RESEARCH ARTICLE

REVISED Predicting the chances of live birth for couples undergoing IVF-ICSI: a novel instrument to advise patients and physicians before treatment

[version 2; peer review: 1 approved with reservations, 2 not approved]

Bruna Estácio da Veiga[1,2], Beatriz Brás de Guimarães[2], Duarte Pedro Tavares[1], José Luis Metello[3], Fernando Ferreira[4], Pedro Ferreira[3], José Manuel Fonseca[1]

[1]CA3 - CTS UNINOVA, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, 2829-516, Portugal
[2]Physics Department, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, Caparica, 2829-516, Portugal
[3]CIRMA, Hospital Garcia of Orta, Almada, 2805-267, Portugal
[4]CTS UNINOVA, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, 2829-516, Portugal

v2 **First published:** 04 Sep 2019, **8**:1585
https://doi.org/10.12688/f1000research.20038.1
Latest published: 22 Dec 2020, **8**:1585
https://doi.org/10.12688/f1000research.20038.2

Abstract

Background: The prevalence of infertility ranges from 3.5% to 16.7% in more developed countries. For this reason, the number of *In Vitro* Fertilization(IVF) technique and Intracytoplasmic Sperm Injection (ICSI) treatments has been significantly increasing. 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, expensive and with a variable probability of success, the most common question asked by IVF patients is "What are my chances of conceiving before starting an IVF/ICSI treatment?". The main aim of this study is to develop a validated model that estimates the chance of a live birth before the start of an IVF/ICSI non-donor cycle.

Methods: A logistic regression model was developed based on the retrospective study of 737 IVF/ICSI cycles. Overall 14 pre-treatment variables were evaluated (*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 the outcome of the treatment was discriminated as "live birth" or "no live birth".

Results: From the 14 variables acquired before starting the IVF/ICSI 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

Open Peer Review

Approval Status ✗ ✗ ?

	1	2	3
version 2 (revision) 22 Dec 2020			? view
version 1 04 Sep 2019	✗ view	✗ view	

1. **Charalampos Siristatidis**, Third Department of Obstetrics and Gynecology, Medical School, National and Kapodistrian University of Athens, "Attikon" University Hospital, Athens, Greece

2. **Jichun Tan**, Shengjing Hospital of China Medical University, Shenyang, China

3. **Edwin Amalraj Raja**, University of Aberdeen, Aberdeen, UK

Any reports and responses or comments on the article can be found at the end of the article.

the Receiver Operating Characteristic (AUROC) curve test for the discriminatory ability of the final prediction model was 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' pre-treatment expectations.

Keywords
IVF, ICSI, prediction model, live birth, assisted reproduction, logistic regression

Corresponding author: Bruna Estácio da Veiga (bf.veiga@campus.fct.unl.pt)

Author roles: Estácio da Veiga B: Conceptualization, Formal Analysis, Investigation, Methodology, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Brás de Guimarães B**: Writing – Review & Editing; **Tavares DP**: Formal Analysis, Investigation, Methodology, Supervision, Writing – Review & Editing; **Metello JL**: Conceptualization, Data Curation, Supervision, Validation, Writing – Review & Editing; **Ferreira F**: Conceptualization, Supervision, Writing – Review & Editing; **Ferreira P**: Conceptualization, Data Curation; **Fonseca JM**: Conceptualization, Project Administration, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: We thank Portuguese Foundation for Science and Technology FCT/ MCTES - UID/EEA/00066/2019 (JF).

How to cite this article: Estácio da Veiga B, Brás de Guimarães B, Tavares DP *et al.* **Predicting the chances of live birth for couples undergoing IVF-ICSI: a novel instrument to advise patients and physicians before treatment [version 2; peer review: 1 approved with reservations, 2 not approved]** F1000Research 2020, **8**:1585 https://doi.org/10.12688/f1000research.20038.2

First published: 04 Sep 2019, **8**:1585 https://doi.org/10.12688/f1000research.20038.1

REVISED **Amendments from Version 1**

We submitted an original research article that has been reviewed by two experts in the field. We want to thank their comments and we have done some revisions in a new version. The entire text was revised (which includes language and grammar revision) to make the article more conspicuous and clearer. Therefore, the abstract was also restructured. The author list was updated by adding Beatriz Brás de Guimarães, who contributed in Writing – Review & Editing. The affiliation of each author was updated and added the e-mail address for each one. The introduction section was modified in order to provide the necessary information, leading to the aim of the study, and enriched with a more recent study based on an AI predictive model. At the same time, the main aim of the study was clarified and reinforced. We highlighted the variables analyzed using italics (in the main text and in Table 1, Table 2 and Table 3) and the data of Table 2 was rectified. Considering the size of the sample available and the reviewer's assessment of the model's evaluation, it was calculated 5 metrics (i.e., accuracy, F score, precision, sensitivity and sensibility) that were listed in a new table (Table 4), reflecting our model performance. This new table was cited in the main text (i.e., results section). The discussion section was also reformulated by reporting more limitations of our model, namely about the large temporal spectrum of the data (2012 to 2016) and on the fact that only an internal validation was made. Finally, due to all the manuscript was verified, some references were excluded and the bibliography undergone some modifications.

Any further responses from the reviewers can be found at the end of the article

Introduction

Infertility is 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[1].

The prevalence of infertility ranged from 3.5% to 16.7% in developed countries[2]. Most recent portuguese data estimate that 9.8% of couples are infertile[3].

