

Analyzing Outcomes of Intrauterine Insemination Treatment by Application of Cluster Analysis or Kohonen Neural Networks

Anna Justyna Milewska¹, Dorota Jankowska¹, Urszula Cwalina¹,
Teresa Więsak², Dorota Citko¹, Allen Morgan³, Robert Milewski¹

¹ Department of Statistics and Medical Informatics, Medical University of Białystok, Poland

² Department of Gamete and Embryo Biology, Institute of Animal Reproduction and Food Research of Polish Academy of Sciences, Olsztyn, Poland

³ Shore Institute for Reproductive Medicine, Lakewood, USA

Abstract. Intrauterine insemination (IUI) is one of many treatments provided to infertility patients. Many factors such as, but not limited to, quality of semen, the age of a woman, and reproductive hormone levels contribute to infertility. Therefore, the aim of our study is to establish a statistical probability concerning the prediction of which groups of patients have a very good or poor prognosis for pregnancy after IUI insemination. For that purpose, we compare the results of two analyses: Cluster Analysis and Kohonen Neural Networks. The k-means algorithm from the clustering methods was the best to use for selecting patients with a good prognosis but the Kohonen Neural Networks was better for selecting groups of patients with the lowest chances for pregnancy.

Introduction

Infertility refers to an inability to conceive after having regular unprotected sex for a period of at least one year (Radwan J., 2011). More and more often women are experiencing difficulty becoming pregnant. The female, male or both partners can contribute to the couple's infertility. It has been estimated that in approximately 20–30% of couples, both partners suffer from infertility (Kurzawa et al., 2010). A study conducted by the World Health Organization showed that males might contribute in 50% to these couples' infertility (Radwan et al., 2011).

Usually, at the beginning of any treatment, the male and female are evaluated to establish the reason(s) of infertility. Many tests need to be performed to establish a diagnosis for the couple. During a basic evaluation, the potency of the fallopian tubes and uterus, and concentrations of repro-

ductive and non-reproductive hormones in the blood need to be determined. The quality of male gametes (semen analysis) and ovulatory status of the woman are also checked. Superovulation therapy is prescribed when an ovulatory problem exists. The follicular growth is controlled by the injection of different hormones and measurement of estradiol level in the blood. 75–85% of women will ovulate after such a treatment (Pierzyński, 2011) and half of them will become pregnant within the first 3–4 months if no major infertility problems exist on both sides (female – tubal factor or endometriosis and/or male – abnormal semen parameters) (Pierzyński, 2011). In some circumstances, intrauterine inseminations (IUI) or In Vitro Fertilization (IVF) procedures need to be performed.

The first IUIs were performed at the end of the XVIII century (Horák, 2004; Radwan P., 2011). There are many types of insemination but the most common is intrauterine insemination. It has been established that after 6 IUI inseminations 30–40% of women will become pregnant. However, efficiency of the IUI is in the range of 6–18% per insemination and depends on the type of diagnosis (Radwan P., 2011). IUI is prescribed for the following medical conditions: cervical factors, endometriosis, and male and/or immunological or idiopathic factors (Derwich et al., 2008; Radwan P., 2011). The collected semen samples are purified in the laboratory before IUI and a minimum 5 million motile spermatozoa, usually in the 0.5 ml volume, are deposited in the uterus (Tkaczuk-Włach et al., 2006; Wainer et al., 2004).

The fertility of women drastically decreases after they are forty years old. (Milewski et al., 2008, 2013). Usually a few IUIs are recommended but when the woman is older (around 40 plus years old) IVF is more often suggested as the first choice of treatment (Pierzyński, 2011; Radwan P., 2011). The best outcome with IUI treatment might be obtained: when the woman is younger than 30 years old, superovulation is carried out with gonadotropins, two follicles larger than 16 mm in diameter are present on the ovaries, endometrium thickness is greater than 9 mm and there are more than 5 mln/ml of motile spermatozoa with the forward progression class A and B (Radwan P., 2011).

To be successful in providing treatment to patients, the medical staff has to be knowledgeable, experienced and equipped with adequate tools to make proper diagnoses. The correct diagnosis is essential for designing an efficient treatment plan and analyzing collected medical information appropriately. Classic statistical analyses are lagging behind in the quality of results in such situations. Therefore, other more efficient statistical analysis methods, such as Data Mining, should be employed because they give

an opportunity for the creation of suitable predictable algorithms. Medical databases are becoming larger and larger. This allows for accommodation of more sophisticated statistical analyses, which lead to the creation of precise treatment options. For example, to predict efficacy of the IVF with embryo transfer procedure – the Artificial Neural Networks (ANNs) (Milewski et al., 2009), Correspondence (Milewska et al., 2012) or Basket (Milewska et al., 2011) or types of analysis with the application of feature selection algorithms to reduce original dimensionality of the original data set (Milewski et al., 2011, 2012) are applicable. Therefore, the aim of our study is to provide a statistical probability of successful IUI treatment outcomes by grouping patients appropriately (with good or poor prognosis).

Material and Methods

The medical information without personal identifiers of the 825 IUI cycles performed at the Shore Institute for Reproductive Medicine, Lakewood, USA was used in these statistical analyses. Segmentation methods such as Cluster Analysis or Kohonen Neural Networks were applied because they provide an option for uniformly grouping data to determine/estimate percentages of successful pregnancies. The quantitative variables: semen parameters, hormone levels, age of the female (Table 1) and qualitative variables: reason for infertility (Table 2) were analyzed. The Statistica Data Miner + QC 10.0 (StatSoft, Tulsa, OK, USA) software was used. Statistical significance was determined at the $p < 0.05$ level.

