Women's Health & Gynecology

ISSN: 2369-307X

Research Article

The Number of Embryos Obtained can offset the Age Factor in IVF Results According to an Artificial Intelligence System

This article was published in the following Scient Open Access Journal:
Women's Health & Gynecology
Received March 28, 2016; **Accepted** April 25, 2016; **Published** May 06, 2016

Ana Karina Bartmann[1,2,3,4*], Milton Faria[4], Giselle Pulice de Barros[1,3], Larissa Sampaio de Paula[1,3] and Nathalia Rodrigues Bettini[1]

[1]*Assisted Reproduction Center - Department Gynecology and Obstetrics - Medical School - University of Ribeirao Preto (UNAERP), Ribeirao Preto -SP - Brazil*

[2]*Human Reproduction Center of the Ana Bartmann Clinic/Ribeirao Preto - SP - Brazil*

[3]*National Council of Scientific and Technological Development (CNPq) - Brazil*

[4]*Department of Biotechnology - University of Ribeirao Preto, Ribeirao Preto- SP - Brazil*

Abstract

Purpose: Creation of an artificial intelligence system able to calculate the chances of pregnancy when fed with the patients' clinical and laboratory information.

Methods: One hundred and twenty-one women who underwent IVF or ICSI were clinically and lab analyzed and sixty of those were used to train the artificial neural network.

Results: The analysis of clinical and laboratory data showed, with 93% confidence, that the number of embryos obtained after the IVF has a greater correlation with pregnancy than the others. The artificial intelligence system also showed a high potential to predict pregnancy, after the training of the neural network.

Conclusions: After training with a greater number of women, the intelligence system will be a highly useful tool in assisted reproduction centers, because it can be used for countless multivariate analyses and to quantify the real chances of a woman's pregnancy in *in vitro* fertilization cycles.

Keywords: Artificial Intelligence,Embryos, IVF, ROC,Pregnancy,Curve,Pearson correlation

Capsule

A new data system model with artificial intelligence is capable to determine which variables are the best predictors of pregnancy.

Introduction

There has been a desire to determine clinical parameters related to the attainment of pregnancy since the beginning of Assisted Reproduction. We know, for example, that natural fecundity and pregnancy rates decline with age, even in assisted reproductive (AR) procedures [1]. Women's fertility decreases after age 35, because their germ cells are not replenished during life [2]. The utilization of follicles leads to a decline in the quantity and quality of oocytes until menopause [3-5].

Antral Follicle Count (AFC) is used as a pregnancy possibility parameter; it demonstrates the potential of oocytes that can be stimulated in the menstrual cycle and thus have the eventual ability to predictively assess ovarian response. A low number of antral follicles correlate with a lower ovarian response duringhyper stimulation for IVF[6-8].Another parameter, FORT (Follicular Output Rate) was created in order to assess the "utilization" rate of the follicles, that is, the ratio between the pre-ovulatory follicle counts on the day of hCG administration × 100 / antral follicle count (AFC). This relates closely to the individual response ability to the ovarian stimulation and consequently to the reproductive prognosis [9].

There is a wide variation in terms of quantity and quality regarding the number of embryos obtained. We over stimulate in order to mathematically increase the chances of pregnancy in AR. In theory, the more embryos the greater the chances of success. The possibility of pregnancy, however,falls drastically in the case of patients older than 38 years of age, regardless of the number of embryos [10]. This is likely due to the declining quality of the embryos in older patients.

Although each of the clinical/laboratory variables described above has already been exhaustively studied, the strength of these parameters, when associated with the others, is still unknown. For this reason, we analyzed the ROC curve of each attribute and the correlation between them [11-15]. We also developed an Artificial Intelligence (AI) program that can predict the likelihood of pregnancy according to the attributes

***Corresponding author:** Dr. Ana Karina Bartmann, Professor, Centro de Reproducao Humana - Hospital Electro BoniniAv. Leao XIII, s/n - 2º andarCEP: 14096-180,Ribeirao Preto - SP, Brazil,
Email: clinica@anabartmann.com.br

of each patient and the result of each step of the IVF treatment.