Infertility is a multifactorial disease[4]. Several treatments have been proposed, with the most innovative being IVF developed by Robert Edwards[5] in 1978. The two most important Medically Assistant Reproduction Techniques (MAR) are IVF and its subtype ICSI[6]. According to the European Society of Human Reproduction and Embryology (ESHRE), the number of MAR cycles between 1997 and 2014 increased by 13%, reaching 776 556 cycles in Europe in 2015[7].

Despite the increasing number of MAR treatments, IVF success is not guaranteed[8]. In healthy young couples, the probability of achieving a live birth varies between 20% and 25% per month. This likelihood may increase up to 60% with MAR techniques[9]. According to the study by Malizia *et al.*, between 38% and 49% of couples who start MAR treatments remain childless, even after undergoing up to six MAR cycles[10]. In Portugal, the last report of MAR showed that the treatments success rate is around 25–30% per started cycle[11].

MAR treatments are expensive, time-consuming, stressful, and may lead to anxiety, depression, or marital problems[12–16]. There are also complications to take into consideration such as ovarian hyperstimulation syndrome, bleeding and infection, as well as multiple or premature births[17].

The success rates fluctuate between studies[18]. There have been various efforts to build prediction models to assist physicians and patients in predicting MAR success[18–22]. To our knowledge, the first predictive model ever built in this context is from Templeton *et al.* in 1996. A logistic regression model to predict the probability of live birth for an individual woman was developed 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[18].

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[18].

Considering the evolution of technology and the inclusion of other variables, Nelson and Lawlor developed a new model in 2011, using the same mathematical techniques. They included 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)[21].

In 2014, Velde *et al.* used their cohort study to validate Templeton and Nelson and Lawlor's model's performance. They found that Templeton's model underestimated success rates while Nelson's overestimated it[23].

Other relevant models were the one developed by Marca *et al.,* which predicts live birth in assisted reproduction based on serum AMH and woman's age[20], and the McLernon *et al.* study[22] that estimated the cumulative first live birth over a maximum of 6 complete cycles using data from 23 417 women in the UK.

Nonetheless, when Leijdekkers performed an external validation of McLernon's model[24] with dutch women, he found that there was an overestimation of the results and he decided to include biomarkers such as AMH, AFC and woman's body weight. Other studies indicate it is important to include covariables such as BMI, ethnicity and ovarian reserve[19], and corroborate that the use of variables such as BMI, AFC, AMH, ethnicity and smoking status should be predictors in the model[4,25–27].

Furthermore, a machine learning approach was also proposed, including decision trees, genetic algorithms, and k-nearest neighbors classifiers[9,28–31]. In 2019, in the study of Qiu *et al.*[32],

4 predictive models based on supervised learning algorithms were created. The purpose of these models was to estimate the cumulative probability of having a live birth before the first IVF treatment starts, using pre-treatment variables including BMI and AMH. One of the models was a logistic regression and achieved an AUROC of 0.71.

The main aim of this study was to find a pre-treatment predictor for achieving a live birth before an IVF/ICSI treatment.

Methods

This was a retrospective study of 739 IVF/ICSI cycles, performed between 2012 and 2016, in Centro de Infertilidade e Reprodução Medicamente Assistida (CIRMA) at Hospital Garcia de Orta, E.P.E., Almada, Portugal.

Only couples with autologous gametes with a live birth or without any frozen embryos available were considered.

The study was approved by the Hospital's Ethics Committee for Health of the institution. Patients were informed that, under complete anonymity, the data may be used in scientific papers for public presentation or publication[33].

The model's dependent variable/primary outcome was classified as 1 (if, at least, one baby was born) and 0 (otherwise). This decision was based on other studies in this area[18–22].

The following variables were analyzed: *woman's* and *man's age* (years), *duration of infertility* (months), *cause of infertility* (categorized 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 Body Mass Index* (*BMI*) (kg/m^2), *Anti-Müllerian Hormone* (AMH) (ng/mL), *Antral Follicle Count* (AFC), *woman's* and *man's ethnicity* (categorized as African, Asian, Caucasian, Gipsy, Indian and Mixture), *woman's* and *man's smoking status* (never, previous and present) and *woman's* and *man's previous live children* (yes or no) were considered. AFC was obtained by transvaginal ultrasound, and serum AMH 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[34].

A binary logistic regression[35] was developed using the primary outcome as binary (no = 0 and yes = 1). An automatic backward selection process based on the Wald statistic was used to determine the final and the best logistic regression model[36].

As recommended by Hosmer and Lemeshow[35], the interactions between some variables were included to increase the reliability of the model. As Fisher[37] explained, the primary outcome can depend not only on one individual baseline characteristic, but also on the relationship between them. The variables in interaction were: *woman's age* and *duration of infertility*[38], *BMI* and *oligospermia*[39], *AFC* and *woman's BMI*[40], and *AFC* and *AMH*[26].

The model's performance was measured with the discriminatory power assessed using the Area Under the Curve (AUC) value[41,42] and the model was internally validated using the bootstrapping technique with 1000 iterations[43]. All statistical procedures were computed using IBM SPSS Statistics 25, and it was considered an α level of 0.05.

The STROBE cross-sectional reporting guidelines were adopted in this article[44].

Results

A total of 737 cycles were evaluated. Overall 31.4% of the cycles had at least one live birth.

Baseline characteristics of couples are presented in Table 1 and Table 2. The average age for females was 34.04 years, and 36.14 years for males. Younger women and men were more likely to achieve a live birth. The same was verified for men and women with lower BMI and shorter infertility duration. However, these results were only statistically significant for woman's age, man'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 male factor was the most prevalent infertility cause. Most men and women had no previous live births (90.1% and 87.4%). Chi-square test revealed that no statistically significant differences for live birth ($p > 0.05$) for categorical variables.

