The Peerian Journal

Open Access | Peer Reviewed

Volume 35, October, 2024

ISSN (E): 2788-0303

Website: www.peerianjournal.com

Email: editor@peerianjournal.com

# Prediction Of Implantation Results In Vitro Fertilization And Intracytoplasmic Sperm Injection Using Data Mining Methods

Candidate of Medical Sciences - **Egamkulov Z.T**

**Abstract:** In vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI) are two important subtypes of assisted reproductive technologies used to treat infertility. Predicting the outcome of IVF/ICSI implantation or the likelihood of pregnancy is essential for infertile couples as these treatments are complex and expensive, and the probability of conception is low.

**Background:** Assisted reproductive technologies (ART) include all therapies that are used to process human oocytes and sperm or embryos in vitro to establish pregnancy (1). Infertility is defined as the inability of a couple to conceive after 12 months of regular unprotected intercourse (2). Among the treatments for ART, in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI) are well-known methods of infertility treatment. The IVF process involves ovarian stimulation, oocyte retrieval, fertilization, embryo culture, and embryo transfer to the uterus (3). ICSI is another treatment method used for infertile couples that involves injecting a selected sperm into the cytoplasm of an oocyte (4). A recent study in Iran (5) shows that the overall average infertility rate is about 10.9% of the population. This study states that among the patients of several infertility clinics in the country, 78.4% had primary and 21.6% had secondary fertility factors. The results give a 34.0% average percentage for the male factor, 43.5% for the female factor, 17.1% for both factors, and 8.1% for unexplained infertility. Ovulatory dysfunction was the most common etiological factor among female causes in this study. Today, many couples suffering from infertility try to have a child with ART and ask about the likelihood of pregnancy for several reasons. Firstly, due to the high cost of IVF and ICSI treatment in Iran, some couples cannot afford the cost of these treatments. Then, the probability of conception is 20–25% in the normal reproductive cycle (3), which in ART increases to about 30–40% in each cycle; However, it is still considered low. Finally, ART consists of several stages that are time-consuming and difficult for infertile women to tolerate. There are also three main clinical reasons that make predicting pregnancy outcomes necessary. First, there are many prognostic factors for this treatment that determine the likelihood of conception, which in turn makes it difficult for doctors to make a decision. Second, using previous cases for this solution seems reliable, although it is a time-consuming task for doctors. And finally, there may be an alternative method of IVF and ICSI, which the specialist offers to couples with a very low probability of pregnancy, such as adoption, which makes them refuse infertility treatment. Data mining (DM) refers to the use of machine learning, pattern recognition, and statistical techniques to extract knowledge from data, in this case patient information, and is a specific step in the process of knowledge discovery in databases (KDD) (6). In medical DM, the classification system predicts the class to which a patient belongs by training a model based on the input dataset. Because DM techniques perform data analysis and extract valuable insights from the data, clinical obstetricians and gynecologists can use this information for diagnosis and treatment (7). According to Sios and Moore (8), medical DM can be useful to patients in finding a solution to analyze different types of clinical data.

**Keywords:** In vitro fertilization, intracytoplasmic sperm injection, data mining.

The Peerian Journal

Open Access | Peer Reviewed

Volume 35, October, 2024

ISSN (E): 2788-0303

Website: www.peerianjournal.com

Email: editor@peerianjournal.com

**Research Materials and Methods:** A dataset of 486 labeled records along with 29 variables was collected at the Infertility Research Center of Mother and Child Hospital in Shiraz, Iran from 2009 to 2015. Each patient signed a consent form at the time of admission to the hospital and before the start of the study. This study was approved by the Ethics Committee of Shiraz University of Medical Sciences. and the sampling method is a census. This dataset contained 131 positive and 355 negative implantations. Because the number of negative samples exceeds the number of positive samples, this dataset is highly imbalanced. The necessary variables for this study were extracted from paper medical records by our trained staff. In order to use these records for computer models, a data entry process was performed. In this study, the results of implantation of frozen embryos were excluded and only fresh embryo transfer was considered due to the differences in some characteristics between the two transfer methods. [9]