Table 1. List of quantitative variables

Quantitative variables	median	min	max
semen – number of motile sperm	17	0.02	614.8
semen vol. – volume of semen – ml	3	0.2	12
sperm ct. – concentration sperm – M/ml	47	1.4	598
sperm mot. – motility sperm – %	66	2	100
age – age of women	35	23	46
no. of follicles – total number of ovulatory follicles per cycle	7	1	46
endometrium thickness at HCG injection – mm	10	5	24
E ₂ at HCG – estradiol level at HCG injection – pg/ml	470	89.3	2624
P ₄ at HCG – progesterone level at HCG injection – ng/ml	1	0.2	8.8

Table 2. List of qualitative variables used in analysis

Qualitative variables	n	%
clinical pregnancy (variable result)	98	11.9%
infertility reason – idiopathic factor	439	53.0%
infertility reason – AMA – Advanced Maternal Age	126	15.3%
infertility reason – endometriosis	62	7.5%
infertility reason – MF – Male Factor	41	5.0%
infertility reason – ovulatory factor	67	8.1%
infertility reason – PCOS – polycystic ovary syndrome	36	4.4%
infertility reason – tubal factor	29	3.5%
infertility reason – secondary infertility	17	2.1%
infertility reason – uterine factor	23	2.8%
stimulation drug – CC – clomiphene citrate	387	46.9%
stimulation drug – I – gonadotropins	352	42.7%
stimulation drug – TX – tamoxifene	86	10.4%

Cluster Analysis

Cluster analysis refers to the methods of organizing data according to certain structures. Basically, it is a process of identifying groups of objects similar to each other in some characteristics but distinctively different from elements in other groups. This indirect technique is ranked among unsupervised learning methods. The variables that decisively determine an observational group are not defined. Cluster Analysis is applicable in exploratory data mining because it allows us to reduce the sizes of enormous databases and to organize information for easy access. It also allows us to discover existing relationships, for example, the relationship between a patient's biochemical parameters and occurrence of illness (McLachlan, 1992).

Algorithm Cluster Analysis groups the objects that are more similar to each other but differ in some way from elements in other clusters. Either cases or variables, which characterize a study group, can be grouped into classes. The distances among them are determined to estimate similarity and/or dissimilarity between the objects $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$. The following metrics can be used:

- Euclidean distance – the most natural geometric distance in a multidimensional space.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

- Squared Euclidean distance – usually used to provide higher weight to distant objects.

$$d(x, y) = \sum_{i=1}^n (x_i - y_i)^2$$

- Chebyshev distance – useful in demonstrating differences among the objects in one dimension that differentiate elements the most.

$$d(x, y) = \max_{i=1, \dots, n} |x_i - y_i|$$

- Manhattan distance (City block) – a sum of differences between the objects in all dimensions (in this metric sphere it is the surface area of the cube).

$$d(x, y) = \sum_{i=1}^n |x_i - y_i|$$

- Exponential (Power) distance – allows the weights that are placed on the differences between objects in individual dimensions to be guided.

$$d(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^a \right)^{\frac{1}{b}}$$

where a, b are the parameters established by the researcher

- Percent disagreement – used in situations where variables are qualitative.

$$d(x, y) = \frac{\text{number of pair of variables such as } x_i \neq y_i}{\text{total number of pair of variables}}$$

$$i = 1, 2, \dots, n$$

Groups of algorithms called the Cluster Analysis can be divided into hierarchical and nonhierarchical subgroups. Among the hierarchical methods are the agglomerative and divisive procedures. The agglomerative algorithm is called a “bottom-up” approach. Basically, it assumes that initially each object is a separate cluster and new clusters are formed by combining the nearest objects and groups, which were joined earlier. Contrary to the previously described bottom up method, the top down approach or divisive method relies on the gradual splitting of clustered content into single elements (Stanisz, 2007). In these methods there is no need to define the number of clusters for the analyzed objects up front but an enormous computer processing power is essential. The applications of the agglomerative analyses are more frequently used and run according to the algorithm presented in diagram form in Figure 1 (Timm, 2002).