Table 3 shows the binary logistic regression parameters computed with SPSS. It shows that from the 14 variables evaluated, only *cause of infertility* (male factor), *man's BMI*, *man's ethnicity* (mixture) and *AMH* were statistically significant. Apart from these variables, the interactions between *duration of infertility* and *woman's age*, *duration of infertility* and *man's BMI*, *AFC* and *AMH*, *AFC* and *woman's age*, *AFC* and *woman's BMI*, and *AFC* and *cause of infertility* (disovulation) were also statistically significant ($p < 0.05$). The interactions between *infertility duration* and *woman's age* and *man's BMI*, *AFC* and *AMH*, *woman's age*, *woman's BMI* and *cause of infertility* (disovulation) were also statistically significant (*p*-value < 0.05).

According to regression coefficients of the final model, an increase of an AMH unit raises the odds of IVF-ICSI success by 0.172 times, *ceteris paribus*. Similarly, the interaction between AFC and woman's BMI decreases the same probability by 0.003 times, and interaction between AFC and woman's age increases

Table 1. Baseline continuous characteristics of couples.

Characteristics	All couples (n=739)			Couples with live birth (n=232)			Couples without live birth (n=507)			*p*-value
	Mean ± Standard deviation	Minimum value	Maximum value	Mean ± Standard deviation	Minimum value	Maximum value	Mean ± Standard deviation	Minimum value	Maximum value	
***Woman's age* (years)**	34.04 ± 3.74	18	39	33.13 ± 3.80	18	39	34.45 ± 3.64	19	39	0.000
***Man's age* (years)**	36.14 ± 5.18	20	54	35.42 ± 4.90	21	52	36.47 ± 5.27	20	54	0.008
***Woman's BMI* (kg/m²)**	24.35 ± 5.09	17	43	24.08 ± 4.21	17	41	24.47 ± 5.44	17	43	0.286
***Man's BMI* (kg/m²)**	26.43 ± 7.29	18	47	25.97 ± 3.87	18	47	26.28 ± 3.85	19	47	0.619
AFC	14.47 ± 9.24	2	52	17.03 ± 9.18	4	50	13.30 ± 9.03	2	52	0.000
***AMH* (ng/mL)**	3.50 ± 3.74	0.09	30.70	4.24 ± 3.81	0.2	20	3.17 ± 3.66	0.09	30.7	0.000
Duration of infertility (months)	56.40 ± 28.68	10	180	55.17 ± 29.37	10	176	56.96 ± 28.34	12	180	0.437

BMI – Body Mass Index; AMH - Anti-Müllerian Hormone; AFC – Antral Follicle Count

F1000Research 2020, 8:1585 Last updated: 30 MAY 2025

Table 2. Baseline discrete characteristics of couples.

Characteristics		All couples (n=739)	Couples with live birth (n=232)	Couples without live birth (n=507)	*p*-value
Woman's smoking status	Present	200 (27.1%)	66 (28.4%)	134 (26.4%)	0.846
	Previous	118 (16.0%)	36 (15.5%)	82 (16.2%)	
	Never	421 (57.0%)	130 (56.0%)	291 (57.4%)	
Woman's ethnicity	Caucasian	662 (89.6%)	208 (89.7%)	454 (89.5%)	0.722
	African	29 (3.9%)	6 (2.6%)	23 (4.6%)	
	Asiatic	3 (0.4%)	1 (0.4%)	2 (0.4%)	
	Gipsy	11 (1.5%)	5 (2.2%)	6 (1.2%)	
	Indian	5 (0.7%)	2 (0.9%)	3 (0.6%)	
	Mixture	29 (3.9%)	10 (4.3%)	19 (3.7%)	
Woman's with previous children	Yes	73 (9.9%)	20 (8.6%)	53 (10.5%)	0.432
	No	666 (90.1%)	212 (91.4%)	454 (89.6%)	
Man's smoking status	Present	243 (32.9%)	82 (35.3%)	161 (31.8%)	0.379
	Previous	124 (16.8%)	33 (14.2%)	91 (17.9%)	
	Never	372 (50.3%)	117 (50.4%)	255 (50.3%)	
Man's ethnicity	Caucasian	678 (91.7%)	203 (87.5%)	475 (93.7%)	0.070
	African	21 (2.8%)	8 (3.4%)	13 (2.6%)	
	Asiatic	3 (0.4%)	1 (0.4%)	2 (0.4%)	
	Gipsy	12 (1.6%)	6 (2.6%)	6 (1.2%)	
	Indian	6 (0.8%)	3 (1.3%)	3 (0.6%)	
	Mixture	19 (2.6%)	11 (4.8%)	8 (1.6%)	
Man's with previous live children	Yes	93 (12.6%)	25 (10.8%)	68 (13.4%)	0.316
	No	646 (87.4%)	207 (89.2%)	439 (86.6%)	
Cause of infertility	Endometriosis	68 (9.2%)	19 (8.2%)	49 (9.7%)	0.482
	Male factor	237 (32.1%)	86 (37.1%)	151 (29.8%)	
	Both female and male factor	67 (9.1%)	16 (6.9%)	51 (10.1%)	
	Unexplained infertility	167 (22.6%)	49 (21.1%)	118 (23.3%)	
	Multiple female factors	23 (3.1%)	6 (2.6%)	17 (3.4%)	
	Disovulation	57 (7.7%)	21 (9.1%)	36 (7.1%)	
	Tubal	114 (15.4%)	33 (14.2%)	81 (16.0%)	
	Other	6 (0.8%)	2 (0.9%)	4 (0.8%)	

it by 0.004 times. Male factor is the only cause of infertility that enters individually into the final model raising the chances of success by 0.420 times.