AI aims to create computer systems that are similar to human behavior. These systems support health professionals, assisting them in their clinical routine, in handling data and knowledge. Thus, they are able to analyze and classify the data of individualized patients [16]. In literature, there are many studies which try to calculate more personalized chances of a woman become pregnant[17-19], but there is no consensus of what method is more accurate. To our knowledge there is no study using AI as fundamental tool. The idea of this study was to analyze clinical parameters of IVF cycles using an AI system. The main outcome wasto identify which variables are more important to achieve pregnancy. Secondly, we want to offer a new tool capable to individuallypredict pregnancy chances.

Materials and Methods

In this retrospective, descriptive, cross sectional study, one hundred and twenty-one cases obtained from the database of the Human Reproduction Center of the Ana Bartmann Clinic/ UNAERP were clinically and lab analyzed and sixty of those were used to train the artificial neural network from October 2015 to January 2016. All patients treated until the end, i.e., with positive or negative pregnancy test, were included. We excluded patients who were still waiting for pregnancy test results or those in treatment.Beingthis a descriptive study, we did no sample size calculation.

Patients had multiple stimulation protocols, including estrogen prime, microdoselupron or GnRHantagonist. Baseline ultrasounds were performed on day two of menstrual cycle and subsequently until when at least 3 follicles reached 18mm in diameter. Oocyte maturation was induced using 10,000 units of subcutaneousrecombinant human chorionic gonadotropin (hCG). Oocyte retrieval was performed 35 hours post trigger injection. Insemination or ICSI occurred 3 to 6 hours post retrieval.The primary outcome was pregnancy, which was defined as presence of embryo with heartbeat in ultrasound.

All patients had clinical and laboratorial data determined to further analysis in AI system.

The neural network architecture used was a multilayer perceptron with a back propagation training algorithm. The AI system was developed in the Exact Sciences Department of the University of Ribeirao Pretoespecially for our work. The study has the institutional review board (IRB) approval.

The analysis used the following variables: Age, Antral follicle count, Mature follicle count (with an average diameter of over 16 mm), FORT (ratio between pre-ovulatory count on the day of the administration of hCG × 100/AFC), Number of oocytes obtained and Number of embryos obtained. We also performed a discriminant multivariate analysis to determine which variables are most important in the result.

During the development phase of the system, we carried out a reliability analysis of the attributes by the ROC curve and their correlation. This method is used to visualize and select classifiers based on their performance.

Results

The area under the ROC curve for the number of embryos obtained was 0.917, with significance of p value < 0.001, sensitivity of < 72.73% and specificity of 95.45%, thus being the best parameter found for pregnancy success in all 121 women.

The area under the ROC curve for the mature follicles count was 0.888, with significance of p value < 0.001, sensitivity of 72.73%, and specificity of 90.91%.

The area under the ROC curve for oocytes collected was 0.881, with significance of p value < 0.001, sensitivity of 72.73%, and specificity of 90.91%.

The area under the ROC curve for FORT was 0.822, with significance of p value < 0.001, sensitivity of 81.82%, and specificity of 69.32%.

The area under the ROC curve for the antral follicle count was 0.773, with significance of p value < 0.001, sensitivity of 81.82%, and specificity of 60.23%. In this case, AFCs higher than 12 are associated with greater likelihood of pregnancy. This shows the study's high reliability, since the literature considers an AFC higher than 10 and 15 appropriate [9].

The area under the ROC curve for age was 0.701, with significance of p value 0.001, Youden index J of 0.3220, association criteria < 37, sensitivity of 69.70%, and specificity of 62.50% (Figure 1). The number of embryos' graph had a larger area compared to those of other clinical parameters.

The Pearson correlation graphs showed that age does not present a good correlation with the other clinical parameters studied. That is noticeable in Figure 2, where the "r" values were between -0.274 and -0.494, although the value of $p < 0.05$(Figure 2).