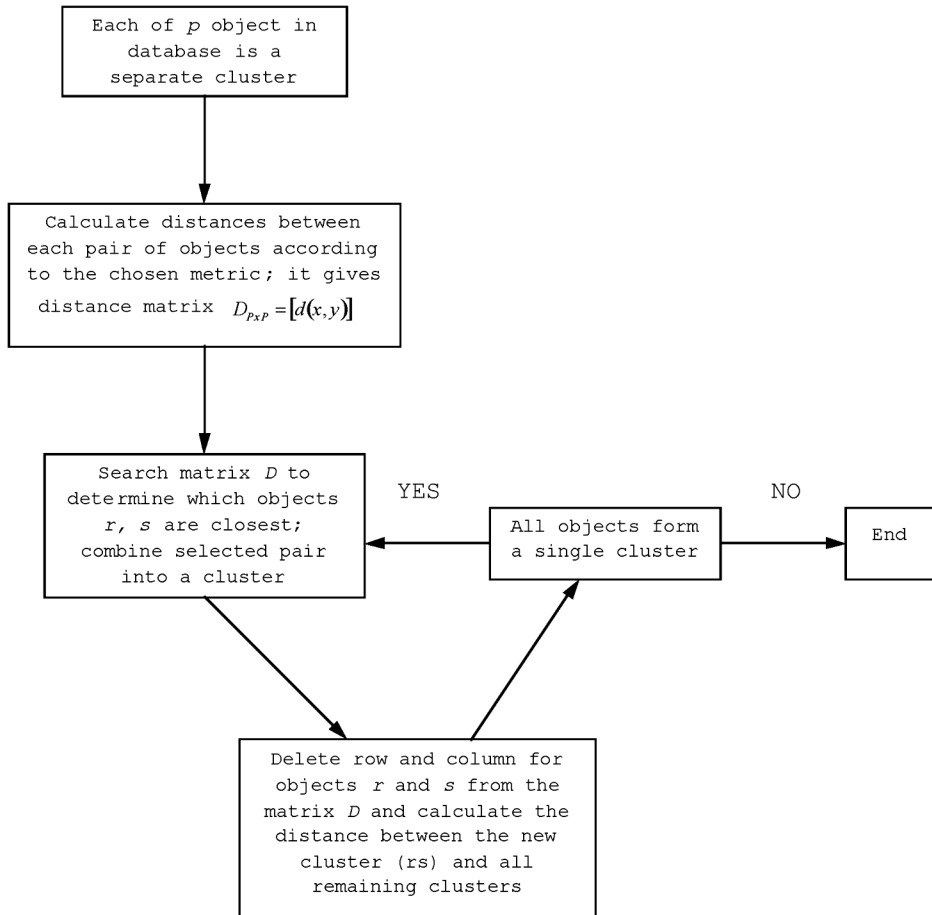


Figure 1. The flowchart of the agglomerative hierarchical clustering procedure

The method of defining distances between the clusters is the major contributing factor to the quality of classification and subsequently the quality of the whole analysis. The following techniques are applied very often (Stanisz, 2007):

- Single Linkage Clustering – also called the Nearest Neighbor Method; it is a distance between the analyzed clusters that is equivalent to the distance between the two nearest objects from two different clusters.
- Complete Linkage Clustering (Farthest Neighbor) – the distance between the clusters is defined as a distance between the two farthest objects in different clusters. This method is in contrast with the single linkage clustering method.

- Average Link Method – UPGMA (unweighted pair-group method using arithmetic averages); distance between clusters is determined as an average of the distances between all pairs of objects from two different clusters.
- Weighted Average Link Method – WPGMA (weighted pair-group method using arithmetic averages); this algorithm is recommended in situations where clusters possibly differ in number from each other. This method is more advanced than the previous one because it uses weights related to the number of elements in clusters.
- Centroid Method – UPGMC (unweighted pair-group method using the centroid average); distance between the clusters is estimated as a distance between their centers of mass.
- Weighted Centroid Method (Median Linkage) – WPGMC (weighted pair-group method using centroid average); determination of the distance is boosted with the weight that accounts for the disproportion in numbers of elements between clusters.
- Ward's Method (Incremental Sum of Squares) – this algorithm minimizes variability within the cluster; among all possible connections of pairs of objects, the one that characterizes the minimal variability is chosen – so this method minimizes the square deviations sum inside the cluster.

The vertical or horizontal tree diagram (dendrogram) is a graphic presentation of hierarchical clustering method results. Figure 2 represents a sample of such a dendrogram.

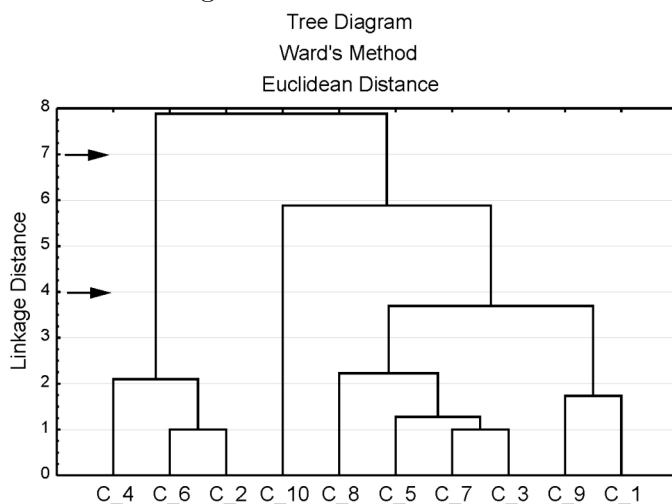


Figure 2. An example of a dendrogram, created as a result of Cluster Analysis using Ward's method

The objects are on OX axes that were initially assigned as single clusters. The most similar objects are linked by gradually diminished criteria of similarity (Stanisz, 2007). Subsequently, the clusters are created by linking more and more different objects. Finally, one cluster is created. The distances at which respective elements were merged into a single new cluster are on the OY axis. The dendrogram allows us to make a decision on the number of groups to be analyzed by cutting the tree diagram at the appropriate height.

The other Cluster Analysis methods are the nonhierarchical analyses, such as the k-means method or Expectation Maximization Method (EM). The principle of these methods is to disperse a group of objects into a known number of separate clusters in such a way that none of those groups are not subgroups of others. In the k-means method, the number of the clusters k has to be determined by the statistician at the beginning of the analysis (for example based on his/her intuition and experience). However, performing v -fold cross-validation is recommended to optimize the analysis. The algorithm divides objects in the study groups into more and more segments while at the same time checking the precision of division for each of them. The precision criterion for the classification is the average of the distances of each object from the center of its own cluster. The number of clusters is chosen such that the increase of precision in divisions is still very distinct. Before conducting analyses using the k-means method, it is necessary to define how the first cluster centers will be chosen. Usually, they are the first k objects. The cluster centers can be chosen using the principle of maximizing initial distances between clusters or by sorting all the distances between objects choosing elements with constant intervals as a center. After setting up initial assumptions, the algorithm runs according the diagram presented in Figure 3. The core of performance k-means is to move objects between clusters to minimize variation inside segments and to maximize variation between the clusters. This algorithm runs towards maximization of the significance of the F -test results in the analysis of variance. The statistical F value in a particular dimension decides on the role of particular variable during creation of clusters (Stanisz, 2007).