The ROC curve test for the discriminatory ability of the final prediction model (Figure 1) had an AUC equal to 0.700 (95% CI 0.660–0.741). This model was also assessed by its performance metrics, shown in Table 4.

Discussion

So far, providing an accurate prediction of the chances of achieving a live birth after FIV-ISCI treatments has not been

Table 3. **Final logistic regression model for live birth (n=737).**

Variables	Regression coefficient	Standard Error	*p*-value	Odds ratio	95% CI Odds ratio	
					inf	sup
Cause of infertility **(male factor)**	0.445	0.180	0.013	1.561	1.097	2.222
Man's BMI	-0.155	0.034	<0.001	0.856	0.801	0.916
Man's ethnicity **(mixture)**	1.330	0.496	0.007	3.781	1.430	9.996
Interaction between *duration of infertility* and *woman's age*	-0.002	<0.001	<0.001	0.998	0.997	0.999
Interaction between *duration of infertility* and *man's BMI*	0.002	<0.001	<0.001	1.002	1.001	1.003
Interaction between *AFC* and *woman's age*	0.004	0.001	<0.001	1.004	1.002	1.006
Interaction between *AFC* and *cause of* infertility (disovulation)	0.043	0.016	0.006	1.044	1.013	1.076
Interaction between *AFC* and *woman's BMI*	-0.003	0.001	0.007	0.997	0.995	0.999
AMH	0.172	0.052	0.001	1.188	1.073	1.315
Interaction between *AFC* and *AMH*	-0.008	0.002	<0.001	0.992	0.989	0.996
Intercept	1.934	0.876	0.027	6.914	-	-

BMI – Body Mass Index; AMH - Anti-Müllerian Hormone; AFC – Antral Follicle Count

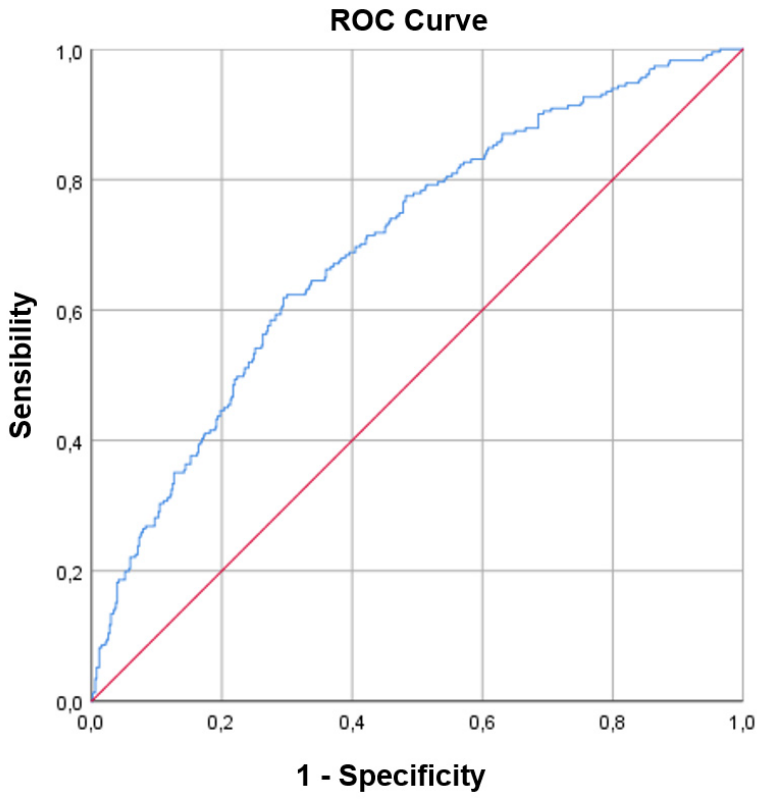

Figure 1. **Receiver Operating Characteristic (ROC) curve of the final model.**

Table 4. Logistic regression model performance.

Performance metric	Value
Accuracy	0.711
F-score	0.324
Precision	0.607
Sensitivity	0.221
Specificity	0.935

an easy task[45]. 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 man's ethnicity. It also includes variables such as *BMI* and *AFC*, which were not included in many models before[19]. This model includes all the characteristics that have been indicated as important in the literature to predict the chances of live birth in MAR[4,18,45].

The Templeton model[18] considers the existence of previous IVF cycles and has been externally validated by Nelson and Lawlor[21]. Therefore, both 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, 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 studies[20,26] and corroborates indirectly the importance of woman's age to predict the result as shown by Templeton *et al.*[18]. Nonetheless, it is crucial to refer that the increase of woman's age, when in interaction with duration of infertility, significantly decreases the probability of a living birth, which is in accordance to Nelson's study[21].

Concerning the main reason for treatment, both male factor and disovulation (with AFC interaction) are the only causes that emerged in the final model. When present, each one of them raises the chances of success, which might be explained with the IVF-ICSI technique itself once the problem of sperm and ovulation anomalies are overcome. Possibly, 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* environment[6].

Another notable finding is that an increase of 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 great risks of miscarriage, in the same way that obesity may adversely affect male reproduction by endocrine, thermal, genetic, and sexual mechanisms[46].

Concerning the man's ethnicity, we found that if the man is of mixture 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 mixture 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 mixture ethnicity in men and increased success in IVF/ICSI 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. That is, the resulting probability of our model should be considered a baseline one. Taking into account that this study was performed with data from only one medical center, one limitation of these models is that only an internal validation was made. Moreover, the large temporal spectrum of the data (2012 to 2016) could mean that treatments made with older technology may have different results than those made with more recent ones (namely, vitrification procedures or culture media).

