DOI 10.26724/2079-8334-2022-4-82-69-73
UDC 618.177-089.888.11

M.K. Ismayilova
The Central Clinic, Baku, Azerbaijan

APPLICATION OF THE LOGISTIC REGRESSION METHOD TO PREDICT THE RESULTS OF IN VITRO FERTILIZATION

e-mail: mahiremk@hotmail.com

Development of the predictors of effectivity of treatment of infertility by means of in vitro fertilization is one of the most important tasks of the modern medicine. We have examined 518 patients applying on account of infertility to the Central Clinic of Baku city. All patients were divided into two groups: the patients with onset pregnancy i.e. positive in vitro fertilization intra cytoplasmic sperm injection results – main group (n=234) and patients with negative in vitro fertilization intra cytoplasmic sperm injection efforts – control group (n=284). The criteria precisely differing in the compared groups were chosen among the big quantity of signs at the women with various in vitro fertilization outcomes. These parameters were used at the factor analysis. Basing upon the result of the factor analysis there were assigned the main factors for forecasting the in vitro fertilization outcomes of infertile women by means of logistic regression. The developed model of forecasting the logistic regression allows precisely forecasting the in vitro fertilization outcome upon certain parameters in 86.93 % cases.

Key words: supplementary reproductive technologies, infertility, logistic regression, predictive value

М.К. Ісмайлова

ЗАСТОСУВАННЯ МЕТОДУ ЛОГІСТИЧНОЇ РЕГРЕСІЇ ДЛЯ ПРОГНОЗУВАННЯ РЕЗУЛЬТАТІВ ЕКСТРАКОРПОРАЛЬНОГО ЗАПЛІДНЕННЯ

Розробка предикторів ефективності лікування безпліддя методом екстракорпорального запліднення є одним із найважливіших завдань сучасної медицини. Нами обстежено 518 пацієнток, які звернулися з приводу безпліддя до Центральної клініки міста Баку. Всі пацієнтки були розділені на дві групи: пацієнтки з вагітністю, тобто, позитивні результати екстракорпорального запліднення та внутрішньоцитоплазматичної ін'єкції сперматозоїдів – основна група (n=234) та пацієнтки з негативними результатами екстракорпорального запліднення та внутрішньоцитоплазматичної ін’єкції (n=284). Критерії, які чітко відрізняються в порівнюваних групах, були обрані серед великої кількості ознак у жінок з різними результатами екстракорпорального запліднення. Ці параметри використовувалися під час факторного аналізу. За результатами факторного аналізу визначено основні чинники для прогнозування результатів екстракорпорального запліднення жінок із безпліддям методом логістичної регресії. Розроблена модель дозволяє точно прогнозувати результат екстракорпорального запліднення за певними параметрами у 86,93 % випадків.

Ключові слова: додаткові репродуктивні технології, безпліддя, логістична регресія, прогностична цінність

The achieved definite border of efficiency of using the in vitro fertilization (IVF) method does not satisfy the reproductologists’ requirements at achievement of further reduction of frequency of infertility at marriage and, it determines the researchers’ further scientific interest in search of further prognostic criteria of effectiveness of the program of supplementary reproductive technologies (SRT) [2, 3, 5].

There are some determined factors connected to interim and final IVF outcomes. First of all, it is anti-Müllerian hormone (AMH) in the follicular fluid [7].

Presently the most researchers consider the patient’s age as the key predictor of effectiveness of treatment of infertility by means of SRT [1, 8] The results of analysis of the observations inserted to the SRT register of Great Britain (HFEA) submitted by Nelson SM e al. in 2011 evidently demonstrate the dynamics of progressing reduction of possibility of birth of a living child subsequently to IVF and embryo transfer (ET) of the patients older than 35 years old [9].

The reproductive anamnesis data of a married couple are also traditionally considered as potential predictors of effectiveness of IVF and ET. The prognostic significance of existence in anamnesis of pregnancy and viviparity is proved with results of researches [4].

