Early Detection of the most Important Factors Leading to Breast Cancer

Aseel Mubder Dawood Al-Marsuomy[1], Mustafa Mohammed Subhi Fadhel[2]
[1] Department of Computer Systems, Middle Technical University/Institute of Administration-Rusafa,
Almarsoumy66@yahoo.com
[2]Department of Banking and Financial Sciences Uruk University/College of Administration and Economics,
Mustafa_mms_1986@yahoo.com.

Abstract:

The aim of this research is to determine the most important and main factors that lead to breast cancer. It is also about finding suitable solutions to eradicate these factors and avoiding them in order to prevent getting breast cancer.

In order to achieve the aim of this research, a sample of (40) patients from Medical City - Oncology Teaching Hospital where a questionnaire was used to collect data. It contained (17) reasons. The statistical programmed (SPSS) was used to compare the results of the data analysis using two methods (Radial Bases Function Network) and (Factorial Analysis). The research came up with important results; the two methods determined the same factors that could be a direct reason to breast cancer, but the first method of analysis was more precise and accurate in identifying these factors. Concerning both methods the focus was on the factors of: breast self examination, the periodical mammography machine examination early and continuously (periodically), and their life style had a major effect on their susceptibility of the risk of breast cancer.

Keywords: Factorial Analysis, Radial Bases Function Network, Gaussian Function, Breast cancer, Eigen values

Corresponding Author:
Aseel Mubder Dagwood Al-Marsuomy
Department of Computer Systems, Middle Technical University/Institute of Administration-Rusafa
Baghdad - Iraq
almarsoumy66@yahoo.com

الكشف المبكر عن أهم العوامل المؤدية لسرطان الثدي

أسيل مبدر داود[1] **مصطفى محمد صبحي[2]**
[1] الجامعة التقنية الوسطى ـ معهد الإدارة ـ الرصافة **[2]جامعة أوروك ـ كلية الإدارة والاقتصاد**

المستخلص: الهدف من هذا البحث هو تحديد أهم وأهم العوامل التي تؤدي إلى الإصابة بسرطان الثدي. كما يتعلق بإيجاد الحلول المناسبة للقضاء على هذه العوامل وتجنبها للوقاية من الإصابة بسرطان الثدي.ولتحقيق هدف هذا البحث تم اختيار عينة من (40) مريضا من المدينة الطبية - مستشفى الأورام التعليمي حيث تم استخدام استبانة لجمع البيانات. احتوت على (17) سببا. تم استخدام البرنامج الإحصائي (SPSS) لمقارنة نتائج تحليل البيانات باستخدام طريقتين (شبكة وظائف القواعد الشعاعية) و (التحليل العاملي). توصل البحث إلى نتائج مهمة؛ حددت الطريقتان نفس العوامل التي يمكن أن تكون سببًا مباشرًا لسرطان الثدي، لكن الطريقة الأولى في التحليل كانت أكثر دقة ودقة في تحديد هذه العوامل. فيما يتعلق بكلتا الطريقتين، كان التركيز على عوامل: الفحص الذاتي للثدي ، والفحص الدوري بآلة التصوير الشعاعي للثدي مبكرًا ومستمرًا (دوريًا)، وأسلوب حياتهم كان لهما تأثير كبير على قابليتهما لخطر الإصابة بسرطان الثدي.

الكلمات المفتاحية: التحليل العاملي، شبكة وظائف القواعد الشعاعية، الوظيفة

Introduction:

Breast cancer is a type of cancer that develops from breast tissue. It usually and most commonly develops in cells from the lining of milk ducts and the lobules that supply the ducts with milk. Men and woman can get breast cancer, but men diagnosed with breast cancer are rare. Breast cancer forms (22.9%) of common cancer cases. It leads to an anomaly and increase in tissue cells of the breast and developing of lumps that are called tumors. These tumors can either be malignant or benign. Malignant tumors destroy the healthy body tissues.

The statistical methods that were used in data analysis are: (Radial Bases Function Network) and (Factorial Analysis). These two methods were chosen specifically because of linear relationship between the variables acquired. The number of variables was large which required the use of the two methods to reduce the large number of variables in order to compare both of them which would help determine the common factors in the two methods that would have a major effect in developing the disease.

1. The Preliminary Part

1.1 The Aim of the Research:

Shedding some light on the importance of determining the main causesfor breast cancer in women by using (Radial Bases Function Network) and (Factorial Analysis) where through these methods reducing the factors and focusing on the important and main factors in developing the disease and then comparing the results from the two methods to get to the direct and real factors in developing the disease.

1.2 The Importance of the research:

The importance of the research lies in using (Radial Bases Function Network) and (Factorial Analysis) as two statistical tools working on reducing the effective factors and to get through them to the factors that cause the disease.

1.3 The Research Hypothesis:

The hypothesis can be formed as follows:

H_0: the data that were acquired as explanatory data for the factors leading to developing breast cancer. It indicates that there are no effective factors that lead to developing breast cancer.