Many studies have accounted for the chances of live birth after IVF or ICSI treatment[18–24]. When compared with other studies, the most similar one is Dhillon *et al.*'s model[19] 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 access[47] and so the results can be generalized to people of all socioeconomic backgrounds. This is an advantage over the Dhillon *et al.*'s model[19], 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 in medicine has been assessed by the AUC value[41,42]. In general, AUC for prediction models in reproductive medicine is rather low, ranging between 0.59 and 0.64[42]. Table 5 shows the values of models' AUC predicting the chances of live birth for couples undergoing IVF-ICSI. Although McLernon *et al.*[22] had the highest value of the AUROC curve, that value decreased on Leijedekkers

Table 5. Comparison of AUROC curve value of existing models.

Model	AUROC curve value of the model
Templeton *et al.*[18]	0.62
Nelson and Lawlor[21]	0.63
La Marca *et al.*[20]	0.66
McLernon *et al.*[22]	0.73
Dhillon *et al.*[19]	0.62
This study	0.70

validation to 0.62[24]. The model developed in this study had 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 heterogeneous group of subfertile patients[42], 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 treatment in their particular case and also aid physicians to compare different treatment strategies. Furthermore, it might support institutions to predict the likelihood of repeat treatment based on the specific characteristics of each couple.

In the near future, we intend to perform an external validation of the developed model with a new dataset from 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 treatment.

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 before the start of a MAR treatment. This is the first model with portuguese data and considers 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. Hence 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 co-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

1. Zegers-Hochschild F, Adamson GD, Dyer S, *et al.*: **The international glossary on infertility and fertility care, 2017.** *Hum Reprod.* 2017; **32**(9): 1786–1801.
PubMed Abstract | Publisher Full Text | Free Full Text
2. Boivin J, Bunting L, Collins JA, *et al.*: **International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care.** *Hum Reprod.* 2007; **22**(6): 1506–1512.
PubMed Abstract | Publisher Full Text
3. Silva-Carvalho JL, Santos A: **Estudo Afrodite: Caracterização da infertilidade em Portugal (Vol. 1. Estudo na Comunidade).** *Porto Fac Med da Univ do.* 2009; 74.
Reference Source
4. Vander Borght M, Wyns C: **Fertility and infertility: Definition and epidemiology.** *Clin Biochem.* 2018; **62**: 2–10.
PubMed Abstract | Publisher Full Text
5. Johnson MH: **Robert Edwards: the path to IVF.** *Reprod Biomed Online.* 2006; **12**(2): 234–253.
6. Chow S, Cheung S, Parrella A, *et al.*: **Intracytoplasmic sperm injection: state of the art in humans.** *Reproduction.* 2004; 1741–7899.
7. Nyboe Andersen A, Goossens V, Bhattacharya S, *et al.*: **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; **24**(6): 1267–1287.
PubMed Abstract | Publisher Full Text
8. De Geyter C, Calhaz-Jorge C, Kupka MS, *et al.*: **ART in Europe, 2014: results generated from European registries by ESHRE: The European IVF-monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE).** *Hum Reprod.* 2018; **33**(9): 1586–1601.
PubMed Abstract | Publisher Full Text
9. Pegah Hafiz BN, Nematollahi M, Boostani R: **Predicting implantation outcome of *in vitro* fertilization and intracytoplasmic sperm injection using data mining techniques.** *Int J Fertil Steril.* 2017; **11**(3): 184–190.
PubMed Abstract | Publisher Full Text | Free Full Text
10. Malizia BA, Hacker MR, Penzias AS: **Cumulative live-birth rates after *in vitro* fertilization.** *Obstet Gynecol Surv.* 2009; **64**(5): 315–316.
Publisher Full Text
11. Conselho Nacional de Procriação Medicamente Assistida: **Relatório da Atividade Desenvolvida pelos Centros de PMA em 2015**. 2015; **27**: 2015.
12. **Associação Portuguesa de Fertilidade**. [Accessed: 24-Jan-2019].
Reference Source
13. Karaca N, Karabulut A, Ozkan S, *et al.*: **Effect of IVF failure on quality of life and emotional status in infertile couples.** *Eur J Obstet Gynecol Reprod Biol.* 2016; **206**: 158–163.
PubMed Abstract | Publisher Full Text
14. Maroufizadeh S, Karimi E, Vesali S, *et al.*: **Anxiety and depression after failure of assisted reproductive treatment among patients experiencing infertility.** *Int J Gynecol Obstet.* 2015; **130**(3): 253–256.
PubMed Abstract | Publisher Full Text
15. Valoriani V, Lotti F, Lari D, *et al.*: **Differences in psychophysical well-being and signs of depression in couples undergoing their first consultation for assisted reproduction technology (ART): an Italian pilot study.** *Eur J Obstet Gynecol Reprod Biol.* 2016; **197**: 179–185.
PubMed Abstract | Publisher Full Text
16. Heredia M, Tenías JM, Rocio R, *et al.*: **Quality of life and predictive factors in patients undergoing assisted reproduction techniques.** *Eur J Obstet Gynecol Reprod Biol.* 2013; **167**(2): 176–180.
PubMed Abstract | Publisher Full Text
17. Paulson R: ***in vitro* Fertilization**. *UpToDate*. 2018; 379–386.
Reference Source
18. Templeton A, Morris JK, Parslow W: **Factors that affect outcome of *in-vitro* fertilisation treatment.** *Lancet.* 1996; **348**(9039): 1402–1406.
PubMed Abstract | Publisher Full Text
19. Dhillon RK, McLernon DJ, Smith PP, *et al.*: **Predicting the chance of live birth for women undergoing IVF: a novel pretreatment counselling tool.** *Hum Reprod.* 2016; **31**(1): 84–92.
PubMed Abstract | Publisher Full Text
20. La Marca A, Nelson SM, Sighinolfi G, *et al.*: **Anti-Müllerian hormone-based prediction model for a live birth in assisted reproduction.** *Reprod Biomed Online.* 2011; **22**(4): 341–349.
PubMed Abstract | Publisher Full Text
21. Nelson SM, Lawlor DA: **Predicting live birth, preterm delivery, and low birth**

weight in infants born from *in vitro* fertilisation: a prospective study of 144,018 treatment cycles. *PLoS Med.* 2011; **8**(1): e1000386.
PubMed Abstract | Publisher Full Text | Free Full Text