The Expectation Maximization Method (EM) is needed at the beginning of the analysis to define the number of segments for the initial data collection. Segregation of objects into separate k clusters is performed based on the principle that distribution of each analyzed variable is a mixture of k distributions. The basis for this method is first, to determine parameters of each component of distribution (such as mean or standard deviation),

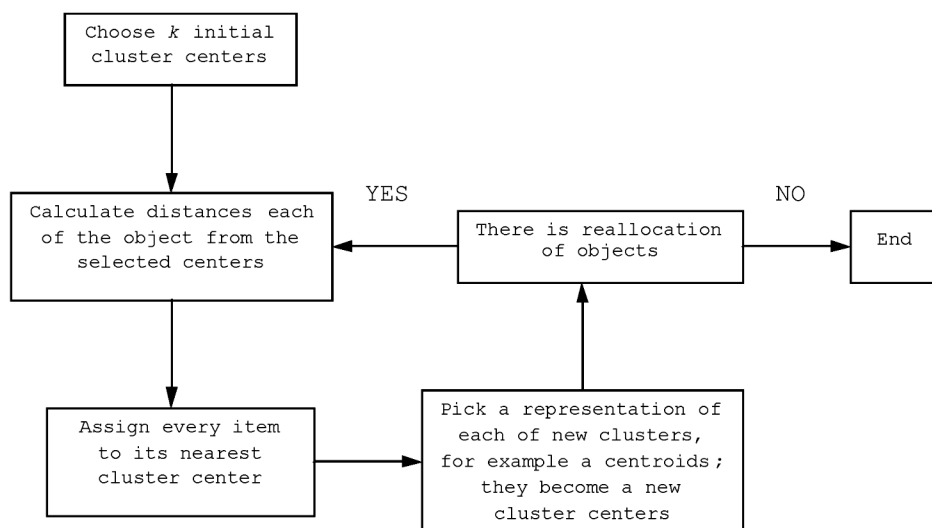


Figure 3. The flowchart of k-means method clustering

then the probability of assigning objects to those particular created clusters is determined. Finally, the objects are assigned to the cluster with their highest probability of belonging to that cluster.

Cluster Analysis is primarily used in analyzing data in the areas of biology, medicine and/or bioinformatics (Xu et al., 2010). The segmentation of data has a wide application and is very useful in: classification of plants or animals, psychological studies to select personalities (Sava et al., 2011), artificial intelligence studies, analyzing results of tomographic studies (distinction of tissue types and blood in the three-dimensional pictures (DeLa-paz et al., 1990)) or in computer-aided diagnoses that assist doctors in the interpretation of medical images. Additionally, Cluster Analysis is applied to analyze microarray results (Shannon et al., 2003) because it allows one to determine groups of genes with similar patterns of expression or to establish similarity in the genotypes (Eisen et al., 1998).

Kohonen Neural Networks

Self-Organizing Feature Maps, also called Kohonen Neural Networks (Kohonen, 1982), are one of the basic self-organizing networks. The Kohonen Neural Networks are an example of unsupervised learning. This means that data are entered into a database without an established earlier pattern. It is similar to the way the human brain functions. Patterns are cre-

ated during the learning process. Such an approach is very useful in situations where new relationships that are not detectable using traditional statistical methods need to be found. Kohonen Neural Networks are able to demonstrate new output patterns, which can be identified with earlier unrecognized relationships. Thus, it allows statisticians to better understand data and, subsequently, to be able to apply different tools in further analysis. This method groups cases or creates grouped cases that are quite similar to each other or similar in their characteristics. At the same time, the groups should be as different as possible from each other. Kohonen Networks is a competitive learning method where the winning neuron (the most similar to the input vector) is chosen in competition with other neurons.

The Kohonen Network is built from two layers: an input layer and an output layer that is built out of neurons. The basic algorithm of Kohonen Networks runs in an interactive way. At the beginning, the nodes are chosen randomly and afterward they go through multiple runs according to the plan in Figure 4.