H_1: the data that were acquired as explanatory data for the factors leading to developing breast cancer. It indicates that there are effective and major factors that lead to developing breast cancer that must be avoided and treated in order to prevent the disease.

1.4 Research Sample:

The data used in this research were acquired from Medical City (Oncology Teaching Hospital) with a sample size of (40) patients and (17) causes(X_i) with different effects that leads to developing the disease and they are as follows:

X1: Age

X2: Social Status

X3: Academic Achievement

X4: Where you live?

X5: Was one of the ovaries or both of them surgically removed?

X6: Were you taking contraception pills?

X7: Source of nutrients is from animals or vegetables

X8: Do you go for a periodic mammography examination?
X9: Do you used a hormonal treatment especially pills that contain progestin?
X10: Do breast feed?
X11: Were you ever exposed to x-rays and chemo therapy?
X12: Do you have a family member who has the disease?
X13: At what age did you start periodical examination?
X14: Do you have children? And how many?
X15: Are you keen on doing breast self examination
X16: At what age did you start breast self-examination?
X17: Are you suffering from being overweight?

1.5 Research Problem:

Studying the effective factors in developing breast cancer and according the reports from the Iraqi Ministry of Health indicates the increase of developing the disease in Iraq and it is obligatory to determine the direct factors of the disease.

1.6 Research Structure:

The research was divided into four main chapters where in chapter one deals with the preliminary frame of the research structure and the concept of developing breast cancer and types of diagnosis. As for chapter two, it deals with the theoretical part of the main concept of the two methods that were used in data analysis (Radial Bases Function Network), and (Factorial Analysis). Wherein chapter three deals with the applicable part using the package (SPSS). Chapter four has the conclusions and recommendations.

1.7 Breast Cancer (Causes and Prevention and Treatment):

Breast cancer means that the number of breast cells start to multiply in an abnormal way. These cells start to divide faster than the healthy cells and start to spread (Metastasis) in all the breast tissue to the lymph nodes and on to other organs in the body.

1.7.1 Causes of the disease:

There are a number of causes that leads todeveloping the disease, some of them are:

1) Genetic: (5 – 10) % of breast cancer cases are results of genetic causes. There are families with a defective gene one or two breast cancer gene number (1) (BRACA.1) or breast cancer gene number (2) (BRACA.2) and that could make the chances of her sons or daughters developing breast cancer or ovarian cancer are very high.
2) Exposure to radiation.
3) Overweight.
4) Menstruation at an early age.
5) Reaching menopause at a relatively late age.
6) Hormonal therapy.
7) Taking contraception.
8) Smoking.

1.7.2 Diagnosis:

There are a number of methods in which the disease can be diagnosed like**:**

1) Self Examination: where these examinations help determine any irregularities in breast tissue that can be discovered through routine examination like: discovering a lump in the breast by touch or by Magnetic Resonance Imaging (MRI).

2) Mammography: Mammography examines the breast tissue through generating x-ray images where nowadays is considered one of the most trusted examinations for early discovery of breast cancerous lumps even before the doctor could feel it by touch and that is why this type of examination is recommended on periodic basis to all women above the age of forty.
3) Taking a node biopsy sentinel lymph since breast cancer spreads towards lymph nodes in the under arms (sentinel lymph node) so all women who were diagnosed with cancer of the invasive type, they should get these glands examined.
4) Examination with Computer Aided Design (CAD).
5) Breast ultrasound.
6) Magnetic Resonance Imaging (MRI).

1.7.3 Prevention from breast cancer:

There are no ways in which we can guarantee not developing breast cancer, but there are numerous steps that we can take to limit the risk of developing breast cancer.

1) Prevention through chemo therapy.
2) Surgical removal of breasts as a precaution.

1.7.4 Treatment:-

There are a number of methods in which the disease can be treated like**:**

1) Surgery: the removal of the whole breast has become a rare procedure these days and instead most women are excellent candidates for partial removal of the breast(the cancerous part) or surgical removal of the tumor only.
2) Radiation Therapy.
3) Chemotherapy.
4) Hormonal therapy.
5) Biological therapy.

2. The Theoretical Part

In this chapter we are going to deal with a number of important and necessary definitions to understand the theoretical part of the research.

2.1 Important definitions of Radial Bases Function Network:

2.1.1 Radial Bases Function Network (RBFN):

It consists of two layers excluding the input layer and they are the hidden layer and the output layer. The reason for calling it Bases Function because the cells of the middle layer represent a group of functions like (Gaussian Function) and scatter plot of the data resembles a ray and that is why it is called as such (Rentals, 1989) (Feral, 1997). The (RBFN) is considered a hybrid network of a group of characteristics in some types of artificial neural networks like (Single layer Perceptron) and (Error back propagation) where the network acquired the property of error back propagation from the aforementioned networks.

