Predicting the chances of live birth for couples undergoing IVF-ICSI: a novel instrument to advise patients and physicians before treatment [version 1; peer review: 2 not approved]

Bruna Estácio da Veiga¹, Duarte Pedro Tavares¹, José Luis Metello², Fernando Ferreira³, Pedro Ferreira², José Manuel Fonseca¹

¹CA3 - CTS UNINOVA, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, 2829-516, Portugal
²CIRMA, Hospital Garcia de Orta, Almada, 2805-267, Portugal
³CTS UNINOVA, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, 2829-516, Portugal

Abstract

Background: In developed countries, the prevalence of infertility ranges from 3.5% to 16.7%. Therefore, the number of in vitro fertilization technique (IVF) and its subtype intracytoplasmic sperm injection (ICSI) treatments has been significantly increasing across Europe. Several factors affect the success rate of in vitro treatments, which can be used to calculate the probability of success for each couple. As these treatments are complicated and expensive with a variable probability of success, the most common question asked by IVF patients is "What are my chances of conceiving?". The main aim of this study is to develop a validated model that estimates the chance of a live birth before they start their IVF non-donor cycle.

Methods: A logistic regression model was developed based on the retrospective study of 737 IVF cycles. Each couple was characterized by 14 variables (woman's and man's age, duration of infertility, cause of infertility, woman's and man's body mass index (BMI), anti-Müllerian hormone (AMH), antral follicle count (AFC), woman's and man's ethnicity, woman's and man's smoking status and woman's and man's previous live children) and described with the outcome of the treatment "Live birth" or "No live birth".

Results: The model results showed that from the 14 variables acquired before starting the IVF procedures, only male factor, man's BMI, man's mixed ethnicity and level of AMH were statistically significant. The interactions between infertility duration and woman's age, infertility duration and man's BMI, AFC and AMH, AFC and woman's age, AFC and woman's BMI and AFC and disovulation were also statistically significant. The area under the receiver operating characteristic (AUROC) curve test for the discriminatory ability of the final prediction model is 0.700 (95% confidence interval (CI) 0.660–0.741).

Conclusions: This model might result in a new validated decision...
support system to help physicians to manage couples’ expectations.

**Keywords**
IVF, ICSI, prediction model, live birth, assisted reproduction, logistic regression
Introduction
Infertility is defined as a disease of the reproductive system defined by the failure to achieve a clinical pregnancy after 12 months or more of regular unprotected sexual intercourse. Taking this definition into account, it is a challenge to calculate the real prevalence or incidence rates for this condition.

According to the study performed by Boivin et al., the prevalence of infertility ranged from 3.5% to 16.7% in developed countries. Based on these authors’ estimates, 72.4 million women are currently infertile and, of these, 40.5 million are currently seeking infertility medical care. Most recent Portuguese data estimate that 9.8% of couples are infertile.

Infertility is a multifactorial disease. A male-related factor can be the cause, especially when there is a sperm or ejaculatory problem or it may be a female-related factor, namely when there is a disovulation or tubal dysfunction. Women’s age, ovarian reserve, weight, and lifestyle factors – such as alcohol consumption and exercise habits – have also been related to this. Nonetheless, in up to 20% of cases, no cause can be found.

Several treatments have been proposed during the last century, with the most innovative bring in vitro fertilization (IVF), which was developed by Robert Edwards in 1978. After 41 years, the two most important medical assistant reproduction techniques (MAR) are in vitro fertilization (IVF) and its subtype intracytoplasmic sperm injection (ICSI). The number of IVF and ICSI treatments has been increasing across Europe. According to the European Society of Human Reproduction and Embryology, the number of MAR cycles between 1997 and 2014 increased by 13%, reaching 776 556 cycles in Europe in 2015.

Despite the increasing number of MAR treatments, IVF success is not guaranteed. In healthy young couples, the probability of achieving a live birth is between 20% and 25% per month. This probability may increase by up to 60% with MAR techniques. The study by Malizia et al. also corroborates this finding, indicating that between 38% and 49% of couples who start IVF remain childless, even after undergoing up to six IVF cycles. In Portugal, the last report of MAR showed that the treatments success rate is around 25–30%.

MAR treatments are expensive, time-consuming, stressful, and may lead to anxiety, depression, or marital problems. Also, these treatments have possible complications such as ovarian hyperstimulation syndrome, bleeding, and infection, as well as multiple or premature births.

