15.8%) in women aged 35–37 years, 20.1% (CI: 19.0–21.1%) in women aged 38–39 years, 29.3% (CI: 27.7–30.9%) in women aged 40–42 years, 43.6% (CI: 38.1–49.2%) in women aged 43–44 years and 45.7% (CI: 30.3–61.7%) in women aged ≥45 years. The ectopic pregnancy rate remained constant across all age groups (Fig. 2b).

Clinical miscarriage rate stratified by time period

There was no particular trend in the clinical miscarriage rate over time (Fig. 2c). The miscarriage rate was 14.8% in 1991 and 15.8% in 2008. Hence, data from 1991 to 2002 were used for the prediction model and data from 2003 to 2008 were used for model validation.

Relationship between number of oocytes and clinical miscarriage rate

There was a strong association between the number of oocytes retrieved and the clinical miscarriage rate. The miscarriage rate fell from 20 to 13% with increasing numbers of oocytes before levelling off (Fig. 3a). Stepwise logistic regression (P < 0.001) identified three cut-points (4, 10 and 15 oocytes) at or beyond which the probability of clinical miscarriage fell. It failed to find any oocyte number above which the miscarriage rate rose again. The miscarriage rates were 20.0% (1–3 oocytes), 15.5% (4–9 oocytes), 13.8% (10–14 oocytes) and 13.1% (≥15 oocytes). After adjusting for female age, the effect of oocyte numbers on the clinical miscarriage rate was reduced, but not eliminated (Fig. 3b). The adjusted rates (based on multiple logistic regression) were 16.9% (1–3 oocytes), 14.4% (4–9 oocytes), 13.7% (10–14 oocytes) and 13.5% (≥15 oocytes).

Predicting clinical miscarriage

The principal predictors of clinical miscarriage in the training group (1991–2002) were age (5 categories: ≤37, 38–39, 40–42, 43–44 and ≥45 years), ≤ or >3 oocytes, cause of infertility, and primary or secondary infertility. Checking variants of this basic model using the log-rank test and the AIC showed that there was no advantage in treating tubal and ovulatory causes as different, so they were combined into female factor. Introducing interaction terms between these components was likewise not useful. The final prediction model could therefore be summarized as 40 separate probabilities (5 age categories × 2, oocyte groups × 2, cause groups × 2 and infertility groups). These are given in Table I. The lowest risk of miscarriage (9.9%) was for women under 38 years of age, with primary infertility without a female cause and producing more than three oocytes. All other groups had risks over 10%. The highest risks (typically over 40%) were for women aged ≥45 years; however, there were relatively few women in this group, and the precise percentages may be less reliable.

Validating the clinical miscarriage prediction model

The above model was used to calculate prediction scores for all women in the training and validation sets. ROC areas were 0.596 (training) and
Table II gives the actual miscarriage rates for different prediction scores in both the training and validation sets. Age is by far the most important determinant of miscarriage rate. While the prediction scores give accurate estimates of the event rates, there are no pregnancies which can be said to have a risk under 9%, and relatively few pregnancies for which the risk is estimated at over 25% (only 2359/47 150 or 5.0% in the validation set). This is borne out by the ROC curves (Fig. 4); and ROC areas: 0.596 in the training set, 0.599 in validation set. Although the model can be said to be validated, the information it will provide for any one patient is limited.

Discussion

The study results show that there is an association between the number of oocytes retrieved and the miscarriage rate following IVF treatment. The miscarriage rate is high with a low number of oocytes and stepwise logistic regression identified cut-off points of 4, 10 and 15 oocytes below which the miscarriage rate increased. It demonstrates that women with poor ovarian response (≤ 3 oocytes) have a higher risk of miscarriage following IVF treatment across all age groups. Women who have a hyper-response to ovarian stimulation do not generally have a significantly increased risk of miscarriage. Other possible factors such as a lower fertilization rate (due to a higher proportion of immature oocytes) and lower implantation rate due to endometrial factors as a result of the high estradiol levels (Valbuena et al., 2001; Mitwally et al., 2006; Joo et al., 2010) could be speculated as possible causes of reduced live birth rates demonstrated in hyperresponders compared with normal responders (Sunkara et al., 2011). The study also demonstrates that female age is the most important cause of miscarriage. The finding that the effect of oocyte numbers on miscarriage rate was maintained but reduced after adjusting for age shows that although age is a predominant factor in influencing miscarriage rate, there is also an association between low oocyte numbers and miscarriage, irrespective of age.

The low live birth rates in poor responders undergoing IVF treatment could therefore be explained by the availability of fewer embryos to
Table II  Prediction score performance.