2.1.2 Constructing Radial Bases Function Network:-

It consists of three layers of cells and they are: input layer, hidden layer, and output layer and each layer in this network is connected with the layer that follows which means any cell from the input layer connects with all cells in the hidden layer and all cells from hidden layer sends its output to every cell in the output layer. The number of cells in the input layer depends on the application given to the network and the number of hidden cells depends on the degree of complication in the problem. The

following figure explains the Radial Bases Function Network which consists of three layers (Kiernan, 1996).

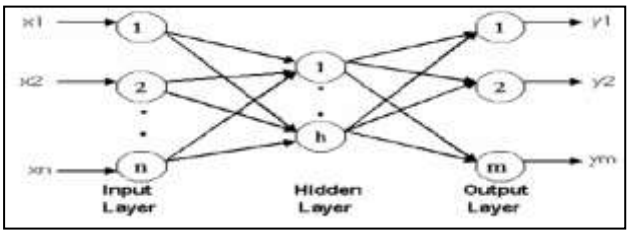

Figure (1) (Construction of Radial Bases Function Network)

2.1.3 Weights in(RBFN):

Primary weights are prepared for first processing (hidden layer) and so for the output layer with random numbers between [0-1] as for changing weights for the layers it would be as follows (Leonard 1991)

1) Hidden layer: No changing weights in this layer.
2) Output layer: changing weights when we don't get the required output and that is done through mathematical equations explained in the algorithm of the Radial Bases Function.

2.1.4 Gaussian Function:

In the hidden layer, the Gaussian function is used. The mathematical structure is used (Buck, 2002)**:**

$$O_i = exp\left[-(X - W)^2/2\sigma^2\right]$$

Wherein:

O_i**:** represent the actual output of the cell (i) and the output value would be between [0-1].

X: is the input vector of the network.

W: is the escort vector weights for the inputs and they are randomly generated values.

2.2 Important definitions of Factorial Analysis:

2.2.1 Factorial Analysis:

It is a group of statistical methods that aim to reduce the number of variables or data related to a certain phenomenon. It is used in data analysis or correlation matrices or variance matrices of variables and the results of their multiplications. The aim would be to explain the relationships between the variables and it results in a number of numerous changes or hypothetical and they are called factors. It aims to analyse a group of correlation coefficients between the number of variables and it reduces them to a lesser number of factors which would help in comprehending the formation of the correlation matrix or the covariancematrix through lesser number of factors.

2.2.2 Eigen values:

It is a sum of squares of every variable on each factor from the matrix factors on its own and it indicates the amount of variance which the factor contributes.

2.2.3 Commonality:

The sum of contributions of the variable in different factors which can be concluded from the factorial matrix and wherein each variable contributes in different amounts in each factor and whether its contributions are fundamental or non-fundamental and so the sum of squares of these contributions of matrix factors is the value of commonality for the variable or contributions.

2.2.4 Rotation:

After the factors reaches its saturation the process of rotating the factors to another place where it can be explained. The main aim of rotating factors is to reach a suitable combination for factors that can be explained and so rotating factors helps in explaining factors a logical explanation.

2.2.5 KMO Test:

We use the test of (Kaiser, Meyer, Olkin) to calculate the sufficiency of the sample and to test whether or not the partial correlations between the variables are small and ranges between (0 to +1) wherein the values close to (+1) indicates to the sufficiency of the sample or that it is adequate.

2.3 Types of factorial analysis:

First: exploratory factorial analysis:

It is used in discovering the factors that the variables can be classified into considering the factors are categories of these variables.

Second: Confirmatory factorial analysis

2.4 The conditions of factorial analysis:

1) The identical distribution of variables.
2) The level of measuring the variables from the categorical level or relative.
3) The existence of linear relationships between the variables.
4) The chosen sample must be random, large, and represents the population (for every variable at least 10 individuals).
5) The independence of errors in every variable and the independence of variables themselves.

3. The Applicable Part

The applicable part includes the analysis of the sample with available statistical methods. As for the research under study two methods were used through program (SPSS) and they are: (Radial Bases Function Network) and (Factorial Analysis) and that is for the purpose of comparing the results of the two analyses to obtain the least error in the estimation.

3.1 The results of data analysis using neural networks.

Because of the scientific advancement in all fields especially in the medical field, it has become necessary to build and develop methods of analysis that specializes in this field. As it has been explained, the research focuses on the most important factors that lead to developing breast cancer using artificial neural network which is a RBFN. It is a function that was counted on during the analysis in this programme (SPSS) and that is because this method of analysis is compatible with the nature of the problem under study.

Table No. 1

Case Processing Summary			
		N	Percent
	Training	32	80.0%
	Sample Testing	8	20.0%
Valid		40	100.0%
Excluded		0	
Total		40	

Table No. 2

Network Information			
Input Layer	Covariates	1	X1
		2	X2
		3	X3
		4	X4
		5	X5
		6	X6
		7	X7
		8	X8
		9	X9
		10	X10
		11	X11
		12	X12
		13	X13
		14	X14
		15	X15
		16	X16
		17	X17
	Number of Units		17
	Rescaling Method for Covariates		Standardized
Hidden Layer	Number of Units		5[a]
	Activation Function		Softmax
Output Layer	Dependent Variables	1	Y
	Number of Units		2
	Activation Function		Identity
	Error Function		Sum of Squares

Where (a) is: Determined by the testing data criterion the "best" number of hidden units is the one that yields the smallest error in the testing data.