The success rates fluctuate between studies and are dependent on several factors. If the chances of live birth are low and the risk or the cost is too high, the couple may consider other options such as adoption or to remain childless. In other words, based on their specific probability of success, the couple may decide whether or not they proceed with the treatment. Therefore, the most common question asked by IVF patients is “What are my chances of conceiving?”

The answer to this tough question usually depends on the woman’s age and infertility diagnosis. Nevertheless, many more parameters are known to affect IVF outcomes such as sperm quality, hormone doses, and physiological factors. Given this, a decision support system would be helpful to ensure that infertile couples are well informed regarding their chances of success with IVF.

There have been various efforts to build prediction models to assist physicians in predicting MAR success. To our knowledge, the first predictive model ever built in this context is from Templeton et al. in 1996 using a logistic regression model to predict the probability of live birth for an individual woman using the woman’s age, number of previous live birth or pregnancies not resulting in a live birth, whether these were a result of previous IVF treatment, female causes of infertility, duration of infertility and the number of previous unsuccessful IVF treatments.

These authors found that the success of IVF decreased with female age and that women between 25 and 30 years were the most likely to have a live birth.

Considering the evolution of technology and the importance of other variables, Nelson and Lawlor developed a new model in 2011, using the same mathematical techniques, but including other factors such as the most prevalent causes of infertility, the source of the egg (donor or patient’s own), type of hormonal preparation used (antioestrogen, gonadotrophin, or hormone replacement therapy), whether or not ICSI was used, and the number of previous cycles (1, 2 or 3).

Taking into account that it is imperative to have validated models, in 2014, Velde et al. used their cohort to validate Templeton and Nelson and Lawlor’s model’s performance. They found that Templeton’s model underestimated success rates while Nelson’s model overestimated it.

Other relevant models of note include the one developed byMarca et al. which predicts live birth in assisted reproduction based on serum anti-Müllerian hormone (AMH) and women’s age, and the Mc Lernon et al. study that estimated the cumulative personal chance of a first live birth over a maximum of six complete cycles of IVF using data from 23 417 women in the UK. The estimates of the last study were only adjusted by the woman’s age, infertility duration, previous pregnancies, infertility causes, and type of treatment.

Nonetheless, when Leijdekkers performed an external validation of Mc Lernon’s model with Dutch women, it was found that there was an overestimation of the results and he decided to include biomarkers such as anti-Müllerian hormone (AMH), antral follicle count (AFC) and women’s body weight. Other studies indicate it is important to consider covariables such as body mass index (BMI), ethnicity and ovarian reserve and corroborate that the use of variables such as BMI, AFC, AMH, ethnicity, and smoking status should be predictors in the model.
Furthermore, a machine learning approach was also proposed, which includes decision trees, genetic algorithms, and k nearest neighbors classifiers. The main aim of this study was to integrate all the variables of interest referred in the literature to develop a validated model that estimates the chance of live birth for couples before they start their IVF non-donor cycle.

**Methods**

The present work is a retrospective study of data from IVF/ICSI cycles. The cycles were performed between 2012 and 2016 in the Centro de Infertilidade e Reprodução Medicamente Assistida (CIRMA) at Hospital Garcia de Orta, E.P.E., Almada, Portugal.

739 couples were considered once they met the criterion of being fresh non-donor cycles or cycles with live birth or without frozen embryos available.

The study was approved by the Hospital’s Ethics Committee for Health, couples signed an informed consent document before treatment begins at the first infertility consultation. Patients were informed that, under complete anonymity, the data may be used in scientific papers for public presentation or publication.

The IVF result for each couple in the study was classified as 1 (if, at least, one baby was born alive and survived for more than 1 month) and 0 (otherwise), which is the model’s dependent variable/primary outcome. This decision was based on other studies in this area.

In terms of the baseline characteristics used to develop this model and based on evidence from the published literature in this area, the woman’s and man’s age (years), duration of infertility (months), cause of infertility (categorised as tubal factor, endometriosis, disovulation, male factor, both female and male factor - depending on whether the cause of infertility underlies the woman or the man - multiple female factors, unexplained infertility or other), woman’s and man’s BMI (kg/m²), AMH (ng/mL), AFC (antral follicle count), woman’s and man’s ethnicity (categorised as Asian, Caucasian, Gipsy, Indian, Black or Mixed), woman’s and man’s smoking status (never, previous, present) and woman’s and man’s previous live children (yes or no) were considered. Antral follicle count was obtained by transvaginal ultrasound, and serum anti-Müllerian hormone levels were measured by blood analysis. There should be no bias associated with this study, however, possible sources of bias may arise from couples responses during the consultation and the entry of the data in the database.