A lot of authors also consider the indicators of the ovarian reserve as factors determining the forecast of the stage of embryos’ implantation and viviparity subsequently to induced pregnancy [6, 7, 10].

Despite of the varying information existing in literature regarding the methods of forecasting the IVF results, all of them have low informational character and are formed not considering a lot of parameters characteristic for infertile women among which the immunological and reproductive factors have important place. Many reproductologists’ dissatisfaction related to the results of the IVF because frequency of onset of pregnancy after application of IVF is not more than 30–35 % even in the best clinics of the world [1].

In this regard, search of new way for improving the effectiveness of the IVF programs is actual.

The purpose of the study was to reveal the predictors of in vitro fertilization outcomes on the basis of the forecasting model indices.

Materials and methods. The researching object was 518 infertile women (initial and secondary infertility) at the age of 20–43 years, directed to the Central Clinical Hospital of Baku city of the Ministry of Health of Azerbaijan for treatment of infertility with IVF/ICSI method. All patients were divided into two groups: the patients with onset pregnancy i.e. positive in vitro fertilization intra cytoplasmic sperm injection results – main group (n=234) and patients with negative in vitro fertilization intra cytoplasmic sperm injection efforts – control group (n=284). The criterion for patients' inclusion to the research was existence of indication for implementation of the IVF program.

An individual map and algorithm of patients' examination were developed for forecasting the IVF results and complication of the IVF procedures (formalized protocol). 160 parameters are exposed to analysis in the analytical map: anamnestic data, anthropometric indicators, the results of the patients' laboratory (biochemical, hormonal, inflectional and hemostasiological) and instrumental examination, the data of gynecological and somatic anamnesis, the results of treatment with IVF method as well as immunological, immunogenetic and genetic examinations.

All qualitative signs are shown in this map in quantitative gradations and, the quantitative indicators are shown in their absolute values.

The acquired data of the clinical and laboratory researches are processed by the variation statistics method in the STATISTICA 10 statistical analysis system (STATISTICA (USA) program package, version 10 for Windows 8). All options contain the character allocation for normality upon criteria of Kolmogorov-Smirnov, Shapiro-Wilkey and Leven. The Student's parametric criterion is used for comparative analysis if allocation is normal or the number of the researching objects is very high (over 100). Mann-Whitney's method was used at allocation of the indicators differing from the normal one.

The impact of patients' separate signs and their weighting coefficients on the IVF results and its complications was researched by means of logistic regression analysis in the STATISTICA 10 (USA) software. The logistic regression is used when the dependent value is binary (i.e. accept value yes/no, has/has not, for instance, positive IVF outcome or negative IVF outcome) and dependent variables of different character (qualitative and/or quantitative) influence its outcome. Possibility of accepting one of the set of assertions under impact of the reviewed signs is factually assessed. Logit of this possibility is the natural logarithm of the relation of the possibility of "positive effect*"* (*p*) to the possibility of "negative effect*"* (1 *p*).

$$\text{Logit (p)} = \ln \frac{p}{1-p}$$

The value of *logit (p)* is permanent and accept value in interval from 0to1 (from negative effect to positive one). The procedure of logistic regression is creation and assessment of the equation of the following type:

Logit (p) $= Y = b_0+b_1x_1+b_2x_2+b_3x_3+\ldots$

X_1, x_2, x_3 – independent variables.

$b_0, b_1, b_2, b_3,$ - permanent coefficients

Whereas, the possibility of positive effect is

$$p = Y = \frac{1}{1+e^{-Y}} = \frac{1}{1+2.72^{-Y}} = A$$

Possibility of negative IVF outcome is equal to "A", and positive IVF outcome – 1–A. Or, possibility of negative IVF outcome is equal to (A×100) %, and positive IVF outcome is [(1-A) × 100] %.

The average values of the parameters acquired during research (at numerous values) or figures 1 (in presence of sign) and 0 (in absence of sign) are compared and Y value is acquired.