Table No.3

Model Summary		
Training	Sum of Squares Error	2.392
	Percent Incorrect Predictions	12.5%
	Training Time	0:00:00.06
Testing	Sum of Squares Error	1.397[a]
	Percent Incorrect Predictions	37.5%

Where (Y) is : Dependent Variable and (a) is: The number of hidden units is determined by the testing data criterion: The "best" number of hidden units is the one that yields the smallest error in the testing data.

Table No.4
(Parameters Estimates)

Predictor		Predicted						
		Hidden Layer[a]					Output Layer	
		H(1)	H(2)	H(3)	H(4)	H(5)	[Y=0]	[Y=1]
Input Layer	X1	-2.009-	.372	.372	.372	.372		
	X2	-.559-	-.559-	.254	.254	.254		
	X3	.536	-.473-	-.113-	-.473-	.215		
	X4	.317	.317	-.166-	-1.372-	.317		
	X5	.317	.317	-.648-	.317	.010		
	X6	.117	-.298-	-.120-	1.360	-.336-		
	X7	.013	-.789-	-.904-	.313	.814		
	X8	.013	.814	-.331-	.313	-.279-		
	X9	-.114-	-.568-	.081	1.137	-.155-		
	X10	-.719-	-.322-	.585	.868	-.214-		
	X11	-.473-	-.473-	.608	.158	-.014-		
	X12	-.712-	.117	.176	.842	-.147-		
	X13	.712	.712	-1.360-	.194	.147		
	X14	.424	-.661-	.424	-.254-	-.069-		
	X15	-1.034-	.317	-.166-	.317	.317		
	X16	.762	.356	-1.271-	.762	.023		
	X17	.818	-.372-	-.372-	1.116	-.372-		
Hidden Unit Width		1.718	1.289	1.516	1.575	1.340		
Hidden Layer	H(1)						.203	.797
	H(2)						1.570	-.570-
	H(3)						1.052	-.052-
	H(4)						-.087-	1.087
	H(5)						-.402-	1.402

Where (a) is displays the center vector for each hidden unit.

Table No. 5 (Classification)				
Sample	Observed	Predicted		
		Not Infected	Infected	Percent Correct
Training	Not Infected	10	3	76.9%
	Infected	1	18	94.7%
	Overall Percent	34.4%	65.6%	87.5%
Testing	Not Infected	2	3	40.0%
	Infected	0	3	100.0%
	Overall Percent	25.0%	75.0%	62.5%

Where: (Y) is dependent Variable.

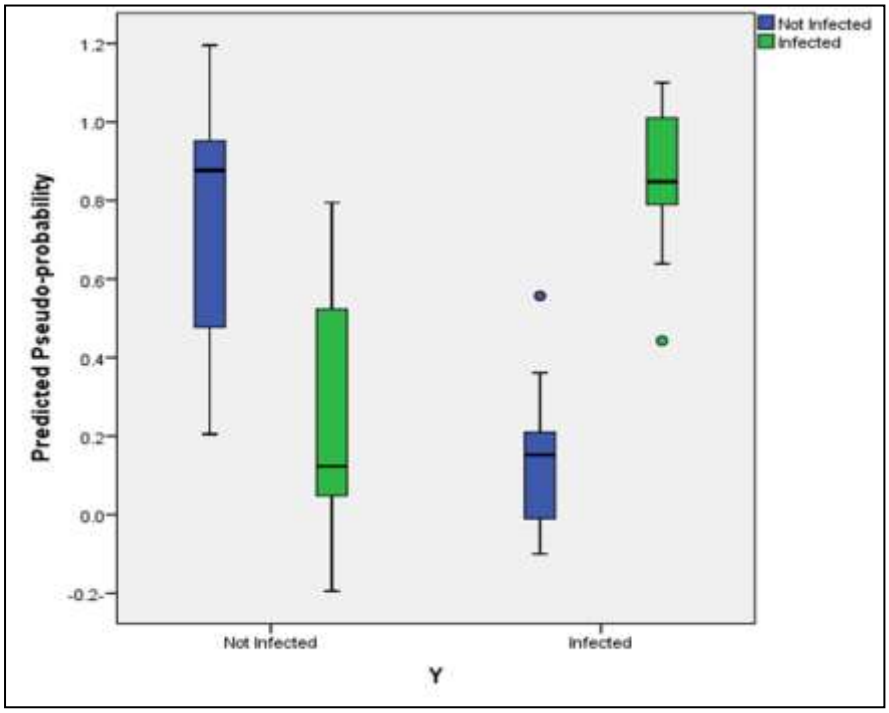

Figure (2)