A summary statistical evaluation was performed for each variable. A univariate analysis was first performed. All continuous variables were compared using the t-student test and the categorical one using the Chi-square test. A p-value <0.05 was considered statistically significant.

A binary logistic regression was developed using the primary outcome as binary (no=0 and yes=1) to achieve the aim of this study. An automatic backward selection process based on the Wald statistic was used to determine the final and the best logistic regression model. The model was based on the data from 737 couples (99.73%), because in 2 couples there was data missing (lack of BMI values in two couples) and as lezzoni indicated, it is not recommended to input missing values in health data.

As recommended by Hosmer and Lemeshow, the interactions between some variables were included, to increase the reliability of the model because as Fisher explained, the primary outcome can depend not only on one individual baseline characteristic but also on the relationship between baseline characteristics. The variables in interaction were: female’s age and infertility duration, BMI and oligospermia, AFC levels and women’s BMI and AFC and AMH levels.

The model’s performance was measured with the discriminative power assessed using the area under the curve (AUC) value, and the model was internally validated using the bootstrapping technique with 1000 iterations. All statistical procedures were computed using IBM SPSS Statistics 25, and it was considered an α level of 0.05, which means that was assumed that there isn’t statistical significance if the p-value of the test is above 0.05.

The STROBE cross-sectional reporting guidelines were adopted in this article.

**Results**

A total of 737 cycles were evaluated. The overall rate of at least one live birth was 31.4%.

Baseline characteristics of couples are presented in Table 1 and Table 2.

The average age for female participants was 34.04 years (Table 1) and 36.14 for male participants. Younger women and men were more likely to achieve a live birth. The same was seen for men and women with lower BMI and shorter infertility durations. However, these results were only statistically significant for women’s age, men’s age, AFC and AMH (p<0.05).

Table 2 shows that most women and men had never smoked. Most men and women were Caucasian and the male factor, which means male infertility, is the most prevalent infertility cause in this data. Most men and women had no previous live births (90.1% and 87.4%). Chi-square test revealed that no discrete characteristics have statistically significant differences for live birth (p>0.05). For this reason, interactions on different characteristics were considered.

In Table 3, the binary logistic regression parameters computed with SPSS are presented. These parameters allow the assessment
### Table 1. Baseline continuous characteristics of couples.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>All couples (n=739)</th>
<th>Couples with live birth (n=232)</th>
<th>Couples without live birth (n=507)</th>
<th>T-student test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± Standard deviation</td>
<td>Minimum value</td>
<td>Maximum value</td>
<td>Mean ± Standard deviation</td>
</tr>
<tr>
<td><strong>Woman’s age (years)</strong></td>
<td>34.04 ± 3.74</td>
<td>18</td>
<td>39</td>
<td>33.13 ± 3.80</td>
</tr>
<tr>
<td><strong>Man’s age (years)</strong></td>
<td>36.14 ± 5.18</td>
<td>20</td>
<td>54</td>
<td>35.42 ± 4.90</td>
</tr>
<tr>
<td><strong>Woman’s BMI (Kg/m²)</strong></td>
<td>24.35 ± 5.09</td>
<td>17</td>
<td>43</td>
<td>24.08 ± 4.21</td>
</tr>
<tr>
<td><strong>Man’s BMI (Kg/m²)</strong></td>
<td>26.43 ± 7.29</td>
<td>18</td>
<td>47</td>
<td>25.97 ± 3.87</td>
</tr>
<tr>
<td><strong>AFC (number of follicles)</strong></td>
<td>14.47 ± 9.24</td>
<td>2</td>
<td>52</td>
<td>17.03 ± 9.18</td>
</tr>
<tr>
<td><strong>AMH (ng/mL)</strong></td>
<td>3.50 ± 3.74</td>
<td>0.09</td>
<td>30.70</td>
<td>4.24 ± 3.81</td>
</tr>
<tr>
<td><strong>Infertility Duration (months)</strong></td>
<td>56.40 ± 28.68</td>
<td>10</td>
<td>180</td>
<td>55.17 ± 29.37</td>
</tr>
</tbody>
</table>