Self-descriptiveness of the independent signs is initially discovered for inclusion of the necessary signs characterizing the infertile women to the regression analysis. Informative signs are selected for formation of the classification functions basing upon discovery of precision of the compared signs of patients with positive and negative IVF results.

The contingency table 2×2 and some indicators like sensitiveness and specificity were applied for assessing the efficiency of the linear forecasting equations.

Sensitiveness is a share of persons with positive IVF result of examined patients. This indicator characterizes the possibility of the true positive IVF result. Its values were calculated upon the SENSITIVENESS TP/TP+FN formula.

Specificity is a share of persons with negative IVF result of examined patients. This indicator was calculated upon the SPECIFICITY TP/TP+FN formula.

TP – true positive
FN – false negative
TN – true negative
FP – false positive

The parameters for researching in the regression analysis were selected close to the forecasting predictors or with the factor analysis method in the Statistica 10 (USA) method.

Results of the study and their discussion. The women passing IVF were characterized with a lot of signs like the anthropometric data and anamnestic indicators as well as gynecological, clinical, immunological, immunogenetic and reproductive anamneses, etc.

The factor analysis i.e. method of the main components was used for definition of the women's main signs and determination of their weighting coefficients.

Using this method, all existing signs were joined into several groups – 4 factors. Each factor is characterized by set of signs with relatively high weighting coefficients upon significance. The signs with low weighting coefficient are excluded from the acquired groups.

The analysis of the acquired data discovered 4 main grouping factors upon the principal signs: 1st factor – gynecological anamnesis; 2nd factor – reproductive anamnesis; 3rd factor – embryological analysis and 4th factor – combined analysis.

Passage from the higher number of signs to the lower one is implemented for the purpose of exclusion of the few informative signs, existence of the doubling information and rationality of consolidation of the frequently combining signs into one group.

The symptomatic complex is named as "Gynecological anamnesis" because it has two signs with high weighting coefficient: "outcome of the previous pregnancies at secondary infertility" (0.787), "loss of spontaneous uterine pregnancy" (0.788). This factor explains14.23 % dispersion of the system.

The 2nd factor is named as "Reproductive anamnesis" because this symptomatic complex contains an indicator with high weighting coefficient belonging to the patients' reproductive anamnesis, in particular, "pregnancy subsequently to IVF with viviparity" (–0.87). However, it is noticeable that this factors contains also "Age" (0,72). The load weight bearing of other parameters on this factor is not high. This grouping factor explains 7.85 % dispersion of the system.

The 3rd grouping factor contains these parameters having high weighing parameters: "Number of mature oocytes preceding SRT" (0.73) and "Number of high-quality embryos preceding SRT on the transfer day" (0.72). Consequently, this grouping factor is conditionally named as "Embryologic" and explains 7.36 % of dispersion of the system. The 4th grouping factor named as "Combined" explaining 6.53 % of dispersion of the system contains a lot of signs with low weighing coefficients (mycoplasmosis, age, syndrome of polycystic ovaries (SPO), clamidiosis). Thus, this factor is excluded from the further research. The data of the examined women' IVF outcomes factors are presented in the Table 1.

Table 1

The main components of the examined women' IVF outcomes

Factor 1	*Gynecological anamnesis – 14.23 %*			
	Outcome of the previous pregnancies at secondary infertility: miscarriage1 – one; 2 – 2 and more; 0 – no	0.787	Loss of spontaneous uterine pregnancy: 1 –yes; 0 – no	0.788
Factor 2	*Reproductive anamnesis – 7.85 %*			
	Age	0.72	ECO cycle with viviparity preceding SRT	–0.87
Factor 3	*Embryological – 7.36 %*			
	Number of the mature oocytes preceding SRT	0.73	Number of high-quality embryos preceding SRT on the transfer day	–0.72
Factor 4	*Combined – 6.53 %.*			
	Mycoplasmosis	–0.609	Age	0.30
	SPO	0.27	Clamidiosis	0.36

These parameters are combined in four groups (grouping factors), symptom complexes explaining 35.94 % of the used dispersion of system. Whereas, the greatest power is of 1st factor – grouping factor – "Gynecological anamnesis" determining 14.23 % of variabilities of the established dispersion model, the 2nd grouping factor – Reproductive anamnesis + age with 7.85 % contribution to the variability of the system, 3rd grouping factor – Embryological, responsible for 7.36 % of dispersion of the system, 4th factor – combined – 6.53 %. The weighing coefficients of the signs in the content of the grouping factors are also determined.