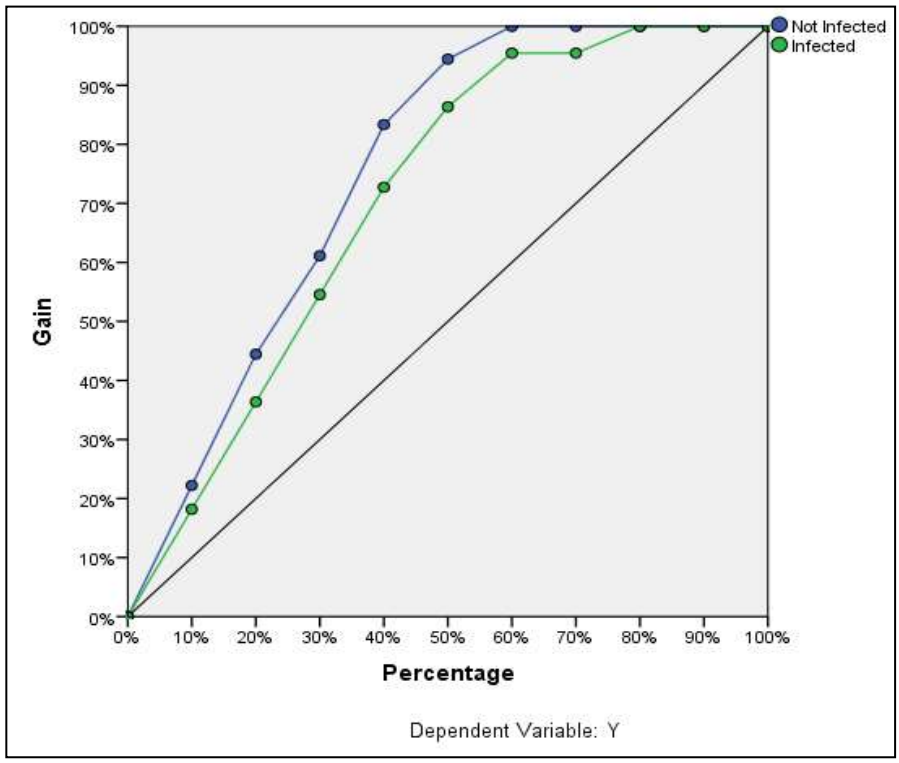

Figure (3)

Table (6) (Independent Variable Importance)

	Importance	Normalized Importance
X1	.083	57.2%
X2	.060	41.1%
X3	.054	36.8%
X4	.071	48.5%
X5	.071	48.7%
X6	.031	21.1%
X7	.146	100.0%
X8	.052	35.4%
X9	.046	31.5%
X10	.024	16.6%
X11	.042	28.9%
X12	.021	14.2%
X13	.065	44.8%
X14	.040	27.4%
X15	.063	43.4%
X16	.058	40.1%
X17	.075	51.2%

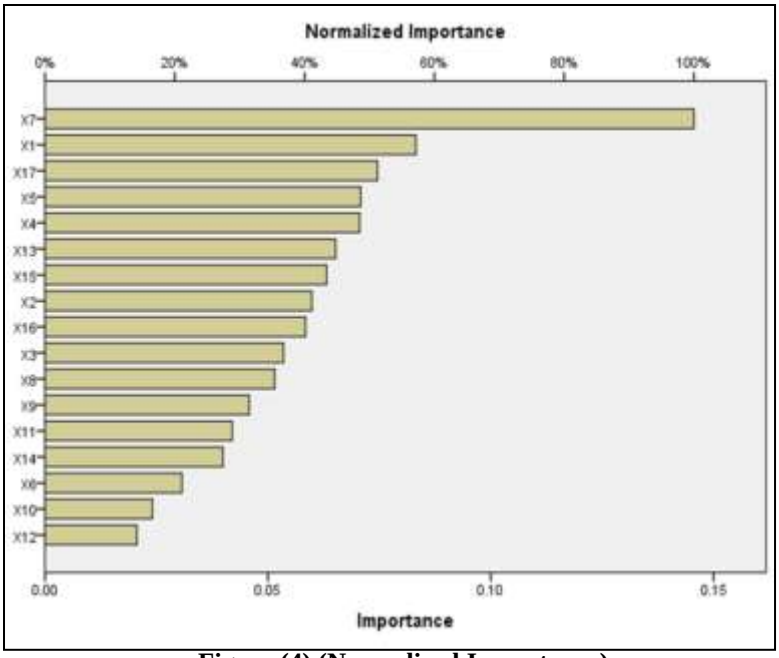

Figure (4) (Normalized Importance)

3.2 Theresults of data analysis the factorial analysis method:-

The applicable part includes the data analysis of the sample using the factorial analysis method through the statistical program (SPSS), before conducting the analysis we testedthe sample to know the extent of its sufficiency using the KMO test as it is shown in table No. (7)

Table No. (7) Shows the sufficiency of the size of the sample for data of breast cancer in women.