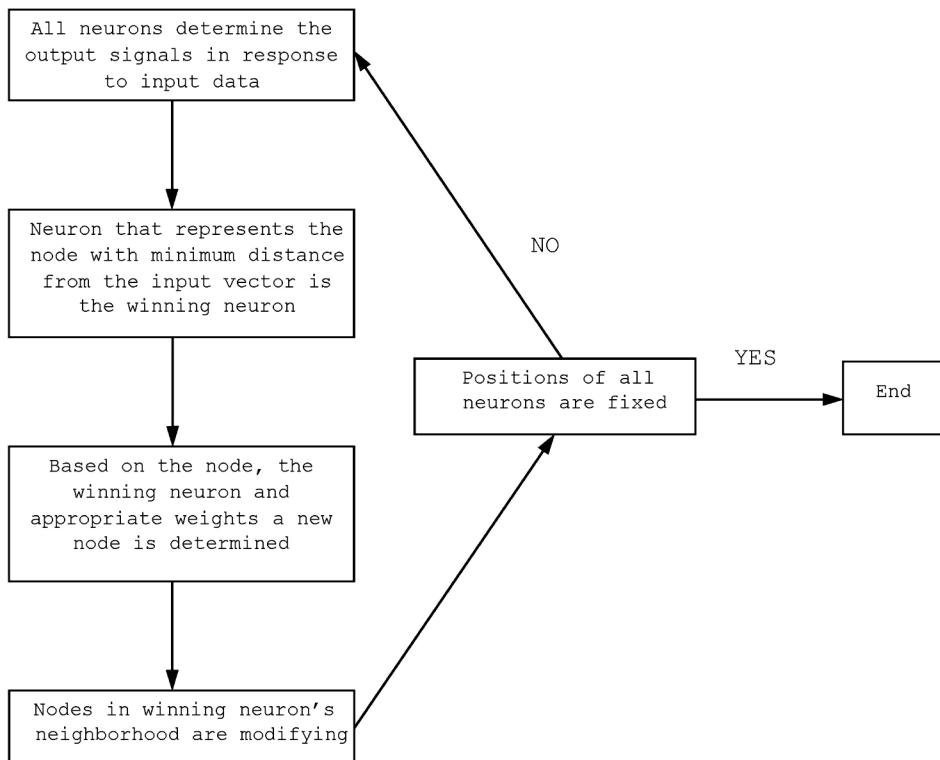


Figure 4. Scheme of Kohonen algorithm

During algorithm runs, the learning coefficient diminishes, affecting vector weights. At the beginning, the changes are usually large but with time they become smaller and smaller. The algorithm tests successive input factors and assigns the most appropriate nodes to them. Next, each node in the neighborhood is modified to resemble the element of the training set. At first an algorithm creates an approximate topological map, which at the end shows neurons that correspond with small clusters. Neurons in Kohonen Network are not connected with each other but they are presented as two-dimensional grids of nodes. This makes interpretation easy and/or allows one to observe similarities between the clusters with ease.

The Kohonen network is usually a one-way network and each neuron is connected to all elements of input vector X . The process starts from the normalization of the N-dimension input vector according to the equation below.

$$x'_i = \frac{x_i}{\sqrt{\sum_{y=1}^N (x_y)^2}}$$

After entering the input vector, the neurons compete with each other. The winning neuron w_w is in the smallest distance to X , so it complies with the following equation:

$$d(x, w_w) = \max_{1 \leq i \leq n} d(x, w_i)$$

Where weight vector is:

$$w_i = \begin{bmatrix} w_{i1} \\ w_{i2} \\ \vdots \\ w_{iN} \end{bmatrix}$$

$d(x, w)$ is a distance between the vectors x and w according to the selected metric. The Kohonen algorithm uses the concept of topological neighborhood neurons w_w . In the content of the neighborhood (with diminishing radius in time) neurons surrounding w_w are included. The winning neuron and neighbor neurons are subjected to adaptation according to Kohonen's formula (Kohonen, 1995):

$$w_i(n+1) = w_i(n) + \eta_i(n)[x - w_i(n)]$$

where η is coefficient factor i-of that neuron in neighborhood $Sw(n)$ in k-time. The value of that coefficient factor decreases with the distance to the winner. The basic Kohonen learning algorithm is:

$$w_i(n+1) = w_i(n) + \eta G(i, x)[x - w_i(n)]$$

Kohonen (Kohonen, 1995) proposed two types of neighborhood:

- rectangular (based on Euclid metric)

$$G(i, x) = \begin{cases} 1 & \text{dla } d(i, w) \leq \lambda \\ 0 & \text{dla } d(i, w) > \lambda \end{cases}$$

- Gaussian

$$G(i, x) = \exp\left(-\frac{d^2(i, w)}{2\lambda^2}\right)$$

Kohonen Neural Networks have a wide application where grouping objects into clusters with similar characteristics allows industrial or diagnostic processes to be improved. An interesting example of an application of Kohonen Networks is attempt the prognosis of daily usage of water (Licznar et al., 2006). This method might be used in marketing (Migut, 2009), as grouping clients with similar habits and preferences plays a key role in designing a marketing strategy. Furthermore, Kohonen Networks are useful in analyzing medical pictures (Ahmed et al., 1997).

Results

The 5 clusters (statistically different $p = 0.01$ in terms of pregnancy percentage) were outlined using k-means. The most efficient treatment was in cluster IV with a more than 27% pregnancy rate. However, the lowest pregnancy rate (only 7%) was in cluster II. In the other clusters the pregnancy rates were close to the mean value of allover analyzed cycles and were in the range of 10–13%. Moreover, in cluster IV, 59% of women were diagnosed with PCOS and 23% with endometriosis, and 7% of men were diagnosed with male factor. There were no cycles with idiopathic and tubal factors. In spite of male factor presence, all the semen parameters were above the median and 61% of patients were stimulated with gonadotropins. The median (30.5 years) for the age was the lowest among the studied clusters but the endometrium thickness was the highest – 11 mm (Table 3).

The lowest pregnancy success was in cluster II, probably due to fact that 84% of women were diagnosed with advanced maternal age (median for the age was 42 years old). The male factor diagnosis was not present in this cluster. However, the sperm concentration and motility were similar to those in cluster IV. Sixty-six percent of patients were stimulated with gonadotropins and the median for the estradiol (E_2) was 560 pg/ml. It was the highest value in comparison to other groups/clusters.

Table 3. Clusters characteristics obtained with k-means method

clusters	I	II	III	IV	V
frequency	249	114	72	44	97
clinical pregnancy	12.9%	7.0%	11.1%	27.3%	10.3%
Qualitative variables (percent)					
idiopathic factor	90.4%	0.9%	87.5%	–	–
AMA	–	84.2%	1.4%	–	2.1%
endometriosis	4.4%	4.4%	4.2%	22.7%	16.5%
MF	3.2%	–	–	6.8%	13.4%
ovulatory factor	–	–	–	9.1%	57.7%
PCOS	–	–	–	59.1%	–
tubal factor	0.8%	6.1%	1.4%	–	11.3%
secondary infertility	–	2.6%	2.8%	18.2%	–
uterine factor	–	3.5%	4.2%	2.3%	4.1%
CC	30.9%	23.7%	20.8%	15.9%	69.1%
I	58.2%	65.8%	73.6%	61.4%	23.7%
TX	10.8%	10.5%	5.6%	22.7%	7.2%
Quantitative variables (median)					
sperm no.	15.3	17.0	43.7	28.9	11.4
sperm vol.	2.8	2.9	2.2	3.0	3.0
sperm ct.	42.2	48.5	120.5	48.8	33.2
sperm mot.	60.0	73.5	86.0	75.0	60.0
age	34.0	42.0	34.0	30.5	31.0
no. of follicles	6.0	6.0	33.5	8.5	6.0
endom. thickness	10.0	10.0	9.0	11.0	9.0
E ₂ at HCG	463.0	560.0	491.5	467.0	415.0
P ₄ at HCG	0.9	0.9	0.8	0.9	0.9