<table>
<thead>
<tr>
<th>Prediction score for clinical miscarriage</th>
<th>Training set</th>
<th>Validation set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of cycles</td>
<td>Predicted rate</td>
</tr>
<tr>
<td>9–10%</td>
<td>16,074</td>
<td>9.9%</td>
</tr>
<tr>
<td>10–15%</td>
<td>28,076</td>
<td>12.3%</td>
</tr>
<tr>
<td>15–20%</td>
<td>8,862</td>
<td>16.6%</td>
</tr>
<tr>
<td>20–25%</td>
<td>3,102</td>
<td>22.2%</td>
</tr>
<tr>
<td>25–30%</td>
<td>1,259</td>
<td>27.7%</td>
</tr>
<tr>
<td>&gt;30%</td>
<td>894</td>
<td>35.7%</td>
</tr>
</tbody>
</table>

*99.5% CIs are used to compensate for narrowing of the intervals due to lack of record linkage for women undergoing multiple cycles.

Figure 4  ROC areas for training and validation sets.

The strength of this study lies in the large data set which allows generalizability of the findings. However individual women would have contributed to more than one cycle and outcome (miscarriage) in the data set which means that the true sample size is unknown, and the conventional significance tests and CIs may be anti-conservative. It would have been ideal to analyse one IVF cycle per woman, but the HFEA data, being anonymized, did not permit us to analyse IVF cycle per woman as the unit of analysis. Nevertheless, the very large data set allowed for precise estimates of quantities of interest; and significance levels could be set conservatively (P < 0.001) for model development and a 99.5% CI could be given for final estimates without weakening the study. The use of a historic training set meant that we could confirm that the model for predicting miscarriages was not affected by changes over time in the population seeking treatment or by developments in treatment practices. The model was adjusted for age, which was the most important confounder, and type of infertility, but the dataset lacked information on other potential confounders such as BMI. Although only IVF cycles having ovarian stimulation were included, there was no information on the gonadotrophin dose, which is another weakness of the study. It was assumed that all cycles had optimal gonadotrophin stimulation. However, irrespective of the stimulation dose of gonadotrophin, women with fewer oocytes had a higher miscarriage rate. Furthermore, as with any national database the HFEA data are reliant on the reporting of individual centres. Hence the reliability of the database would be influenced by the stringency of the measures adopted by individual centres to ensure accurate reporting.

Previous studies have found a higher risk of fetal aneuploidy associated with an elevated day 3 serum FSH and a reduced ovarian follicle pool (Nasseri et al., 1999; Haadsma et al., 2010). Another study found a significantly higher miscarriage rate in women with elevated basal serum FSH levels (Levi et al., 2001). A study by De Sutter and Dhont (2003) which showed no significant difference in the miscarriage rate between poor responders and normal responders defined poor ovarian response using an arbitrary cut-off of <5 oocytes. Whether the findings would be different if poor ovarian response was defined as ≤3 oocytes needs to be considered. In a study by Setti et al. (2011) which showed no significant difference in the aneuploidy rate and miscarriage in poor responders and normal responders, 34 poor responders were compared with 50 normal responders and embryos reaching at least the five cell stage were biopsied on Day 3 and analysed for seven chromosomes (X, 7, 13, 16, 18, 21, 22) using fluorescent in situ hybridization (FISH) techniques. This study could be criticized for its small sample size and for the inclusion of women with a mean age >40 years. Using stepwise logistic regression, we identified four oocytes as a cut point below which the risk of miscarriage increased significantly. There have been several debates over the years as to what defines a poor ovarian response to simulation. We can say,
from the findings of our study, that it can be justified to consider ≤3 oocytes retrieved to define a poor responder. This is the cut-off that has also been suggested in the recent ESHRE consensus definition of poor ovarian response (Ferraretti et al., 2011).

In conclusion, the analysis of this extensive data set has found an association between poor ovarian response and an increased risk of miscarriage. This information could potentially be valuable in informing women of the increased miscarriage risk associated with reduced ovarian reserve. Finally it would also be important to explore in future research the question of whether poor responders have an increased incidence of adverse obstetric outcomes.

Acknowledgement
We thank the HFEA for providing the data.

Authors’ roles
S.K.S. conceived the hypothesis. S.K.S. directed the data analysis by P.S. and drafted the manuscript. A.C. and P.S. appraised the write-up of the analysis. A.C., A.M. and Y.K. appraised the manuscript of its content.

Funding
No funding was used.

Conflict of interest
None declared.

References
Valbuena D, Martin J, de Pablo JL, Remohí J, Pellicer A, Simón C. Increasing levels of estradiol are deleterious to embryonic implantation because they directly affect the embryo. Fertil Steril 2001;76:962–968.