KMO and Bartlett's Test		
665Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.503
Bartlett's Test of Sphericity	Approx. Chi-Square	278.824
	Df	153
	Sig.	.000

From table No. (7): We notice that the measurement value of the (KMO) test is (0.503) and as long as the value is larger than (0.5) so the size of the sample is considered adequate for the factorial analysis. As for the (Bartlett's) test the significance level was (0.000) which is less than the accredited significance level (0.5) which indicates that the matrix represents the identity matrix.

Table No. (8) Shows the contribution of data in effecting the development of breast cancer in women

Communalities		
	Initial	Extraction
Age	1.000	0.727
Social status	1.000	0.865
Academic achievement	1.000	0.753
Where do you live	1.000	0.674
Was one of the ovaries or both surgically removed	1.000	0.639
Have you ever used contraception	1.000	0.630

The source of nutrients is animal or vegetable	1.000	0.889
Do you have a periodic mammography examination	1.000	0.791
Do you use hormonal therapy especially the one that contains progestin	1.000	0.705
Do you breast feed	1.000	0.842
Have you ever been exposed to x-ray and chemotherapy	1.000	0.762
Do you have members of family with the disease	1.000	0.748
At what age did you start periodic examination	1.000	0.805
Do you have children and how many	1.000	0.822
Are you keen on doing the periodic self-examination	1.000	0.808
At what age did you start the periodic self-examination	1.000	0.815
Are suffering from overweight	1.000	0.640
How many children	1.000	0.799
Extraction Method: Principal Component Analysis.		

The common abstract value for the first variable indicates to (0.727) of the data in the values of the first variable (age) explained by the common factors which is (0.630) (did you use contraception pills).

Table No. (9) Explains the variance of effective factors

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative%
1	3.803	21.128	21.128	3.803	21.128	21.128	2.454	13.631	13.631
2	2.754	15.301	36.429	2.754	15.301	36.429	2.418	13.435	27.066
3	2.074	11.523	47.952	2.074	11.523	47.952	2.180	12.112	39.178
4	1.777	9.871	57.822	1.777	9.871	57.822	2.020	11.225	50.403
5	1.222	6.788	64.610	1.222	6.788	64.610	1.892	10.514	60.917
6	1.086	6.031	70.641	1.086	6.031	70.641	1.508	8.378	69.295
7	1.000	5.557	76.198	1.000	5.557	76.198	1.243	6.903	76.198
8	.865	4.808	81.007						
9	.628	3.488	84.494						
10	.570	3.166	87.661						
11	.481	2.670	90.330						
12	.436	2.421	92.751						
13	.398	2.209	94.960						
14	.293	1.630	96.590						
15	.212	1.179	97.769						
16	.175	0.974	98.743						
17	.143	0.792	99.535						
18	.084	0.465	100.000						
Extraction Method: Principal Component Analysis.									

We notice that: there are 18 linear relationships with the data. The value of the Eigen values for the first factor reached (3.803) where it explains the variations of this factor (21.128%) from total variance as for the value of Eigen values for the second factor reached (2.754) and explains (15.301%) as for the third factor it was (2.074) and explains (11.523%) of the total variance. As for the fourth factor it was (1.777) and explains (9.871%) of the total variance. As for the fifth factor it was (1.222) and explains (6.788%) of the total variance. As for the sixth factor it was (1.086) and explains (6.031%) of the total variance. As for the seventh factor its value was (1.000) and the variations of this factor explains (5.557%) of total variance. Despite that these factors include different proportions of the variance, at the end of the table it is mentioned that the sum of variance (total variance) of the factors is (76.198%).

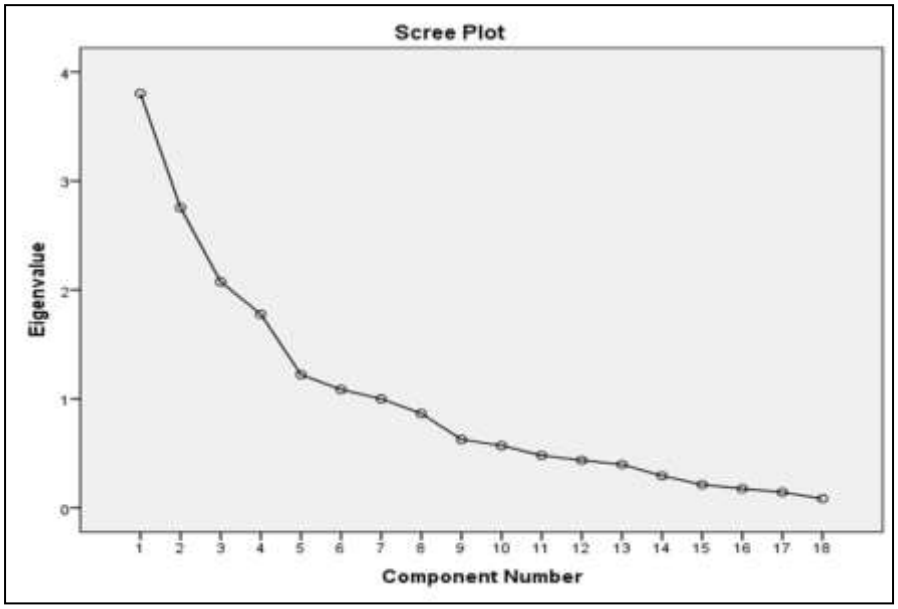

Figure (5) (the variance of effective factors)