Among the graphs obtained, we found that the most correlated variables are those involving the number of embryos obtained and the number of oocytes collected. The graphs show a positive linear correlation with a value of Pearson's "r" moment greater than 0.9. That is, there is a positive and very strong correlation between the parameters mentioned (Figure 3).

Pearson correlation graph (A) Between the number of embryos obtained and the MFC. (B) Between the number of embryos obtained and the number of oocytes collected. There was a strong correlation (r > 0.9) in both.

We have achieved a highly effective AI system after training the neural network. The analysis of the attributes using the ROC curve ensured the reliability of the system. During the data analysis carried out by the AI system developed, we obtained a 93.33 % accuracy, with an error Ɛ = 0.002 in a training with 60 cases.

Discussion

We found that the number of embryos obtained was the best discriminant variable and the one that correlated the most with the analyzed variable in order to predict pregnancy, both by ROC curve analysis and Pearson correlation. That means that a woman with more embryos, even with age, FORT, AFC and MFC below standard, tends to have greater chances of pregnancy according to the AI system analysis. The number of oocytes collected and the MFC were respectively the data that correlated the most with the number of embryos, thus, a woman with a high number of mature follicles, collected oocytes and obtained embryos, has the best chances of getting pregnant.

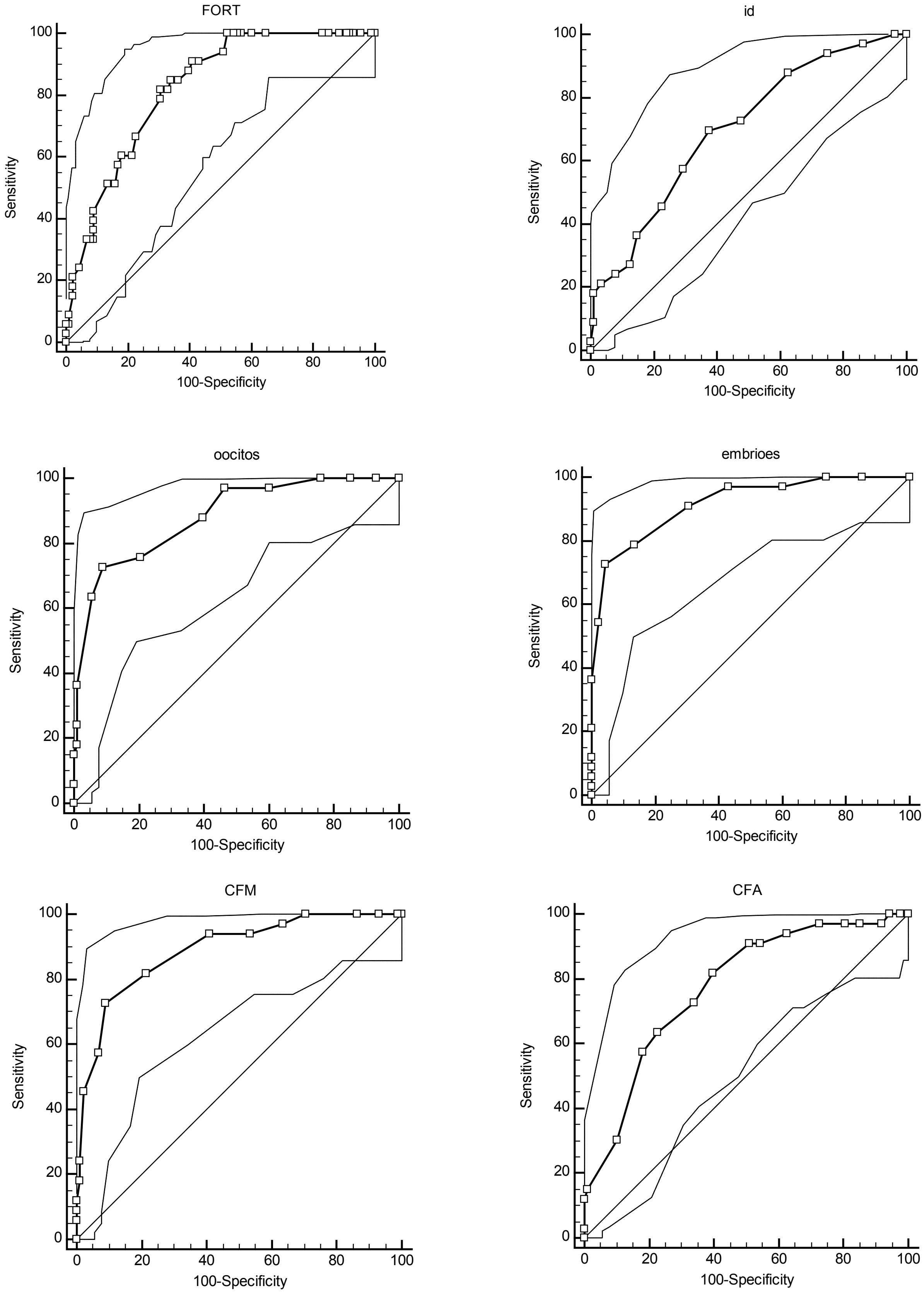

Figure 1: ROC curve graphs. FORT= Follicular Output Rate, MFC= Mature follicle count, AFC= Antral follicle count.

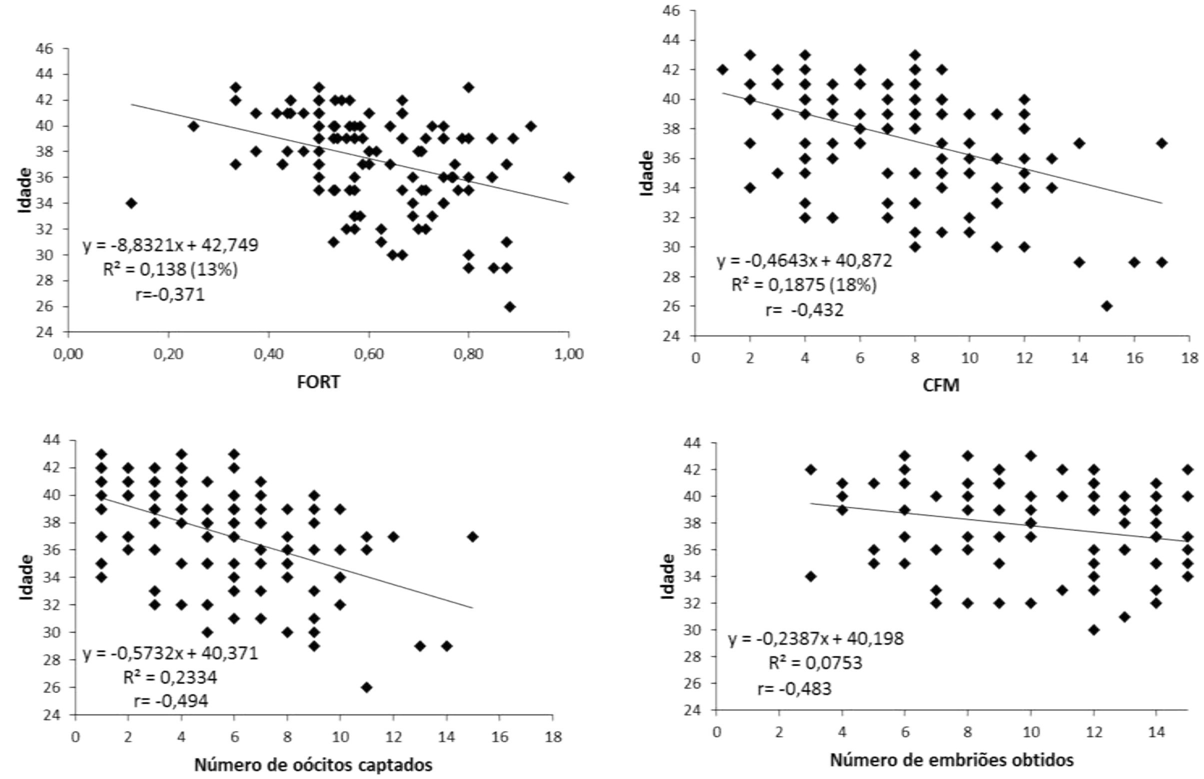

Figure 2: Graph of Pearson correlation between age and other parameters. FORT= Follicular Output Rate, MFC= Mature follicle count, AFC= Antral follicle count.