The application of the hierarchical agglomerative clustering method produced the most optimal separation of objects into 5 clusters (Figure 5, Table 4). The pregnancy rate (%) among the clusters was not statistically different and was in the range of 9–15%. In cluster III, the pregnancy rate was the highest (15%) and medians of other characteristics were higher than in other clusters: age of women (36 years of age), endometrium thickness (11 mm), ovulatory follicle numbers (12), and estradiol level in the blood (1316 pg/ml). However, the medians for sperm concentration and motility were lower in comparison to other clusters. Additionally, there were almost 41% of women with idiopathic factors and more than 20% with advanced maternal age diagnosis. Cluster V was characterized by the lowest

Table 4. Clusters characteristics obtained with agglomerative method

clusters	I	II	III	IV	V
frequency	151	133	59	155	78
clinical pregnancy	10.6%	14.3%	15.3%	12.3%	9.0%
Qualitative variables (percent)					
idiopathic factor	50.3%	46.6%	40.7%	53.5%	56.4%
AMA	14.6%	17.3%	20.3%	16.8%	20.5%
endometriosis	9.3%	9.8%	5.1%	9.7%	–
MF	6.0%	3.8%	1.7%	4.5%	2.6%
ovulatory factor	12.6%	15.0%	2.4%	7.7%	9.0%
PCOS	5.3%	3.8%	11.9%	2.6%	2.6%
tubal factor	2.0%	3.8%	6.8%	2.6%	6.4%
secondary infertility	0.7%	2.3%	5.1%	1.3%	5.1%
uterine factor	2.0%	0.8%	5.1%	3.2%	–
CC	44.4%	33.8%	13.6%	34.8%	24.4%
I	27.2%	55.6%	86.4%	63.2%	75.6%
TX	28.5%	10.5%	–	1.9%	–
Quantitative variables (median)					
sperm no.	15.5	26.1	16.0	19.8	14.0
sperm vol.	2.5	2.6	2.8	2.7	2.8
sperm ct.	44.6	49.2	41.0	49.9	59.1
sperm mot.	66.0	67.0	63.0	66.0	70.0
age	33.0	35.0	36.0	35.0	35.0
no. of follicles	6.0	6.0	12.0	7.0	9.0
endom. thickness	10.0	9.2	11.0	9.0	10.0
E ₂ at HCG	238.0	379.0	1316.0	590.0	890.0
P ₄ at HCG	0.8	0.8	1.2	0.9	1.0

pregnancy rate (only 9%). Cluster V contained women with idiopathic factors (56%) and with advanced maternal age diagnosis (20%). For 76% of patients, ovulation induction was with gonadotropins. The median age was 35 years. The sperm concentration and motility were the highest in comparison to the other clusters.

After running multiple analyses, the most optimal network that contained 9 clusters was chosen (Figure 6, Table 5). The pregnancy rate was in the range of 9–18% and there were no statistical differences between the clusters. In cluster IV, the pregnancy rate was the highest (18.2%) and 59% of women were with secondary infertility (they were all of the women with secondary infertility), the male factor contributed in 41% of