BMI – body mass index; AMH - anti-Müllerian hormone; AFC – antral follicle count
Table 2. Baseline discrete characteristics of couples.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>All couples (n=739)</th>
<th>Couples with live birth (n=232)</th>
<th>Couples without live birth (n=507)</th>
<th>Chi-square test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Woman's smoking status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smokes presently</td>
<td>200 (27.1%)</td>
<td>66 (28.4%)</td>
<td>134 (26.4%)</td>
<td>0.846</td>
</tr>
<tr>
<td>Smoked previously</td>
<td>118 (16.0%)</td>
<td>36 (15.5%)</td>
<td>82 (16.2%)</td>
<td></td>
</tr>
<tr>
<td>Never smoked</td>
<td>421 (57.0%)</td>
<td>130 (56%)</td>
<td>291 (57.4%)</td>
<td></td>
</tr>
<tr>
<td><strong>Woman's ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.722</td>
</tr>
<tr>
<td>Caucasian</td>
<td>662 (89.6%)</td>
<td>208 (89.7%)</td>
<td>454 (89.5%)</td>
<td></td>
</tr>
<tr>
<td>Asiatic</td>
<td>3 (0.4%)</td>
<td>1 (0.4%)</td>
<td>2 (0.4%)</td>
<td></td>
</tr>
<tr>
<td>Gipsy</td>
<td>11 (1.5%)</td>
<td>5 (2.2%)</td>
<td>6 (1.2%)</td>
<td></td>
</tr>
<tr>
<td>Indian</td>
<td>5 (0.7%)</td>
<td>2 (0.9%)</td>
<td>3 (0.6%)</td>
<td></td>
</tr>
<tr>
<td>Mixture</td>
<td>29 (3.9%)</td>
<td>10 (4.3%)</td>
<td>19 (3.7%)</td>
<td></td>
</tr>
<tr>
<td><strong>Woman's with previous children</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.432</td>
</tr>
<tr>
<td>Yes</td>
<td>73 (9.9%)</td>
<td>20 (8.6%)</td>
<td>53 (10.4%)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>666 (90.1%)</td>
<td>212 (91.4%)</td>
<td>451 (89.6%)</td>
<td></td>
</tr>
<tr>
<td><strong>Man's smoking status</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.379</td>
</tr>
<tr>
<td>Smokes presently</td>
<td>243 (32.9%)</td>
<td>82 (35.5%)</td>
<td>161 (31.8%)</td>
<td></td>
</tr>
<tr>
<td>Smoked previously</td>
<td>124 (16.8%)</td>
<td>33 (14.2%)</td>
<td>91 (17.9%)</td>
<td></td>
</tr>
<tr>
<td>Never smoked</td>
<td>372 (50.3%)</td>
<td>117 (50.4%)</td>
<td>255 (50.3%)</td>
<td></td>
</tr>
<tr>
<td><strong>Man's ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.070</td>
</tr>
<tr>
<td>Caucasian</td>
<td>678 (91.7%)</td>
<td>203 (87.5%)</td>
<td>475 (93.7%)</td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>21 (2.8%)</td>
<td>8 (3.4%)</td>
<td>13 (2.6%)</td>
<td></td>
</tr>
<tr>
<td>Asiatic</td>
<td>3 (0.4%)</td>
<td>1 (0.4%)</td>
<td>2 (0.4%)</td>
<td></td>
</tr>
<tr>
<td>Gipsy</td>
<td>12 (1.6%)</td>
<td>6 (2.6%)</td>
<td>6 (1.2%)</td>
<td></td>
</tr>
<tr>
<td>Indian</td>
<td>6 (0.8%)</td>
<td>3 (1.3%)</td>
<td>3 (0.6%)</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>19 (2.6%)</td>
<td>11 (4.8%)</td>
<td>8 (1.6%)</td>
<td></td>
</tr>
<tr>
<td><strong>Man's with previous live children</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.316</td>
</tr>
<tr>
<td>Yes</td>
<td>93 (12.6%)</td>
<td>25 (10.8%)</td>
<td>68 (13.4%)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>646 (87.4%)</td>
<td>207 (89.2%)</td>
<td>439 (86.6%)</td>
<td></td>
</tr>
<tr>
<td><strong>Infertility cause</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.482</td>
</tr>
<tr>
<td>Endometriosis</td>
<td>87 (11.8%)</td>
<td>19 (8%)</td>
<td>68 (13%)</td>
<td></td>
</tr>
<tr>
<td>Male factor</td>
<td>323 (43.7%)</td>
<td>86 (37%)</td>
<td>237 (47%)</td>
<td></td>
</tr>
<tr>
<td>Both female and male factor</td>
<td>83 (11.2%)</td>
<td>16 (16%)</td>
<td>67 (13%)</td>
<td></td>
</tr>
<tr>
<td>Unexplained infertility</td>
<td>216 (29.2%)</td>
<td>49 (21%)</td>
<td>167 (33%)</td>
<td></td>
</tr>
<tr>
<td>Multiple female factors</td>
<td>29 (3.9%)</td>
<td>6 (3%)</td>
<td>23 (5%)</td>
<td></td>
</tr>
<tr>
<td>Disovulation</td>
<td>78 (10.6%)</td>
<td>21 (9%)</td>
<td>57 (11%)</td>
<td></td>
</tr>
<tr>
<td>Tubal</td>
<td>147 (19.9%)</td>
<td>33 (14%)</td>
<td>114 (22%)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>8 (1.1%)</td>
<td>2 (1%)</td>
<td>6 (1%)</td>
<td></td>
</tr>
</tbody>
</table>