Thus, among the entire massive of the researched signs precisely differing between women with positive and negative IVF results with method of factor analysis, upon deduction of some factors, there were acquired the parameters characterizing the women preparing to IVF and deemed as significant for forecasts. The parameters of women having high weighing coefficients acquired subsequently to the factor analysis are used for forecasting the IVF results.

These indicators are “Age”, “Outcome of the previous pregnancies at the secondary infertility-miscarriage” (OP/S); “Loss of spontaneous uterine pregnancy” (LSP); “Pregnancy subsequently to IVF with viviparity” (P/IVF); “Number of the mature oocytes preceding SRT” (NMO); “Number of high-quality embryos preceding SRT on the transfer day” (NHQE).

The model of regression analysis is – positive outcome (1) – Existing pregnancy (EP) at existence of pregnancy and negatie outcome (0), at absence of pregnancy – Absent pregnancy (AP). EP (1) and AP (0) are dependents variables and the aforesaid signs are independent ones.

The equation coefficients are significant at the 5 % level (p-level < 0,05) for indicators of “Outcome of the previous pregnancies at the secondary infertility-miscarriage” with coefficient of 4.702, “Loss of spontaneous uterine pregnancy” – with coefficient of 5.84, “Number of the mature oocytes preceding SRT” with coefficient of 0.22 and “Number of high-quality embryos preceding SRT on the transfer day” with coefficient of 2.28 and “Age” 'with coefficient of 0.72 (Table 2).

Table 2

Coefficients of equation of regression for forecasting the IVF outcomes

Statisical parameters	Factors						
	Absent pregnancy	Outcome of the previous pregnancies at the secondary infertility-miscarriage	Age	Loss of spontaneous uterine pregnancy	IVF cycle with viviparity preceding SRT	Number of mature oocytes preceding SRT	Number of high-quality embryos preceding SRT on the transfer day
Estimate	5.6116	4.702045	0.72659	5.489602	–0.4128	–0.2284	–2.288715
StandardError	1.0137	2.453776	1.457591	2.859676	286.4096	0.088713	0.405077
t(229)	5.5356	1.916249	–0.49848	–1.91965	–0.00144	–2.57492	–5.650073
p-value	0.0000	0.0508761	0.041835	0.050685	0.008851	0.011020	0.0000000
–95 %CL	3.60814	–0.1474647	–2.59019	–11.1413	–566.457	–0.40375	–3.089287
+95 %CL	7.61509	9.551555	9.137018	0.162105	541.631	-0.05310	–1.488142
Wald's Chi-square	30.6432	3.672009	0.248489	3.685091	0.365322	6.63022	31.92332
p-value	0.00000	0.05534235	0.41835	0.054909	0.008851	0.010030	0.0000000
Odds ratioo (unit ch)	273.587	110.1722	3.783554	0.004129	0.661774	0.795781	0.1013967
–95 %CL	36.8975	0.8628929	0.027592	0.000014	0	0.667805	0.0455344
+95 %CL	2028.58	14066.55	8.474132	1.175984		0.948283	0.2257917
Odds ratio (range)		12137.92	0.483554	0.004129	0.661774	0.040842	0.0000001
–95 %CL		0.7445841	0.02759	0.000014	0	0.003508	0.0000000
+95 %CL		197867700	8.474132	1.175984		0.475482	0.0000299

Note: Estimate – free equation member; StandardError – standard error; p-value – significance level of the equation coefficients; –95 % CL; +95 %CL–confidence interval; Wald's Chi-square – Chi-square by Wald; p-value – Wald’s significance level

Thus, it may be concluded that the aforesaid selected variables (predicting signs) impact on the IVF outcome. The result of the regression analysis for forecasting the IVF outcomes is statistically precise (Chi^2=117,80 p=0,0000).