22. McLernon DJ, Steyerberg EW, Te Velde ER, *et al.*: **Predicting the chances of a live birth after one or more complete cycles of *in vitro* fertilisation: population based study of linked cycle data from 113 873 women.** *BMJ.* 2016; **355**: i5735.
PubMed Abstract | Publisher Full Text | Free Full Text

23. te Velde ER, Nieboer D, Lintsen AM, *et al.*: **Comparison of two models predicting IVF success; the effect of time trends on model performance.** *Hum Reprod.* 2014; **29**(1): 57–64.
PubMed Abstract | Publisher Full Text

24. Leijdekkers JA, Eijkemans MJC, van Tilborg TC, *et al.*: **Predicting the cumulative chance of live birth over multiple complete cycles of *in vitro* fertilization: an external validation study.** *Hum Reprod.* 2018; **33**(9): 1684–1695.
PubMed Abstract | Publisher Full Text

25. Künzle R, Mueller MD, Hänggi W, *et al.*: **Semen quality of male smokers and nonsmokers in infertile couples.** *Fertil Steril.* 2003; **79**(2): 287–291.
PubMed Abstract | Publisher Full Text

26. Keane K, Cruzat VF, Wagle S, *et al.*: **Specific ranges of anti-Mullerian hormone and antral follicle count correlate to provide a prognostic indicator for IVF outcome.** *Reprod Biol.* 2017; **17**(1): 51–59.
PubMed Abstract | Publisher Full Text

27. Dechanet C, Anahory T, Mathieu Daude JC, *et al.*: **Effects of cigarette smoking on reproduction.** *Hum Reprod Update.* 2011; **17**(1): 76–95.
PubMed Abstract | Publisher Full Text

28. Jurisica I, Mylopoulos J, Glasgow J, *et al.*: **Case-based reasoning in IVF: prediction and knowledge mining.** *Artif Intell Med.* 1998; **12**(1): 1–24.
PubMed Abstract | Publisher Full Text

29. Guh RS, Wu TCJ, Weng SP: **Integrating genetic algorithm and decision tree learning for assistance in predicting *in vitro* fertilization outcomes.** *Expert Syst Appl.* 2011; **38**(4): 4437–4449.
Publisher Full Text

30. Güvenir HA, Misirli G, Dilbaz S, *et al.*: **Estimating the chance of success in IVF treatment using a ranking algorithm.** *Med Biol Eng Comput.* 2015; **53**(9): 911–920.
PubMed Abstract | Publisher Full Text | Free Full Text

31. Milewska AJ, Jankowska D, Cwalina U: **Prediction of Infertility Treatment Outcomes Using Classification Trees.** *Stud LOGIC Gramm Rhetor.* 2016; **47**(60): 7–19.
Publisher Full Text

32. Qiu J, Li P, Dong M, *et al.*: **Personalized prediction of live birth prior to the first in vitro fertilization treatment: a machine learning method.** *J Transl Med.* 2019; **17**(1): 317.
PubMed Abstract | Publisher Full Text | Free Full Text

33. cnpma: **Conselho Nacional de Procriação Medicamente Assistida - Modelos de Consentimento Informado**. [Accessed: 13-Aug-2019].
Reference Source

34. Daniel W, Cross C: **Biostatistics: A Foundation for Analysis in the Health Sciences**. 10th ed. Hoboken: John Wiley & Sons, 2013.
Reference Source

35. Hosmer DW, Lemeshow S: **Applied Logistic Regression**. 2nd ed. NY: John Wiley & Sons, Inc, 2000.
Publisher Full Text

36. Marill T, Green DM: **On the Effectiveness of Receptors in Recognition Systems.** *IEEE Trans Inf theory.* 1963; **9**(1): 11–17.
Publisher Full Text

37. Fisher RA: **The Arrangement of Field Experiments.** *J Minist Agric Gt Britain.* 1926; **33**: 503–513.
Reference Source

38. Hart RJ: **Physiological Aspects of Female Fertility: Role of the Environment, Modern Lifestyle, and Genetics.** *Physiol Rev.* 2016; **96**(3): 873–909.
PubMed Abstract | Publisher Full Text

39. Hajshafiha M, Ghareaghaji R, Salemi S, *et al.*: **Association of body mass index with some fertility markers among male partners of infertile couples.** *Int J Gen Med.* 2013; **6**: 447–451.
PubMed Abstract | Publisher Full Text | Free Full Text

40. Freour T, Masson D, Mirallie S, *et al.*: **Active smoking compromises IVF outcome and affects ovarian reserve.** *Reprod Biomed Online.* 2008; **16**(1): 96–102.
PubMed Abstract | Publisher Full Text

41. Swets JA: **Measuring the accuracy of diagnostic systems.** *Science.* 1988; **240**(4857): 1285–1293.
PubMed Abstract | Publisher Full Text

