International Journal of Biomedical Data Mining
ISSN: 2090-4924
Hamouda et al., Int J Biomed Data Min 2017, 6:2
DOI: 10.4172/2090-4924.1000128
Research Article
OMICS International

Intelligent System for Predicting, Diagnosis and Treatment of Breast Cancer

Hamouda SKM[1*], Abo El-Ezz HR[2] and Wahed ME[3]

[1]Zagazig University, Zagazig, Egypt
[2]Al-Azhar University, Cairo, Egypt
[3]Suez Canal University, Ismailia, Egypt

Abstract

Cancer is a disease that makes the cells in the body grow out of control. When cancer starts in the breast, it is called breast cancer. Breast cancer is one of the major death causing diseases of the women in the world due to delays and inaccuracies in diagnosis of the disease. The high accuracy in cancer prediction is important to improve the treatment quality and the survivability rate of patients.

In this paper, we hope to reduce the risk of the breast cancer at the earlier stage. So, we have propose a contain two parts: First: We use a Rough Set Theory (RST) as an efficient and intelligent technique, to analyze breast cancer dataset, evaluate approximate sets, and improve the accuracy of diagnosis with reduce redundancies, and evaluate the importance attributes of data. Second: We will construct MATLAB program to diagnosis and treatment of breast cancer depending on the decision rules which can generate by Rough Set Theory from dataset, and the information's taken from records of patient's data base. This system may help doctors to predict and diagnose breast cancer early, also helps in follow-up of the record and medical history of any patient.

Keywords: Breast cancer predictive; Diagnosis; Survivability; Rough set theory; MATLAB; Treatment

Introduction

Breast cancer [1,2] is characterized by the uncontrolled growth of abnormal cells of the breast. Most of the breast cancers start from the ducts that carry milk to the nipple called ductal cancers. Other type of cancer start in the glands that make breast milk called lobular cancers. There are also other types of breast cancer that are less common. There is no sure way to prevent from the breast cancer, but the women can reduce their breast cancer risk, breast cancer also develops in men. Earlier detection of cancer is curable and may increase the survivability rate of patients. Women must be checking breasts for cancer before she has any symptoms, there are various tests available for prediction and diagnosis of breast cancer [1-4].

Medical databases contain large amounts of data about the patients with their medical conditions. Analyzing all these databases is one of the difficult tasks in the medical environment. Thus, we need to extract the knowledge about the patients from these databases.

The rough set theory (RST) is a mathematical tool for extracting the knowledge from uncertain and incomplete database information [5,6]. Rough set methods utilize the comparison between elements, e.g., discernibility, indiscernibility, and similarity. Additionally, we construct program by MATLAB [7-10] to diagnosis and treatment of breast cancer based on the decision rules which are generated by Rough Set Theory from patient's data base.

Breast Cancer Overview

Breast cancer

The anatomy of normal breast [1-3] is shown in Figure 1. This picture shows the lobes and ducts inside the breast, it also shows lymph nodes near the breast.

Breast cancer types

Ductal carcinoma: The first and most common type of breast cancer is ductal carcinoma (Figure 2). In this type, the cancer begins in cells that line a breast duct; about 7 of every 10 women with breast cancer have ductal carcinoma.

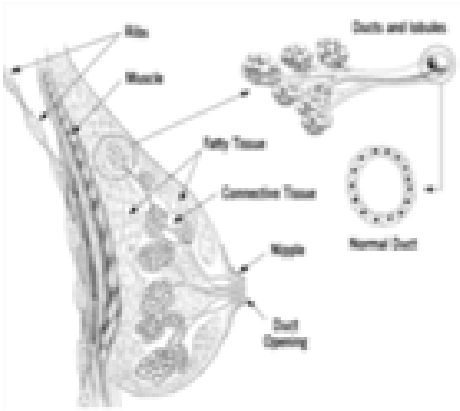

Figure 1: Normal breast anatomy.

Lobular carcinoma: Another type of breast cancer is lobular carcinoma. This cancer begins in a lobule of the breast. About 1 of every 10 women with breast cancer has lobular carcinoma.