The results of the implemented calculation discovered the following equations of regression: (Equation 1):

$$Y=5.61+4.70\times OPP/M+0.72\times B+5.84\times LSP–0.41PIVF/PV–0.22NMO–2.28\times NHQE$$

Note: OPP/M – Outcome of the previous pregnancies at the secondary infertility-miscarriage; LSP – Loss of spontaneous pregnancy; PIVF/PV – IVF cycle with viviparity preceding SRT; NMO – Number of mature oocytes; NHQE – Number of high-quality embryos; A – Age.

The IVF results of the patients with positive and negative IVF results were additionally checked for determination of sensitivity and specificity of the established equations of regression.

The equation of regression is statistically significant, p<0.001. The aforesaid issue shows that the established forecasting model is able to forecast the IVF outcome upon the aforesaid IVF outcome for 86.93 %. This combination of signs may predict the IVF outcome and it is precise; furthermore, the forecast is quite high. The sensitivity of this forecast is equal to 84.7 % and the specificity of equations is 88.8 %. These indices are able to predict the positive and negative IVF outcomes of infertile women before implementation of IVF.

Thus, basing upon the result of the factor analysis there were assigned the main factors for forecasting the IVF outcomes of infertile women by means of logistic regression. These factors are: outcome of the previous pregnancies at the secondary infertility-miscarriage, loss of spontaneous uterine pregnancy, age, pregnancy subsequently to IVF with viviparity, IVF cycle with viviparity preceding SRT, number of mature oocytes preceding SRT, number of high-quality embryos preceding SRT on the transfer day.

Identification of prognostic informative factors to improve the results of SRT, of course, increases the effectiveness of these technologies and the number of positive outcomes. Therefore, other researchers also are working in this area. So, Kovalenko Ya.A., et al., studying the influence of the age of patients on the number and quality of oocytes obtained, the quality of embryos and the outcomes of treatment after IVF and ET, came to the conclusion that the age of patients affects the quality of oocytes, embryos and the outcome of SRT programs. In patients of the older age group (over 35 years), these figures are significantly reduced. In our study, age was also one of the factors influencing IVF outcome. However, we analyzed a larger number of factors, grouped them and identified the most informative ones based on the regression analysis model, which significantly increases the predictive value [2]**.** Due to logistic regression analysis improve the value of results, some authors use this method for assess the influence of many factors on reproductive outcomes not only in IVF cases. Zhao X, et al revealed that age, severity of intrauterine adhesion, increased menstrual volume, and hysteroscopic adhesiolysis procedures were the dominant factors affecting reproductive outcomes and may be regarded as potential predictors for evaluating intrauterine adhesion prognosis. So, it is obvious that logistic regression analysis may be more useful for this purpose (as we recommended) [12].

Zhang Z, et al. conducted retrospective observational study included 493 cycles underwent IVF/intracytroplasmatic sperm injection aimed to investigate the effects of different factors on the euploidy of blastocysts undergoing IVF. In addition to the factors of our study the authors also assess predictive value of preimplantation genetic testing results. They revealed that the parental karyotype, maternal age, number of cycles, and number of blastocysts per cycle were the dominant factors affecting the ability to achieve at least 1 euploid blastocyst for transfer and therefore could be regarded as potential predictors for genetic counseling [11].

Conclusions

1. The total forecasting value of the established model of forecasting the IVF outcomes upon the shown parameters is equal to 86.93 %.

2. The established forecasting model is able to forecast the IVF outcome upon the aforesaid IVF outcome for 86.93 %. The sensitivity of this forecast is equal to 84.7 % and the specificity of equations is 88.8 %.

3. The created se indices are able to predict the positive and negative IVF outcomes of infertile women before implementation of IVF.