**Results of the study:** The processed samples were applied to each classifier to calculate AUC and accuracy over a 5x CV and presented as averages. Each experiment is repeated 20 times to examine a complex combination of data samples. The average of these experiments for each classifier is reported in addition to the standard deviation. In addition, specificity is also calculated for each classifier, sensitivity, PPV, NPV, LR+ and LR-. Our results show that RF and RPART outperform other classifiers in terms of specificity, PPV and NPV. RPART predicts positive cases better than RF; however, negative cases are classified better by RF than RPART. The higher PPV value in RF is due to fewer false positives. Apparently, the higher NPV value in RPART is due to fewer false negatives in the confusion matrices of both models. Adaboost has better values overall, especially in terms of sensitivity compared to SVMs and 1NNs. While the specificity of SVM is 88.73% and higher than 1NN, its specificity value (14.5%) is very low. Interestingly, in a positive pregnancy, a high positive likelihood ratio of RF shows a large increase in the likelihood of pregnancy, and the corresponding value for RPART implies a moderate increase. However, the rest of the models result in minimal magnification. Negative likelihood ratios of all classifiers, which are almost between 0.5 and 1, represent a minimal decrease in the probability of pregnancy.
Of all the classifiers tested in this study, RPART leads to the most useful information beyond the IVF/ICSI success rate. Therefore, we present the significance of the 20 IVF/ICSI traits using RPART. The second column shows the scores of each trait. Note that only 11 traits have specific values for a positive pregnancy because these traits were significant in the RPART's decision. predicting the outcome of IVF/ICSI, since they do not have certain values for a positive pregnancy. Figure 1 shows the ROC curves for predictive models using all data samples. As you can see, RF and RPART have a higher AUC compared to Adaboost, SVM, and 1NN, and the SVM curve is closer to the top two classifiers than 1NN and Adaboost. [9]

**Conclusions:** According to the results obtained in the current study, RF and RPART outperformed other methods for predicting pregnancy with AUCs of 84.23 and 82.05%, respectively. In addition to the issue of classifiers, knowledge in the form of selected traits is extracted from the RPART model. Female age, number of embryos developed, and serum estradiol (E2) level on the day of administration of human chorionic gonadotropin (hCG) are presented as the top three predictive signs for IVF/ICSI.

**References:**

1. Zegers-Hochschild F, Adamson GD, de Mouzon J, Ishihara O, Mansour R, Nygren K, et al. International committee for monitoring assisted reproductive technology (ICMART) and the

The Peerian Journal

Open Access | Peer Reviewed

Volume 35, October, 2024

ISSN (E): 2788-0303

Website: www.peerianjournal.com

Email: editor@peerianjournal.com

world health organization (WHO) revised glossary of ART terminology. Fertil Steril. 2009; 92(5):1520–1524. doi: 10.1016/j.fertnstert.2009.09.009.

2. Practice Committee of the American Society for Reproductive Medicine. Definitions of infertility and recurrent pregnancy loss. Fertil Steril. 2008; 90(5 Suppl):S60–S60. doi: 10.1016/j.fertnstert.2008.08.065.
3. Van Voorhis BJ. Clinical practice. In vitro fertilization. N Engl J Med. 2007; 356
4. Van Steirteghem A, Devroey P, Liebaers I. Intracytolasmic sperm injection. Mol Cell Endocrinol. 2002; 186(2):199–203. doi: 10.1016/s0303-7207(01)00658-x.
5. Parsanezhad ME, Jahromi BN, Zare N, Keramati P, Khalili A, Parsa-Nezhad M. Epidemiology and etiology of infertility in Iran, systematic review and meta-analysis. J Womens Health, Issues Care. 2013; 6:2–2.
6. Fayyad U, Piatetsky-Shapiro G, Smyth P. From data mining to knowledge discovery in databases.
7. Wang YF, Chang MY, Chiang RD, Hwang LJ, Lee CM, Wang YH. Mining medical data: a case study of endometriosis. J Med Syst. 2013; 37(2):9899–9899. doi: 10.1007/s10916-012-9899-y.
8. Cios KJ, Moore GW. Uniqueness of medical data mining. Artif Intell Med. 2002; 26(1-2):1–24.
9. Pegah Hafiz, Mohtaram Nematollahi, Reza Boostani, Bahia Namavar Jahromi. Predicting Implantation Outcome of In Vitro Fertilization and Intracytoplasmic Sperm Injection Using Data Mining Techniques. 2017 Sep 3; 11(3):184–190.