Ductal and lobular carcinoma: Other women have a mixture

***Corresponding author:** Hamouda SKM, Computer Laboratories, Zagazig University, Zagazig, Egypt, Tel: +201124776320; E-mail: HamoudaSaeed@yahoo.com

Received March 10, 2017; **Accepted** March 20, 2017; **Published** March 27, 2017

Citation: Hamouda SKM, Abo El-Ezz HR, Wahed ME (2017) Intelligent System for Predicting, Diagnosis and Treatment of Breast Cancer. Int J Biomed Data Min 6: 128. doi: 10.4172/2090-4924.1000128

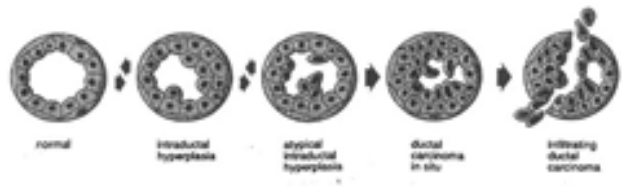

Figure 2: Ductal carcinoma.

of ductal and lobular type or they have a less common type of breast cancer.

Inflammatory breast cancer: The inflammatory cancer is other type of breast cancer. It occurs in about 1 of every 100 women with invasive breast cancer [1-4].

Breast cancer stages

The stage of breast cancer depends on the size of the tumour and determines whether it has spread and how far it has spread. Treated describe the stages of breast cancer: stage 0, stage I, stage II, stage III, and stage IV. Stage I is early-stage breast cancer, and Stage IV is advanced-stage breast cancer that has spread to other parts of the body. The stage of a cancer is one of the most important factors in selecting treatment options, and it uses the tumors, nodes and metastasis (TNM) system. When a patient's TNM categories have been determined, then this information is combined in a process known as stage grouping to determine a woman's disease stage [1,2,11,12]. In stage 1, tumor size less than 2 cm and the cancer hasn't spread beyond the breast. Stage II: tumor size less than or equal 5 cm and cancer may have spread to the lymph nodes. Stage III: tumor size greater than 5 cm and tumors have spread to the lymph nodes, and possibly the chest wall. Stage IV: means cancer has reached other, remote parts of the body.

Effect of breast cancer stages on survival

The effect of breast cancer stages on the survival [13,14] vary depending on the stages of breast cancer.

Non-invasive and the early stages have a better chances of survival than that for the metastatic breast cancer (stage IV) which is the stage wherein the cancer has spread beyond the neighboring tissues. Table 1 shows the 5-year and 10-year survivability rate of a cancer patient.

Breast cancer risk factors

Women who have one or more breast cancer risk factors never develop breast cancer, while many women with breast cancer have no known risk factors. Even when a woman with risk factors develops breast cancer, it's hard to know just how much these factors might have contributed. Some risk factors can't be changed, like a person's age or race, other risk factors can change, such as smoking, drinking, and diet [1-3,15]. Some factors affect risk more than others, and the risk for breast cancer can change over time.

Risk factors you cannot change

Getting older: The risk of developing breast cancer increases with your age. Approximately, 75% of all breast cancers are diagnosed in women 50 years and over.

Genetic risk factors: About 5-10% of breast cancers are inherited, most commonly the genes *BRCA1* and *BRCA2*.

Family history: If a woman has a personal or family history of breast cancer, she is at increased risk of developing breast cancer.

Stage	Description	5–year survival	10–year survival
Stage 0	T=0, LN=0, M=0	95%	90%
Stage I	T≤2 cm ,LN=1:3, M=0	85%	70%
Stage II	T>2 and ≤ 5 cm,LN=3:9, M=0	70%	50%
Stage III	T>5 cm, LN>9, M=0	55%	30%
Stage IV Metastasis	T=any, LN=any, M=1	5%	2%

Table 1: Breast cancer survivability rate.

HER2/neu Status: Cancers that are HER2-positive have too many copies of the *HER2/neu* gene, these cancers tend to grow and spread more aggressively than other breast cancers. All newly diagnosed invasive breast cancers should be tested for HER2/*neu*.