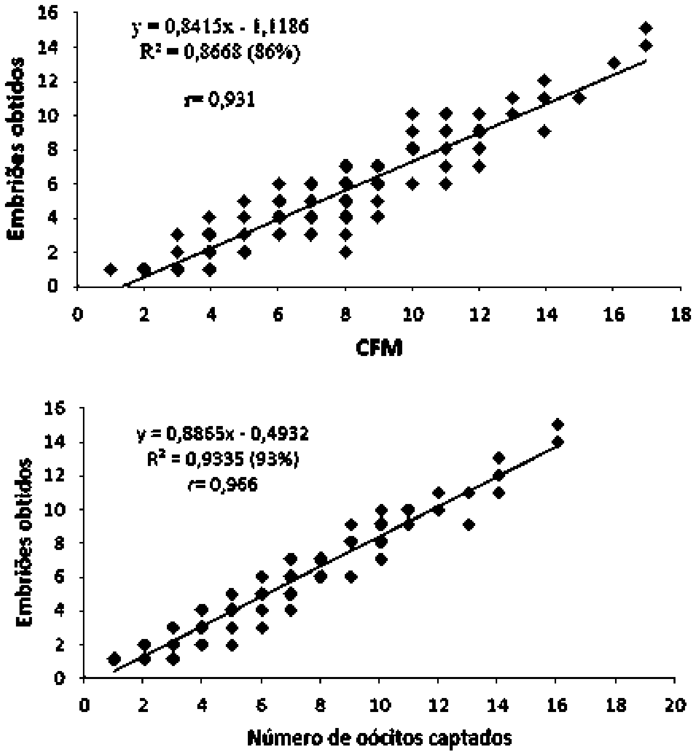

Figure 3: Pearson correlation graph. (A) Between the number of embryos obtained and the MFC. (B) Between the number of embryos obtained and the number of oocytes collected. There was a strong correlation (r > 0.9) in both. MFC= Mature follicle count.

Many other studies have been made to establish the best IVF predictors[17-19]. Some took into account fertilization rate, AMH levels, embryo score or even the number of embryos transferred. Others constructed mathematical models to predict success rates before the start of ovarian stimulation [17,18].The difference between our study and others is that AI can provide customized prediction of pregnancy according to what is being achieved during treatment but before embryo transfer. Moreover, our findings are radically different from the current concept in which age is a major predictor.

Although age is an important parameter [1] when explored separately, when this variable is analyzed with the others, it loses its significance and the number of embryos eventually becomes the most important factor for the success of pregnancy.

As to the AFC, MFC and FORT, we obtained a higher correlation between the number of embryos and the MFC. The correlation between embryos and AFC was also high (r = 0.75), but, surprisingly, the same does not occur between embryos and FORT. Hence, we inferred that the AFC is responsible for decreasing the correlation between FORT (MFC/AFC) and the number of embryos, making the FORT a weak correlation variable. Again, the AFC was not, in this study, considered a good predictor of pregnancy when compared to other grouped variables.

As to the AI system, we can conclude that the developed system offers a high potential to predict pregnancy. The system needs more data to further increase its accuracy, because it aims to determine each patient's individual chance to achieve pregnancy, according to their personal characteristics and response to treatment. The purpose of the system is to be distributed free of charge to interested reproduction centers, so that with thousands of reported cases, we can define with greater accuracy the real probability of pregnancy. We also wish to add new analysis parameters, such as the formation of top embryos, BMI, embryo morphology, metabolome, etc.