The logistic regression results showed that from the 14 variables evaluated on the pre-treatment procedures, only male factor, the man’s BMI, mixed ethnicity for men, and the level of AMH were statistically significant. Apart from these variables, the interactions between infertility duration and women’s age, infertility duration and men’s BMI, AFC and AMH, AFC and women’s age, AFC and women’s BMI and AFC and disovulation were also statistically significant (p-value less than 0.05). The interactions between infertility duration and woman’s age and man’s BMI; AFC and AMH, woman’s age, woman’s BMI and disovulation were also statistically significant.

According to regression coefficients of the final model, an AMH unit increase raises the probability of IVF-ICSI success by 0.172 times, ceteris paribus. Similarly, the interaction between

of the chance of living birth on couples before they start their IVF non-donor cycle based on the variables previously mentioned.
Table 3. Final logistic regression model for live birth (n=737).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Regression coefficient</th>
<th>Standard Error</th>
<th>P-value</th>
<th>Odds ratio</th>
<th>95% I.C. Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male factor</td>
<td>0.445</td>
<td>0.180</td>
<td>0.013</td>
<td>1.561</td>
<td>1.097–2.222</td>
</tr>
<tr>
<td>Man’s BMI</td>
<td>-0.155</td>
<td>0.034</td>
<td>&lt;0.001</td>
<td>0.856</td>
<td>0.801–0.916</td>
</tr>
<tr>
<td>Mixed ethnicity on Man</td>
<td>1.330</td>
<td>0.496</td>
<td>0.007</td>
<td>3.781</td>
<td>1.430–9.996</td>
</tr>
<tr>
<td>Interaction between Infertility duration and Woman’s age</td>
<td>-0.002</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.998</td>
<td>0.997–0.999</td>
</tr>
<tr>
<td>Interaction between Infertility duration and Man’s BMI</td>
<td>0.002</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>1.002</td>
<td>1.001–1.003</td>
</tr>
<tr>
<td>Interaction between AFC and Woman’s age</td>
<td>0.004</td>
<td>0.001</td>
<td>&lt;0.001</td>
<td>1.004</td>
<td>1.002–1.006</td>
</tr>
<tr>
<td>Interaction between AFC and Disovulation</td>
<td>0.043</td>
<td>0.016</td>
<td>0.006</td>
<td>1.044</td>
<td>1.013–1.076</td>
</tr>
<tr>
<td>Interaction between AFC and Woman’s BMI</td>
<td>-0.003</td>
<td>0.001</td>
<td>0.007</td>
<td>0.997</td>
<td>0.995–0.999</td>
</tr>
<tr>
<td>AMH</td>
<td>0.172</td>
<td>0.052</td>
<td>0.001</td>
<td>1.188</td>
<td>1.073–1.315</td>
</tr>
<tr>
<td>Interaction between AFC and AMH</td>
<td>-0.008</td>
<td>0.002</td>
<td>&lt;0.001</td>
<td>0.992</td>
<td>0.989–0.996</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.934</td>
<td>0.876</td>
<td>0.027</td>
<td>6.914</td>
<td>-</td>
</tr>
</tbody>
</table>