Gender: Main risk factor of breast cancer is being a woman. The disease is about 100 times more common amongst women than men.

Risk factors you can change:

Lifestyle factors: For example, being overweight or obese after the menopause, physical inactivity, smoking, a high fat diet, and high alcohol consumption can play an important role in the development of breast cancer.

Hormone replacement therapy: Estrogen hormone therapy has been used to relieve symptoms of menopause and to help prevent osteoporosis but studies reveal that it also causes more risk of breast cancer.

A late first pregnancy: Women who have a late first pregnancy (after the age of 35) are more likely to develop breast cancer.

Breast cancer symptoms

Some people have no symptoms, other people may notice a change in the breast or doctor may find an unusual breast change during a physical examination [1,15]. The major symptoms of breast cancer are: New lump in the breast or underarm (armpit), thickening or swelling of part of the breast., irritation or dimpling of breast skin, redness or flaky skin in the nipple area or the breast, pulling in of the nipple or pain in the nipple area, nipple discharge other than breast milk, including blood, any change in the size or the shape of the breast, pain in the breast.

Rough Set Theory Overview

The theory of rough set theory [16] was developed by Zdzislaw Pawlak in the early 1980's, and other researchers. The main goal of the rough set analysis is induction of (learning) approximations of concepts, it is intelligent mathematical tool for managing uncertainty that is used for the discovery of data dependencies, to evaluate the importance of attributes, to discover patterns in data, to reduce redundancies, and to recognise and classify objects.

The theory has found many interesting applications. The rough set approach seems to be of fundamental importance to AI and cognitive sciences, especially in the areas of machine learning, knowledge acquisition, decision analysis, knowledge discovery from databases, decision support, engineering, environment, expert systems, banking, medicine and others. The main advantage of rough set theory in data analysis is that it does not need any preliminary or additional information about data like probability in statistics [16,17].

Rough set basics concept

RS [16-18] starts from Information System (IS) table, the columns of which are labeled by attributes, the rows by objects of interest and entries of the table are attribute values**.** An Information System is a pair S=(U,A) where U is a non-empty finite set of objects called the universe and A is a non-empty finite set of attributes such that a: U→Va for every a ∈ A. The set Va is called the value set of a, elements of U are called objects.

A Decision System (DS) is a special case of information systems, S=(U, A ∪ {d}), where d (is not element of A) is the decision attribute. The elements of A are called conditional attributes or simply conditions.

Approximation sets

Let X is a subset of U, i.e., X ⊆ U.

Lower approximation: Consists of all objects which surely belong to the set.

$\underline{R}(X)=\{x \in U\ [x]\ R\ \subseteq X\}$.

Upper approximation: Contains all objects which possibly belong to the set.

$R^{-}(X)=\{x \in U\ [x]\ R \cap X \neq \varphi\}$.

Boundary region: The difference between the upper and the lower approximation constitutes the boundary region of the rough set. Boundary positive and negative regions [19] are described as below.

$BN_R(X)= R^{-}(X) - \underline{R}(X)$.

$POS_R(X)=\underline{R}(X)$.

$NEG_R(X)=U - R^{-}(X)$.

A member of the negative region NEG (X) does not belong to X, a member of the positive region POS(X) belongs to X, and only one member of the boundary region BND(X) belongs to X, these approximation set and regions shown in Figure 3.

Approximation accuracy: The accuracy of approximation (accuracy of roughness) of any subset $X \subseteq U$ with respect to $R \subseteq A$, denoted $\alpha_R(X)$ is measured by:

$\alpha R(X)=\underline{R}(X)/R^{-}(X)$.

where X denotes the cardinality of X. For empty set φ , we define $\alpha R(\varphi)=1$.

Obviously, $0 \leq \alpha R(X) \leq 1$.

If $\alpha R(X)=1$. Thus, the set X is crisp with respect to R, otherwise

If $\alpha R(X)<1$, X is rough with respect to R.