References

1. Yermolenko KS. Radzinskiy VYe, Rapoport SI. Sovremennoye sostoyaniye problemy realizatsii fertilnoy funktsii zhenshchin pozdnego reproduktivnogo vozrasta. Klinicheskaya meditsina.. 2016; 94(1): 10-15. DOI 10.18821/0023-2149-2016-94-1-10-15 [in Russian]

2. Kovalenko YaA, Malko AV, Ryazantsev II, Trunyan DG, Filippov YeF, Krutova VA. Vliyaniye vozrasta patsiyentok na kachestvo poluchayemykh ootsitov, embrionov i iskhodov programm vspomogatelnykh reproduktivnykh tekhnologiy. Kubanskiy nauchnyy meditsinskiy vestnik. 2018;25(1):18–22 [in Russian]

3. Mishchenko VP, Rudenko IV, Likhachov VK, Gromova AM, Tarasenko KV, Mishchenko VP. Profilaktyka nevynoshuvannya vahitnosti shlyakhom zastosuvannya folativ u kompleksniy terapiyi. Svit medytsyny ta biolohiyi. 2021.2 (76). 93–98. http://dx.doi.org/10.26724/2079-8334-2021-2-76-93-98 [in Ukrainian]

4. Bhattacharya S, Maheshwari A, Mollison J. Factors Associated with Failed Treatment: an Analysisof 121,744 Women Embarking on Their First IVF Cycles. PLOS ONE. 2013; 8 (12): 1–13.

5. Dang TT, Phung TM, Le H, Nguyen TB, Nguyen TS, Nguyen TL, et al. Preimplantation Genetic Testing of Aneuploidy by Next Generation Sequencing: Association of Maternal Age and Chromosomal Abnormalities of Blastocyst. Open Access Maced J Med Sci. 2019 Dec 20;7(24):4427–4431. doi: 10.3889/oamjms.2019.875.

6. Iliodromiti S, Kelsey TW, Wu O, Anderson RA, Scott M. Nelson, The predictive accuracy of anti-Müllerian hormone for live birth after assisted conception: a systematic review and meta-analysis of the literature, Human Reproduction Update, July/August 2014; 20(4): 560–570. https://doi.org/10.1093/humupd/dmu003

7. Majumdar G, Majumdar A, Verma IC, Upadhyaya KC. Relationship Between Morphology, Euploidy and Implantation Potential of Cleavage and Blastocyst Stage Embryos. J Hum Reprod Sci. 2017 Jan-Mar;10(1):49-57. doi: 10.4103/0974-1208.204013. Erratum in: J Hum Reprod Sci. 2017 Apr-Jun;10 (2):142–150. PMID: 28479756.

8. 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 Jan 4;8(1):e1000386. doi: 10.1371/journal.pmed.1000386.

9. Wang N. Luteal-phase ovarian stimulation vs conventional ovarian stimulation in patients with normal ovarian reserve treated for IVF: a large retrospective cohort study. Clinical Endocrinology. 2015; 0:1–9.

10. Yao L, Zhang W, Li H, Lin W. The role of serum AMH and FF AMH in predicting pregnancy outcome in the fresh cycle of IVF/ICSI: a meta-analysis. Int J Clin Exp Med. 2015 Feb 15;8(2):1755–67. PMID: 25932104.

11. Zhang Z, Zhang L, Wang Y, Bi X, Liang L, Yuan Y, Su D, Wu X. Logistic regression analyses of factors affecting the euploidy of blastocysts undergoing in vitro fertilization and preimplantation genetic testing. Medicine (Baltimore). 2022 Jul 1;101(26):e29774. doi: 10.1097/MD.0000000000029774.

12. Zhao X, Liu Y, Zhang A, Gao B, Feng Q, Huang H, Zhu X, Sun X, Xu D. Logistic regression analyses of factors affecting fertility of intrauterine adhesions patients. Ann Transl Med. 2020 Feb;8(4):49. doi: 10.21037/atm.2019.11.115.

Стаття надійшла 25.10.2021 р.