BMI – body mass index; AMH - anti-Müllerian hormone; AFC – antral follicle count

Figure 1. Receiver operating characteristic (ROC) curve of the final model. This shows the ROC curve test for the discriminatory ability of the final prediction model is 0.700 (95% confidence interval (CI) 0.660–0.741).
AFC and the woman’s BMI decreases the same probability by 0.003 times and interaction between AFC and the woman’s age increases it by 0.004 times. Male factor is the only infertility cause that enters individually into the final model raising the chances of success by 0.420 times.

Discussion

So far, providing an accurate prediction of the chances of achieving a live birth after FIV-ICSI treatments has not been an easy task[39]. In this study, a novel prediction model was built, including almost all clinical factors reported as important in the literature.

To our knowledge, this is the first model for live birth FIV-ICSI prediction that accounts for men’s ethnicity. It also includes variables such as BMI and AFC, which were not included in many models before34. This model includes all the characteristics that have been indicated as important in the literature to predict the chances of live birth in MAR[5,38,39].

The Templeton model10 considers the existence of previous IVF cycles and has been externally validated by Nelson and Lawlor36. Therefore, both these models can be applied before IVF is started and predict the success of IVF/ICSI treatment. During this study, it was not possible to obtain information about the number of previous cycles per couple, and so this factor was not included.

This model supports earlier findings in this scientific area such as the importance of AMH and AFC to predict live birth as predicted by other studies35,40 and indirectly corroborates the importance of woman’s age to predict the result as shown by Templeton et al.30. Nonetheless, it is crucial to refer that the increase of women’s age, when in interaction with infertility duration, significantly decreases the probability of a living birth which is in accordance to Nelson’s study36.

Concerning the main reason for treatment, both male factor and disovulation (with AFC interaction) are the only causes that emerged in the final model. These causes raised the chances of success, which might be explained with the IVF-ICSI technique itself once the problem of sperm and ovulation anomalies are overcome38. In other words, women with ovulation problems and men with sperm anomalies have higher chances of success with IVF-ICSI technique because sperm and oocytes are medically collected and interact directly, although in an in vitro environment31.

Another notable finding is that an increase of the man’s BMI (per si) or woman’s BMI (with AFC interaction) decreased the chances of live birth, which is in accordance with the scientific literature. Women who are overweight are known to have ovulatory problems, and increased risks of miscarriage39 and obesity may adversely affect male reproduction by endocrine, thermal, genetic, and sexual mechanisms30. For instance, increased testicular temperature can result from prolonged periods of sitting in men with excessive lower abdominal fat deposited in the lower abdomen59. Also, obesity tends to increase estrogen levels and reduce testosterone levels in men59.

Concerning the man’s ethnicity, we found that if the man is of mixed ethnicity, the odds of success are around 3.8. This finding must be interpreted cautiously as the sample is composed of 92% Caucasian males, with only 2.6% of males being of mixed ethnicity. Up to now, there doesn’t seem to be any biological plausibility for this find, further researchers is required to validate this possible relationship between mixed ethnicity in men and increased success in IVF treatments.

A limitation of our model is that it is restricted to pre-treatment stages. Thereby, the physician must explain to patients when using our model that their success probability invariably changes during the cycles process. The resulting probability of our model should be considered a baseline one.

Many studies have accounted for the chances of live birth after IVF or ICSI treatment35,38–39. When compared with other studies, the most similar one is Dhillon et al.’s model34 due to the variables integrated into the model. However, in our study, there were no limitations on socioeconomic status since our data has been extracted from a public hospital with universal access30 and so the results can be generalized to people of all socioeconomic backgrounds. This is an advantage over the Dhillon et al.’s model34, where it is estimated that 75% of couples paid for their treatment and therefore their model could not be generalized to all social classes.

Predictive ability of the prediction models on medicine has been assessed by the AUC34,35. In general, AUC for prediction models in reproductive medicine is rather low, ranging between 0.59 and 0.6435. Table 4 shows the values of models’ AUC predicting the chances of live birth for couples undergoing IVF-ICSI. Although McLernon et al.37 have the highest value of the area under the receiver operating characteristic (AUROC) curve, that value decreased on Leijedekkers validation for 0.6235. The model developed in this study has an AUC of 0.700, which is the second-highest value, and so has a comparable discriminatory ability with these previous models.