Reducts and core

Reduct: Reducts are minimal subsets of attributes which contain a necessary portion of information from the set of all attributes [20,21]. Some attributes in decision table may be superfluous (redundant), we reduce unnecessary attributes and only keep those that are required to tell two objects that have different values on some attribute in A apart of all objects that have the same values on the attributes in B are indiscernible, or in the IND**S**(B)-relation to each other.

In order to express the idea of reduct, let B ⊆ A and a ∈ B in an information system I=(U, A) where U is the universe of objects, A is set of attributes, and R(B) is a binary relation: a is dispensable in B if R(B)=R(B –{a}); otherwise a is indispensable in B. Set B is independent if all its attributes are indispensable. B'⊆ B is a reduct of B if B' is independent and R(B')=R(B).

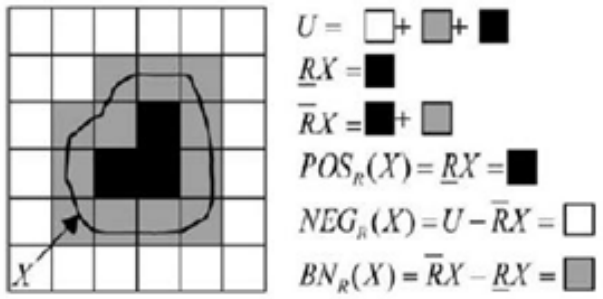

Figure 3: The approximations and regions of set *X*.
Note: POSR (X): the positive region, where
POSR (X)=R (X). i.e., POSR(X)=Lower Approximation
NEGR (X): The negative region, where:
NEGR (X)=U - R^{-}(X). i.e., NEGR (X)=Universe minus Upper Approximation, (U is a non-empty finite set of objects called the universe)
BNR (X): the boundary region, where :
BNR(X)=| R^{-}(X) - R(X) |. i.e., BNR(X)=Upper Approximation minus Lower Approximation

Core: The core is the set of all indispensable attributes, i.e., it is the intersection of all reducts.

Let Red (B) is the set of all reducts of B in an information system I=(U, A) where B ⊆ A then the core of B is defined as:

Core (B)=∩Red (B)

The core is included in every reduct, i.e., each element of the core belongs to some reduct. Thus, the core is the most important subset of attributes, for none of its elements can be removed without affecting the classification.

The Proposed Approach

Materials and methods

Breast cancer data are often presented as a table, columns of which are labeled by attributes, rows by objects of interest and entries of the table are attribute values, in a table containing information about patients suffering from a certain disease objects [22].

The data sets used in our experiments consists of 60 samples, each sample consists of ten of measurement features and decision attribute. All attributes have a data type value ranging from 0 to 5. Table 2 briefs the attributes of breast cancer dataset. Information system table contains dataset about 60 patients suffering from breast cancer disease taken from NCI, Egypt, (Table 3) the breast cancer dataset. All patients have been divided by experts into five classes (stages) corresponding to their health status. The stages are: stage 0, stage I, stage II, stage III, and stage IV.

The problem was to find the description of each class in terms of data available for each patient of this class, check whether the set of attributes is dependent or independent, find reducts for each class, and compute the core and accuracy of descriptions. We use the MATLAB programing to compute these processes [22]. The structure of the

Symbol	Attribute Name	Discrete Values					
		0	1	2	3	4	5
AG	Age Group	-	Age ≤ 35	Age > 35	-	-	-
FH	Family History	No	Yes	-	-	-	-
HER	HER2/neu Status	No	Yes	-	-	-	-
LNS	Lymph Node Status	No Node	1 < Node and < 3	3< Node <9	Node >9	-	-
TS	Tumor Size (cm)	No Tumor	Tumor <2	Tumor >2 and ≤ 5	Tumor >5	Chest Wall	-
HT	Histological Type	No	Ductal	Lobular	Ductal-Lobular	-	-
HG	Histological Grade	No	Grade I	Grade II	Grade III	-	-
ERS	Estrogen Receptor Status	-	<20	20-49	50-100	>100	-
PRS	Progesterone Receptor Status	-	<5	5-30	31-50	51-100	>100
M	Metastasis	No		-	-	-	-
D	Decision (Stage)	Stage 0	Stage I	Stage II	Stage III	Stage IV	-

Table 2: Description of breast cancer dataset.