It is another criterion to determine the factors in the factorial analysis and to only keep the ones in the extreme slope (criterion for keeping factor that has an Eigen value more than a whole one)

Table No. (10) (Component Matrix)

	Component						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Age	.601	.218	-.358-	-.167	-.274		.287
Social status	.456	.122	-.310	.485	.361	.375	-.199-
Academic achievement	-.236-			.295	.759	-.170-	
Where do you live	.276	.679			.109	.192	-.290-
Was one of the ovaries or both surgically removed	-.528-	-.178-	-.311-	.324	-.249	.252	
Have you ever used contraception	.105	.638	.374	-.262-			
The source of nutrients is animal or vegetable	-.281-		.247	.236	.168	.479	.656
Do you have a periodic mammography examination	-.414-		.207	-.461-		.408	_.445 -
Do you use hormonal therapy especially the one that contains progestin	.655	.201	.360	.257	-.191		
Do you breast feed	-.344-	.341	.341	.520	-.391		-.247
Have you ever beenexposed to x-ray and chemotherapy	.484	-.295-	.643				.141
Do you have members of family with the disease		.368	.686	-.173-	.189	.254	
At what age did you start periodic examination		.605	-.262	-.545-	.135	-.204-	
Do you have children and how many	-.527-	.586		.331	-.186	-.239-	
Are you keen on doing the periodic self-examination	-.800-	-.104-		-.281	-.142	.205	.129
At what age did you start the periodic self-examination	-.267-	.708	-.439			.141	.169
Are suffering from overweight	.657	.243	-.192	.276	-.176		
How many children	.589	-.359-	-.278	-.232-		.418	-.113-
Extraction Method: Principal Component Analysis.							
a. 7 components extracted.							

The matrix shows the factors before rotation and shows every factor containing some variable. The seven factors are:

The first factor: it contains four variables (age, do you use hormonal therapy especially the ones containing progestin, are you suffering from overweight, how many children)

The second factor: it contains five variables (where do you live, have you used contraception, at what age did you start periodic self examination, do you have children and how many, at what age did you start self examination)

The third factor: it contains two variables (have you been exposed to x-rays and chemotherapy, do have members of the family with the disease)

The fourth factor: it contains three variables (social status was one of the ovaries or both of them surgically removed, have you ever been exposed to x-rays and chemotherapy).

The fifth factor: it contains one variable (academic achievement)

The sixth factor: it contains two variables (do you have periodically go for mammography examination, are you keen on doing self examination)

The seventh factor: it contains one variable (the source of nutrients is it from animals or vegetables)

Table No. (11) (Rotated Component Matrix)

	Component						
	1	2	3	4	5	6	7
Age	.253	.108	-.315	-.447	.181	.565	
Social Status				-.210	.880	-.149	
Academic Achievement					.129	-.851	
Where do you live	.381	.505	.117		.478		-.164
was one of the ovaries or both surgically removed?	.129	-.624	.338	.264			.212
Have you ever used contraception	.201	.732	.183			.134	
The source of nutrient is it animal or vegetable							.932
Do you have a periodic mammography examination		.157		.870			
Do you use hormonal therapy especially the one with progestin in it	-.369	.421	.150	-.412	.351	.260	
Do you breast feed			.891	.122	.108		
Have you even been exposed to x-ray and chemotherapy	-.734	.362	-.143	-.238			
Do you have a member of family with the disease	-.160	.769		.184			.290
At what age did you start with periodic examination	.693	.439	-.229		-.162		-.211
Do you have children and how many	.433		.761		-.172	-.144	
Are you keen on the periodic self examination	.237	-.251	.131	.595	-.478		.300
At what age did you start periodic examination	.843		.219		.114		.191
Are you suffering from overweight				-.552	.457	.332	
How many children	-.182	-.130	-.601		.434	.432	

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 11 iterations.

The table contains the same data as in table no. (10), but summarizes the diagnosis for each factor acquired after rotation.

The first factor: it contains two variables (at what age did you start periodic examination, at what age did you start self examination)

The second factor: it contains five variables (where do you live, have you ever used contraception, do you use a hormonal therapy especially ones containing progestin, have you have ever been exposed to x-rays and chemo therapy, do have a family member with the disease)

The third factor: it contains three variables (was one or both of the ovaries surgically removed, do you breast feed, do have children and how many)

The fourth factor: it contains two variables (do you go for periodic mammography examination, are you keen on self examination)

The fifth factor: it contains four variables (social status, academic achievement, are you suffering from overweight, how many children)

The sixth factor: it contains one variable (age)

The seventh factor: it contains one variable (the source of nutrients is it from animals or vegetables).