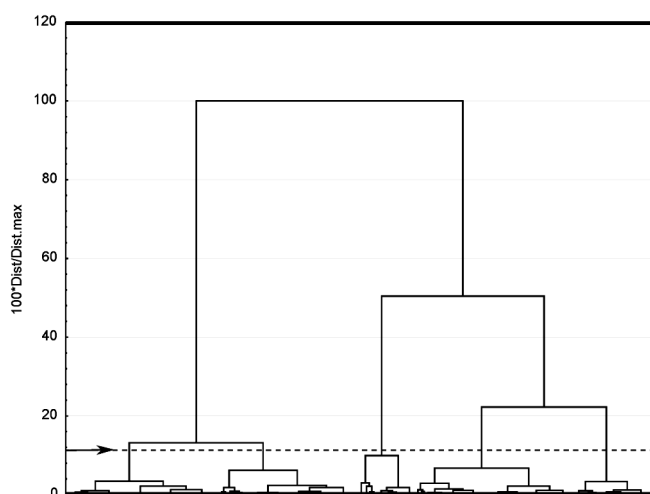


Figure 5. Dendrogram presenting agglomeration of cases

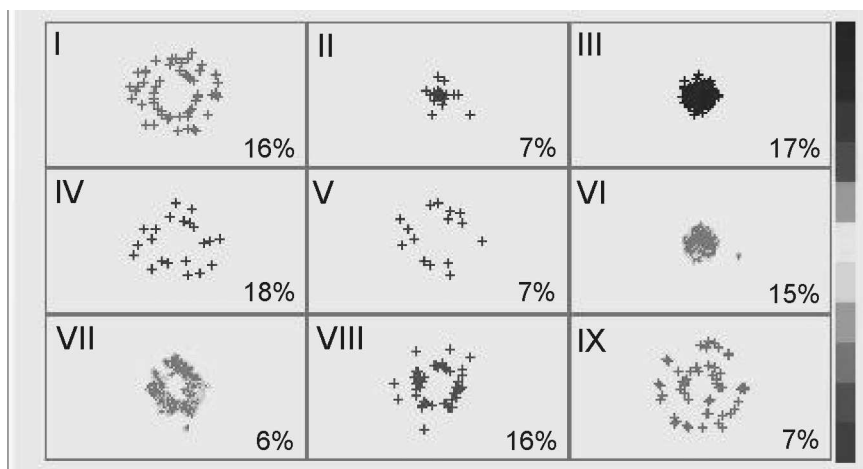


Figure 6. Clusters obtained with Kohonen Networks

the cases (the median for sperm concentration was 45 million/ml and for motility was 72%), the median age of female was 32 years old, 77% of patients were stimulated with gonadotropins and the median for the number of ovulatory follicles/oocytes was 11. The lowest pregnancy rate (6.2%) was in cluster VII with the AMA as a major infertility diagnosis. The median for the female age was 42 years of age and it was the highest among the studied clusters. The median number of ovulatory follicles was 5 after stimulation.

Table 5. Cluster characteristics obtained with Kohonen Networks

clusters	I	II	III	IV	V	VI	VII	VIII	IX
frequency	56	28	169	22	15	92	95	45	54
c. pregnancy	16.1%	7.1%	16.6%	18.2%	6.7%	15.2%	6.2%	15.6%	7.4%
Qualitative variables (percent)									
idiopathic f.	–	100%	100%	–	–	100%	–	–	–
AMA	–	3.6%	–	–	–	1.1%	100%	–	3.7%
endometriosis	–	–	–	–	–	–	–	100%	–
MF	–	–	–	40.9%	–	–	–	8.9%	20.4%
ovulatory f.	48.2%	–	–	–	–	–	–	–	61.1%
PCOS	26.8%	–	–	–	60%	–	–	4.4%	–
tubal factor	17.8%	–	–	–	–	–	–	–	20.4%
secondary i.	–	–	–	59%	–	–	–	–	–
uterine f.	8.9%	–	–	–	40%	–	1.1%	–	–
CC	–	–	–	–	60%	100%	28%	35.6%	90.7%
I	100%	–	100%	77.3%	–	–	61%	51.1%	–
TX	–	100%	–	22.7%	40%	–	10.5%	13.3%	9.3%
Quantitative variables (median)									
sperm no.	17.1	28.4	23.2	25.6	25.3	17.4	15.8	12.8	8.9
sperm vol.	3.0	2.7	2.5	2.0	2.5	2.5	2.8	2.7	3.0
sperm ct.	42.0	60.6	68.0	45.4	51.0	50.3	50.0	29.9	32.7
sperm mot.	72.0	77.5	65.0	72.5	72.0	69.0	72.0	59.0	55.0
age	32.0	33.0	34.0	32.0	29.0	34.0	42.0	33.0	33.5
no. of follicles	12.0	4.0	9.0	10.5	5.0	5.0	5.0	6.0	6.0
endom. thick.	11.0	10.0	10.0	10.5	10.0	9.0	10.0	10.0	8.2
E ₂ at HCG	653.0	249.0	605.0	578.0	204.0	390.0	532.0	406.0	361.5
P ₄ at HCG	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.8	0.9

The idiopathic diagnosis was present in clusters II, III and VI; the highest (16.6% and 15.2%) pregnancy rates were in clusters III and VI, respectively, but in cluster II it was 7.1%. Generally, the medication used for stimulation was different among the clusters; in cluster II all women were stimulated with Tamoxifene (TX), whereas in cluster III they were stimulated with gonadotropins and with clomiphene citrate (CC) in cluster VI. The median ages of females in all of these clusters were quite similar (approx. 33–34 years of age).

According to the the Kohonen algorithm of creation of Neural Networks, the neighboring clusters should be similar to each other. Figure 6 demonstrates the similarities between clusters I and IV in terms of pregnancy rate as well as the significant differences between clusters I and II.

Discussion and Conclusions

The k-means method seems to be the best, in comparison to the other two segmentation methods, in demonstrating the percentage of pregnancies achieved in IUI treatment. The pregnancy rate for the analyzed 825 cycles was 11.9%. However, the cluster with the highest pregnancy rate (27.3%) included the young women with the best chances for successful outcome of IUI treatment and all of the women with PCOS. Only the k-means method was able to demonstrate statistical differences in the pregnancy rates between the studied clusters.

In contrary to the k-means method, the agglomerative method was not able to demonstrate a high pregnancy rate in any of the studied clusters. The clusters were quite similar to each other; therefore pregnancy rates were comparable (9–15%) among them. The application of Kohonen Neural Networks into analysis of our data also did not produce the anticipated results. The internal structure of clusters was more varied here than in other methods. It is evident that Kohonen Neural Networks grouped predominantly qualitative variables. The highest pregnancy rate was 18.2% for that method. It is about 9% lower than using the k-means method. However, the Kohonen Neural Networks was able to detect the lowest chances for pregnancy and it was at the level of 6%.

To conclude – the k-means algorithm from the clustering methods was the best method for selecting patients with good prognosis but the Kohonen Neural Networks was better in selecting groups of patients with the lowest chances for achieving pregnancy.

R E F E R E N C E S

- Ahmed, M. N., & Farag, A. A. (1997). Two-stage neural network for volume segmentation of medical images. *Pattern Recognition Letters*, 18, 1143–1151.
- DeLapaz, R. L., Herskovits, E., Di Gesu, V., Hanson, W. J., & Bernstein, R. (1990). Cluster analysis of medical magnetic-resonance images data: diagnostic application and evaluation. *Proceedings of SPIE*, 1259, Extracting Meaning from Complex Data: Processing, Display, Interaction, 176. DOI: 10.1117/12.19984.
- Derwich, K., Jędrzejczak P., & Pawelczyk, L. (2008). Metody wspomaganego rozrodu. In Z. Słomko (Eds.), *Ginekologia* (pp. 516–532). Warszawa: PZWŁ.
- Eisen, M. B., Spellman, P. T., Brown, P. O., & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America*, 95(25), 14863–14868.