Limitations: Our AI model is unable to predict pregnancy before treatment begins. It was designed to use variables that become available as treatment progresses so that pregnancy chances are calculated only before embryo transfer. Sixty individuals composed the population utilized for the neural training. Although the AI system was able to correctly set more than 95% of the cases, this population size is not by far enough to validate our findings. More individuals, preferentially from other ART centers should be included to increase the power of the model.

Conclusion

The study concluded that the AI tool created has high potential to be used in health programs, including Assisted Human Reproduction Services. The same method can be used for numerous multivariate analyses in a Human Reproduction Center, including the selection of embryos for intrauterine transfer, considering that the method is able to create a good score with high sensitivity and specificity. Although our findings are not new, to our knowledge, this is the first study to present an AI model predicting IVF outcome. Following validation from other ART Centers, AI model may help decision-making for clinicians and patients.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgements

We thank the collaboration of the professionals of the "Hospital Electro Bonini" and of the Human Reproduction Center of the Ana Bartmann Clinic.

References

1. Van Rooij IA,Bancsi LF, Broekmans FJ, Looman CW, Habbema JD, Velde ER. Women older than 40 years of age and those with elevated follicle-stimulating hormone levels differ in poor response rate and embryo quality in in vitro fertilization. *FertSteril*. 2003;79(3):482-488.
2. Abdalla H, Thum MY. An elevated basal FSH reflects a quantitative rather than qualitative decline of the ovarian reserve. *Hum Reprod*. 2004;19(4):893-898.
3. Baird DT, Collins J, Egozcue, J, et al. Fertility and ageing. *Hum Reprod Update*. 2005;11(3):261-276.
4. Toner JP. Modest follicle-stimulating hormone elevations in younger women: warn but don't disqualify. *FertilSteril.* 2004;81(6):1493-1495.
5. Toner JP. Age = egg quality, FSH level = egg quantity. *Fertil Steril.* 2003;79(3):491.
6. Silva GM, Diniz AL, Neto MB, et al. Numero de foliculos antrais e o sucesso da fertilizacao in vitro: uma análise multivariada. *RevistaBrasileira de GinecologiaObstetricia*. 2014;36:473-479.
7. Adibi A, Mardanian F, Hajiahmadi S. Comparison of ovarian volume and antral follicle count with endocrine test for prediction of responsiveness in ovulation induction protocols. *AdvBiomed Res.* 2012;1:71.
8. Carvalho BR, Rosa e Silva, AC. Hormonio anti-mulleriano e contagem de folículos antrais como marcadores de reserva ovariana. Sobrage. 2007;7:5-8.
9. Gallot V, BerwangerAL, Genro V, Grynberg M, Frydman N, Fanchin R. Antral follicle responsiveness to follicle-stimulating hormone administration assessed by the Follicular Output RaTe (FORT) may predict in vitro fertilization-embryo transfer outcome. *Hum Reprod.* 2012;27(4):1066-1072.
10. 10.Boomsma CM, Macklon NS. What can the clinician do to improve implantation? *Reprod BiomedOnline*. 2007;14:27-37.
11. Prati RC, Flach P. "ROCCER: an algorithm for rule learning based on ROC analysis", in Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI'2005). Edinburgh (UK). 2005.
12. Fawcett T, Flach PA. "A response to Webb and Ting's on the application of ROC analysis to predict classification performance under varying class distributions". *Machine Learning.* 2005; 58(1):33-38.
13. Fluss R, Faraggi D, Reiser B. Estimation of the Younden Index and its associated cutoff point. *Biom J.* 2005;47(4):458-472.
14. Filho DB, Júnior JA. Desvendando os Mistérios do Coeficiente de Correlação de Pearson (r). *RevistaPolíticaHoje.* 2009;18(1).
15. Anscombe FJ. Graphs in Statistical Analysis.*The American Statistician.*1973;27(1):17-21.
16. Coeira, E. Inteligencia Artificial na Medicina. Informatica Medica,1998;1(4).
17. 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.
18. 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.
19. teVelde 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;29:57-64.