42. Coppus SF, Van Der Veen F, Opmeer BC, *et al.*: **Evaluating prediction models in reproductive medicine.** *Hum Reprod.* 2009; **24**(8): 1774–1778.
PubMed Abstract | Publisher Full Text

43. Efron B, Tibshirani JT: **An Introduction to the Bootstrap**. 1st ed. 1993.
Reference Source

44. von Elm E, Altman DG, Egger M, *et al.*: **The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.** *J Clin Epidemiol.* 2008; **61**(4): 344–349.
PubMed Abstract | Publisher Full Text

45. Zarinara A, Zeraati H, Kamali K, *et al.*: **Models Predicting Success of Infertility Treatment: A Systematic Review.** *J Reprod Infertil.* 2016; **17**(2): 68–81.
PubMed Abstract | Free Full Text

46. Best D, Bhattacharya S: **Obesity and fertility.** *Horm Mol Biol Clin Investig.* 2015; **24**(1): 1–6.
PubMed Abstract | Publisher Full Text

47. Assembleia da República: **Lei n.º 56/79 - Diário da República n.º 214/1979, Série I de 1979-09-15**. 1979; 2357–2363.
Reference Source

Open Peer Review

Current Peer Review Status: ✗ ✗ ?

Version 2

Reviewer Report 11 June 2024

https://doi.org/10.5256/f1000research.30624.r285602

?

Edwin Amalraj Raja
University of Aberdeen, Aberdeen, UK

Summary

Thank you very much for asking me to review a manuscript entitled "Predicting the chance of live birth for couples undergoing IVF: a novel instrument to advise patients and physicians before treatment" by Bruna Estacio da Veiga et al,. The authors claim that this is the first model with Portuguese data that includes AFC, AMH and women's age. The objective of the study is to develop a pre-treatment prediction model to estimate the cumulative probability of a live birth among couples seeking infertility treatment. They used 739 IVF cycles performed between 2012 and 2016 at the CIRMA Hospital, Almada, Portugal. Logistic regression was used to create a predictive model for the live birth based on the demographic characteristics, and clinical factors of the couple.
A Few points to consider (not in any particular order) to improve the scientific presentation of the manuscript:

Major comments:

1. The data from 739 (or 737) cycles were used in the study. Has each cycle come from a (different) couple? If there was more than one cycle per woman, please provide details. Since it is a pre-treatment model, the baseline characteristics and clinical factors, though collected at multiple visits before treatment, might have been used for the analysis. Further, the table heading has 'couples'. Please decide whether the data are from the first cycle (without considering the outcome from further frozen embryo transfer) or woman, or couple. Be consistent in the use of the study population. If the data were from the first cycles, the word 'cumulative' in the objective is meaningless.
2. You defined the model's dependent variable/primary outcome clearly. Similarly, you must define other predictors with numbers used for coding each category. This will help researchers to reproduce your results, validate your model in their dataset (external) and use it for prediction in the clinical setting.
3. STROBE is the guideline for reporting observational study. Your study is developing and

validating a prediction model. It is recommended to use guidelines for reporting a prediction model such as TRIPOD or any updated guidelines.

4. I don't think reference 36 is appropriate for 'the use of backward selection process' for the model building. The guidelines for selecting variables for the prediction model have changed since 1963. Please provide a reference to justify the variable selection process from the recent literature on prediction modelling.
5. In the methods, there are general statements. "A summary statistical evaluation was performed for each variable. A univariate analysis was first performed". This should be avoided. Instead, specify how you decided which summary measures are appropriate for each variable. For example., The continuous variable is presented as mean (SD) if it is normally distributed or median(25^{th} percentile , 75^{th} percentile) if it is skewed. The categorical variable is presented as number(percent).
6. The results showed no evidence that the authors checked for the normality assumption before summarising data. The variables: AFC, AMH and duration of Infertility are reported as skewed data in the previous literature. Please check the assumption before summarising a continuous variable and provide the appropriate summary measures in Table 1.
7. Check the assumption of linearity between a continuous variable and log odds of live birth. If the relationship is not linear, use the spline/polynomial function to model. It increases the predictive ability of the model.
8. In the model building, after the main effects, check whether adding interaction term improves the AUC.
9. Please check the multicollinearity of the variables in the model (age vs age, BMI vs BMI, AFC vs AMH, etc.,)
10. The model's performance was assessed mainly with discrimination and calibration. There was no report of calibration (graph) in the manuscript.
11. "the model was internally validated using the bootstrapping technique with 1000 iterations". There was no evidence in this in the results section.
12. Interpretation of the results: When the model contains interaction terms(, but not main effects), the interpretation of the regression coefficients is not straightforward. Please consult a medical statistician to interpret the results.
13. Please explain clearly (with an example) how a clinician/patient can utilise your novel tool in the clinical practice before treatment.

Minor comments:

1. The number of cycles/women/couples mentioned is different in Abstract (737), Methods (739) and Results (737).
2. Please mention the number of live births and the percentage within brackets.
3. When you provide mean or median values, support them with SD/ IQR (p25, p75) respectively.
4. Within methods, there are two statements about the level of significance: (i) A p-value < 0.05 was considered statistically significant and (ii) it was considered an alpha of 0.05. Retain the latter statement.
5. Within methods, the coding for the primary outcome is repeated (no=0 and 1= yes). This is not necessary.
6. The statement "There should be no bias associated with this study" is meaningless unless it refers to a specific procedure in the methods.
7. 'Methods' should be 'Methods and Materials'
8. FIV-ICSI appears in a few places. Is it IVF-ICSI?