4. Conclusions and Recommendation

4.1 Conclusions:

1) When using two methods of data analysis that allows the researcher to freely compare between the results of the analysis to reach exact results, the researcher can determine the factors that have a direct impact of the disease in order to avoid it and avoid infection to the disease.
2) The results of the radial base function network analysis and factorial analysis confirm both of them that the feeding method (animal or vegetable) has a big effect On the ready of physical to infection breast cancer, this means that the culture of society is of great importance to avoid the disease
3) The self-checking and periodic checking by mammography system early and continuously (periodically) both of them have a great effect in reducing the infected breast cancer and this depends on the extent of health awareness in women.
4) We note through radial analysis that married women are more likely to infection of breast cancer.
5) We note from the results of the analysis that the analysis of the radial base function network is the best and most explaining of the explanatory variables and their importance from through infection the disease.

4.2 Recommendations:

1) Throughout research and studies of the problem (breast cancer), there should be awareness about nutrition since childhood must depend on vegetable source rather than animal one.
2) Every girl and woman must do the breast self-check especially the married ones and the periodic mammography examination.
3) The work of awareness seminars for women periodically starting from the age of 14 years for the purpose of increasing awareness among women.
4) Building more specialized centers for breast cancer treatment to include the largest number of women.

Reference

اولاً : المصادر العربية:

1) البياتي، محمود تحليل البيانات بإستخدام برنامج إحصائي SPSS، دار ومكتبة الحامد ، عمان 2004م .
2) البلداوي .د. عبد الحميد عبد المجيد (اساليب الاحصاء) دار وائل للنشر / الطبعة الاولى .

ثانياً : مصادر الأنترنت:

1) https://www.webteb.com
2) https://www.ibtesamah.com

Third: English reference:

[1] Buck, Sbastian, 2002,"A C++ Library for Function Approximation using Networks of Radial Basis Functions", Munich University of Technology. http://www.in.tum.de/people/buck/index-e.html.
[2] Feral, Sahin, 1997, "A Radial Basis function Approach To a color Image Classification Problem in a Real Time Industrial Application", formerly the scholarly communications projects, Electrical engineering.
[3] Kiernan, L., al., 1996, "Robust Initialization of Gaussian Radial Basis Function Networks Using Partitioned K-means is clustering," Electronic Letters, vol. 32, no.7, pp. 671-3.
[4] Koenig, A., ET. 1994, "A Novel Method for the Design of Radial Basis Function Networks and Its Implication for Knowledge Extraction," IEEE International Conference on Neural Networks. IEEE World Congress on Computational Intelligence, p.7 vol. 1804-9 vol. 3.
[5] Leonard, J. A. and M. A. Kramer, 1991, "Radial Basis Function Networks for Classifying Process Faults," IEEE Control Systems Magazine, vol. 11, no.3, pp.31-8.
[6] Martin D. Buhrmann, M. J. Ablowitz (2003). "Radial Basis Functions: Theory and Implementations". Cambridge University.
[7] Renal, S., 1989, "Radial Basis Function Network for Speech Pattern Classification," Electronics Letters, vol. 25, pp. 437-439.
[8] Richard E. Behrman and Rebort M. Kliegman, 1990, M. D., "Nelson Essentials of Pediatrics", by W. B. Saunders Company.
[9] Sang, On ghee, 1999, "Recent Advances in Radial Basis Functions Network ", International Journal of Neural System , http://www.anc.ed.ac.uk/~mjo/rbf.html.
[10] Shin, C., et al., 1996, "Radial Basis Function Network Design for Chaotic Time Series Prediction," Transactions of the Korean Institute of Electrical Engineers, no. 4, pp. 602-11.
[11] Yee, Paul V. and Hay kin, Simon (2001). "Regularized Radial Basis Function Networks: Theory and Applications". John Wiley. http://en.wikipedia.org/w/index.php

Appendices

Questionnaire

The questionnaire in your hands is a part of the requirement for the research entitled for the health statistical study on (The Early diagnosis of Breast Cancer and the Most Important Factors that Leads to Developing (Breast Cancer)

Please place (√) in the suitable place.

1. Age
1 More than 40 years old 0 Less than 40 years old

2. Social Status
1 Single 0. Married.

3. Academic Achievement
1 College 0. Less than college.

4. Where do you live?
1 Country 0. City

5. Was one of the ovaries or both of them surgically removed?
1 Yes. 0. No.

6. Have you ever used contraception?
1 Yes, 0. No.

7. What is the source of nutrients that you depend on?
1 Animal 0 Vegetable

8. Do you have a periodic Mammography examination?
1 Yes. 0 No.

9. Do you use hormonal treatment especially pills that contain Progestin?
1 Yes, 0. No.

10. Do you breast?
1 Yes. 0. No.

11. Were you ever exposed to x-rays and chemotherapy?
1 Yes, 0. No

12. At what age did you start a periodic examination?
1 Early 0. Late.

13. Do you have children?
1 Yes. 0. No.

14. Are you keen on self examination?
1 Yes. 0. No.

15. Are you suffering from overweight?
1 Yes. 0. No.