- Horák, S. (2004). Insemination – indications, methods and efficiency. *Ginekologia Praktyczna*, 12(6), 41–49.
- Kohonen, T. (1982). Self-Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics*, 43, 59–69.
- Kohonen, T. (1995). *Self-Organizing Maps*. Springer.
- Kurzawa, R., Kaniewska, D., & Bączkowski, T. (2010). Infertility from clinical and social perspective. *Przewodnik Lekarza*, 2, 149–152.
- Licznar, P. & Łomotowski, J. (2006). Zastosowanie sieci neuronowych Kohonena do prognozowania dobowego poboru wody, *Ochrona Środowiska*, 28, 45–48.
- McLachlan, G. J. (1992). Cluster analysis and related techniques in medical research. *Statistical Methods in Medical Research*, 1(1), 27–48.
- Migut, G. (2009). Zastosowanie technik analizy skupień i drzew decyzyjnych do segmentacji rynku. *StatSoft Polska*. Retrieved from: http://www.statsoft.pl/czytelnia/artykuly/Zastosowanie_teknik.pdf.
- Milewska, A. J., Górka, U., Jankowska, D., Milewski, R., & Wołczyński, S. (2011). The use of the basket analysis in a research of the process of hospitalization in the gynecological ward. *Studies in Logic, Grammar and Rhetoric. Logical, Statistical and Computer Methods in Medicine*, 25(38), 83–98.
- Milewska, A. J., Jankowska, D., Górka, U., Milewski, R., & Wołczyński, S. (2012). Graphical representation of the relationships between qualitative variables concerning the process of hospitalization in the gynecological ward using correspondence analysis. *Studies in Logic, Grammar and Rhetoric. Logical, Statistical and Computer Methods in Medicine*, 29(42), 7–25.
- Milewski, R., Jamiołkowski, J., Milewska, A. J., Domitrz, J., Szamatowicz, J., & Wołczyński, S. (2009). Prognosis of the IVF ICSI/ET procedure efficiency with the use of artificial neural networks among patients of the Department of Reproduction and Gynecological Endocrinology. *Ginekologia Polska*, 80(12), 900–906.
- Milewski, R., Malinowski, P., Milewska, A. J., Czerniecki, J., Ziniewicz, P., & Wołczyński, S. (2011). Nearest neighbor concept in the study of IVF ICSI/ET treatment effectiveness. *Studies in Logic, Grammar and Rhetoric. Logical, Statistical and Computer Methods in Medicine*, 25(38), 49–57.
- Milewski, R., Malinowski, P., Milewska, A. J., Ziniewicz, P., Czerniecki, J., Pierzyński, P., & Wołczyński, S. (2012). Classification issue in the IVF ICSI/ET data analysis. *Studies in Logic, Grammar and Rhetoric. Logical, Statistical and Computer Methods in Medicine*, 29(42), 75–85.
- Milewski, R., Milewska, A. J., Czerniecki, J., Leśniewska, M., & Wołczyński, S. (2013). Analysis of the demographic profile of patients treated for infertility using assisted reproductive techniques in 2005–2010. *Ginekologia Polska*, 84(7), 609–614.
- Milewski, R., Milewska, A. J., Domitrz, J., & Wołczyński, S. (2008). In vitro fertilization ICSI/ET in women over 40. *Przegląd Menopauzalny*, 7(2), 85–90.

- Pierzyński, P. (2011). *Zajść w ciążę*. Białystok: CMR.
- Radwan, J. (2011). Badanie nieplodnej pary. In J. Radwan & S. Wołczyński (Eds.), *Nieplodność i rozród wspomagany* (pp. 47–66). Poznań: Termedia.
- Radwan, J., Krasinski, R., & Gruszczyński, W. (2011). Badanie nasienia. In J. Radwan & S. Wołczyński (Eds.), *Nieplodność i rozród wspomagany* (pp. 67–80). Poznań: Termedia.
- Radwan, P. (2011). Inseminacja domaciczna. In J. Radwan & S. Wołczyński (Eds.), *Nieplodność i rozród wspomagany* (pp. 165–178). Poznań: Termedia.
- Sava, F. A., & Popa, R. I. (2011). Personality types based on the big five model. A cluster analysis over the Romanian population. *Cognition, Brain, Behavior. An Interdisciplinary Journal*, 15(3), 359–384.
- Shannon, W., Culverhouse, R., & Duncan, J. (2003). Analyzing microarray data using cluster analysis. *Future Medicine*, 4(1), 41–52.
- Stanisz, A. (2007). *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. T. 3. Analizy wielowymiarowe*. Kraków: StatSoft.
- Timm, N. T. (2002). *Applied Multivariate Analysis*, Springer.
- Tkaczuk-Włach, J., Robak-Chołubek, D., & Jakiel, G. (2006). Male infertility. *Przegląd Menopauzalny*, 5, 333–338.
- Wainer, R., Albert, M., Dorion, A., Bailly, M., Berge're, M., Lombroso, R., Gombault, M., & Selva, J. (2004). Influence of the number of motile spermatozoa inseminated and of their morphology on the success of intrauterine insemination. *Human Reproduction*, 19(9), 2060–2065.
- Xu, R., & Wunsch, D. C. 2nd. (2010). Clustering algorithms in biomedical research: a review. *IEEE In Biomedical Engineering*, 3, 120–154.