9. Please report AUC to two digits precision
10. Present mean (SD) or median (p25, p75) wherever appropriate and not as mean plus or minus SD
11. reporting minimum and maximum is ok for all couples but not necessary for couples with live birth or without live birth
12. the titles of table 1 and table 2 should reflect the content of the tables and not a general title like 'Baseline continuous characteristics of couples'.
13. There is no need to mention SPSS in the results
14."Chi-square test revealed that no statistically significant differences for live birth (P>0.05) for categorical variables.". This is a general statement. Please provide actual results supported by actual p-value.
15. The regression coefficients are not odds/probabilities these are log-odds of live birth.
16. When a p-value=0.000 then report it as p<0.001

I think the manuscript requires a major revision.

Thank you.

Edwin Amalraj Raja
Medical Statistician
University of Aberdeen &
Statistical Editor, Human Fertility

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: Medical Statistician working in reproductive medicine

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have

significant reservations, as outlined above.

Version 1

Reviewer Report 07 September 2020

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

Jichun Tan
[1] Center of Reproductive Medicine, Department of Obstetrics and Gynecology, Shengjing Hospital of China Medical University, Shenyang, China
[2] 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.

Author Response 03 Nov 2020

Bruna Estácio da Veiga

Lisbon, November the 3rd 2020

We submitted an original research article entitled “Predicting the chances of live birth for couples undergoing IVF ICSI: a novel instrument to advise patients and physicians before treatment” for consideration by F1000Research journal that has been reviewed by two experts in the field, suggesting revisions to this paper. We want to thank their comments and we have done the following revisions accordingly for the second reviewer (all modifications can be seen through the “Track Changes” option):

Point 1: This paper focused on an interesting topic of "What are my chances of conceiving?". However, the aim of this study is not clear.
Response: We followed the reviewer’s advice by modifying the abstract and introduction section, reinforcing the aim of our study: “The main aim of this study is to develop a validated model that estimates the chance of a live birth before the start of an IVF/ICSI non-donor cycle.” and “The main aim of this study was to find a pre-treatment predictor for achieving a live birth before an IVF/ICSI treatment.”, respectively.

Point 2: 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.
Response: Considering the size of the sample available and the reviewer’s assessment of the model’s evaluation, we added a table (Table 4) that lists 5 performance metrics (i.e., accuracy, F-score, precision, sensitivity and sensibility) and reflects our model performance, in order to respond to the reviewer’s observation. Particularly about the sample size, we still achieved a ROC curve with a 95% Confidence Interval (CI) that is not very wide (“The ROC curve (...) had an AUC equal to 0.700 (95% CI 0.660–0.741)”). Although other important parameters are found in other studies, the 14 variables tested in our study were all the variables collected by CIRMA and made available for this study. Therefore, we can not follow the reviewer’s suggestion and include more variables, because we did not collect and select the couple’s data.

Point 3: In addition, more latest studies should be added in the introduction section.
Response: We thank the reviewer for this final comment and we tried to enrich the introduction section by adding more information about AI in predictive models, considering the recent study of Qiu et.al (reference 32) based on supervised learning algorithms.

Sincerely,

The authors.

Competing Interests: No competing interests were disclosed.

Reviewer Report 31 July 2020

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

Charalampos Siristatidis
[1] Assisted Reproduction Unit, Third Department of Obstetrics and Gynecology, Medical School, National and Kapodistrian University of Athens, "Attikon" University Hospital, Athens, Greece
[2] 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; mmethodology; 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.

Author Response 03 Nov 2020

Bruna Estácio da Veiga

Lisbon, November the 3rd 2020

We submitted an original research article entitled "Predicting the chances of live birth for couples undergoing IVF ICSI: a novel instrument to advise patients and physicians before treatment" for consideration by F1000Research journal that has been reviewed by two experts in the field, suggesting revisions to this paper. We want to thank their comments and we have done the following revisions accordingly for the first reviewer (all modifications can be seen through the "Track Changes" option):

Point 1: 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.
Response: We followed the reviewer's advice by modifying the introduction section, including more information about AI in predictive models, considering the recent study of Qiu et.al (reference 32) based on supervised learning algorithms.

Point 2: Language and grammar need revision.
Response: The reviewer asked for English changes and we reviewed the entire text.

Point 3: The sample size is small. A proper power calculation is missing.
Response: Considering the size of the sample available and the reviewer's assessment of the model's evaluation, we added a table (Table 4) that lists 5 performance metrics (i.e., accuracy, F-score, precision, sensitivity and sensibility) and reflects our model performance, in order to respond to the reviewer's observation. Particularly about the sample size, we still achieved a ROC curve with a 95% Confidence Interval (CI) that is not very wide ("The ROC curve (...) had an AUC equal to 0.700 (95% CI 0.660–0.741)").

Point 4: Some important parameters that have been tested in other studies that contribute to live birth are missing.
Response: Although other important parameters are found in other studies, the 14 variables tested in our study were all the variables collected by CIRMA and made available for this study. Therefore, we can not follow the reviewer's suggestion and include more variables, because we did not collect and select the couple's data.

Point 5: Limitations are not reported.
Response: We thank the reviewer for this final comment and we tried to enrich the discussion section by adding more clinical limitations found in our model to the limitation reported before: "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. That is, the resulting probability of our model should be considered a baseline one. Taking into account that this study was performed with data from only one medical center, one limitation of these models is that only an internal validation was made. Moreover, the large temporal spectrum of the data (2012 to 2016) could mean that treatments made with older technology may have different results than those made with more recent ones (namely, vitrification procedures or culture media).".

Sincerely,
The authors.

Competing Interests: No competing interests were disclosed.