Taking into account that Coppus et al.’s systematic review concluded that prediction models in reproductive medicine would be limited to an AUC value of 0.65 due to the relatively

---

**Table 4. Comparison of area under the receiver operating characteristic (AUROC) curve of existing models.**

<table>
<thead>
<tr>
<th>Model</th>
<th>AUROC curve of the model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Templeton et al.30</td>
<td>0.62</td>
</tr>
<tr>
<td>Nelson and Lawlor36</td>
<td>0.63</td>
</tr>
<tr>
<td>La Marca et al.36</td>
<td>0.66</td>
</tr>
<tr>
<td>Mc Lernon et al.37</td>
<td>0.73</td>
</tr>
<tr>
<td>Dhillon et al.34</td>
<td>0.62</td>
</tr>
<tr>
<td>This study</td>
<td>0.70</td>
</tr>
</tbody>
</table>
homogeneous group of subfertile patients, this study can be considered to improve upon the current available models.

The present model predicts the specific probability of a live birth based on easily acquirable couple characteristics before starting a treatment. It might help patients to understand the limitations of an IVF/ICSI in their particular case and also physicians to compare different treatment strategies. Furthermore, it might support institutions to predict the probability of the need for treatment repetition based on the specific characteristics of each couple.

We intend in the near future to perform an external validation with the developed model with a new dataset from Centro de Infertilidade e Reprodução Medicamente Assistida (CIRMA). Follow this, it is also planned to perform an external geographical validation to check if there is a geographical influence on chances of live birth for couples undergoing IVF-ICSI.

Conclusions

Many couples with fertility problems ask themselves if they should undergo an IVF-ICSI treatment. Our novel model provides an estimated probability of their chances of live birth. This is the first model with Portuguese data and takes in account the important variables described in literature such as AFC, AMH, and woman’s age. This tool may help physicians to shape couples’ expectations conceding them the opportunity to plan their treatments and to prepare both emotionally and financially to them and so we are developing a user-friendly interface to help physicians in their clinical practice.

Data availability

Underlying data

The data that support the findings of this study are available on request from the author José Metello (jose.metello@hgo.min-saude.pt). The data cannot be made publicly available in accordance to paragraph d), number 2, Article 9 of the Regulation no. 2016/679 of the European Parliament and the Council of 27 April 2016 and due to the sensitive data containing information that could compromise the privacy of research participants. Informed consent was provided by patients who participated in the study for the use of their data for scientific purposes only and to safeguard their anonymity.

References

Reference Sources


Open Peer Review

Current Peer Review Status:  ❌  ❌

Version 1

Reviewer Report 07 September 2020

https://doi.org/10.5256/f1000research.21999.r69512

© 2020 Tan J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Jichun Tan
Center of Reproductive Medicine, Department of Obstetrics and Gynecology, Shengjing Hospital of China Medical University, Shenyang, China

This paper focused on an interesting topic of "What are my chances of conceiving?". However, the aim of this study is not clear. The logistic regression model is not enough to establish a novel prediction model. To the best of our knowledge, similar studies with larger sample size, more clinical variables, and advanced calculating models have been reported to predict the rates of live birth. In addition, more latest studies should be added in the introduction section.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Assisted reproduction
I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Reviewer Report 31 July 2020

https://doi.org/10.5256/f1000research.21999.r68000

© 2020 Siristatidis C. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Charalampos Siristatidis
Assisted Reproduction Unit, Third Department of Obstetrics and Gynecology, Medical School, National and Kapodistrian University of Athens, "Attikon" University Hospital, Athens, Greece

Authors developed a logistic regression model based on the retrospective study of 737 IVF cycles. They concluded that the model might result in a new validated decision support system to help physicians to manage couples’ expectations concerning their possibility to have a live birth.

The introduction section is unreasonably extended, without providing the necessary information only, leading to the rationale of the study. Important similar models using more sophisticated networks, such as artificial intelligence, are missing.

Language and grammar need revision. The sample size is small. A proper power calculation is missing. Some important parameters that have been tested in other studies that contribute to live birth are missing. Limitations are not reported.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly
**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Assisted reproduction; methodology; artificial intelligence

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com