Objects	Condition Attribution										Decision (Stage)
	AG	FH	HER	LNS	TS	HT	HG	ERS	PRS	M	
1	2	1	0	0	3	1	1	2	3	0	2
2	2	1	1	1	4	1	2	3	5	1	4
3	1	0	0	0	0	0	0	1	1	0	0
4	1	0	0	0	1	1	1	2	2	0	1
5	1	0	0	0	2	2	0	3	2	0	2
6	2	0	0	1	1	1	1	2	2	0	2
7	2	1	1	1	1	1	1	3	3	0	2
8	2	1	1	2	4	1	2	4	5	1	4
9	2	0	1	2	4	2	3	4	4	1	4
10	2	1	0	3	1	1	1	3	5	0	3
11	2	1	1	3	1	3	2	4	4	0	3
12	2	1	0	2	1	1	1	3	2	0	2
13	2	1	1	0	3	1	2	4	5	1	4
14	2	1	1	1	4	2	2	4	5	1	4
15	2	1	1	3	2	1	1	4	3	0	3
16	2	1	1	3	3	1	2	3	4	0	3
17	2	0	1	3	4	1	3	4	4	1	4
18	2	0	1	3	3	1	2	4	5	1	4
19	2	1	1	1	1	3	3	4	5	0	2
20	2	1	1	1	2	2	2	4	3	0	2
21	2	1	1	1	3	1	2	4	5	0	3
22	1	0	0	0	0	0	0	1	2	0	0
23	1	0	0	0	0	0	1	2	1	0	0
24	1	0	0	0	1	1	1	1	2	0	1
25	2	1	0	1	2	1	2	2	3	0	2
26	1	1	0	1	3	3	1	2	2	0	3
27	2	1	1	3	2	1	2	4	4	0	3
28	2	1	1	3	3	2	2	3	3	0	3
29	1	1	0	1	2	1	1	2	3	0	2
30	2	1	0	2	3	1	3	3	4	0	3
31	2	1	0	2	3	3	2	4	3	0	3
32	2	1	1	0	4	1	3	4	3	1	4
33	1	0	0	0	0	0	0	2	2	0	0
34	1	0	0	0	1	1	1	1	2	0	1
35	1	1	0	1	2	1	2	2	3	0	2
36	1	1	0	1	1	1	1	2	2	0	2
37	1	1	0	1	2	3	1	2	3	0	2
38	2	1	0	2	3	1	1	3	4	0	3
39	2	1	0	2	3	1	2	4	3	0	3
40	1	1	0	2	1	1	3	3	4	0	3

Objects	Condition Attribution										Decision (Stage)
	AG	FH	HER	LNS	TS	HT	HG	ERS	PRS	M	
41	1	0	0	0	2	2	1	3	2	0	2
42	2	0	1	2	2	1	2	4	4	0	3
43	1	0	0	0	1	2	1	3	3	0	2
44	2	1	1	2	1	1	2	3	4	0	3
45	1	1	1	2	3	1	3	3	3	0	3
46	1	0	0	1	1	1	1	2	3	0	2
47	1	0	0	0	0	0	0	1	2	0	0
48	1	0	0	0	2	1	1	2	2	0	2
49	2	1	1	2	3	1	3	4	5	1	4
50	2	1	1	2	2	1	2	3	4	0	3
51	2	1	0	0	3	3	1	2	2	0	2
52	1	1	0	0	3	2	0	3	3	0	2
53	2	0	0	0	3	1	1	2	2	0	2
54	2	1	1	1	4	3	1	4	5	0	3
55	2	1	1	3	1	1	3	4	5	1	4
56	1	0	0	0	0	0	1	1	1	0	0
57	2	1	0	1	2	1	1	2	2	0	2
58	2	0	0	1	1	1	0	3	3	0	2
59	2	1	0	2	2	1	2	3	4	0	3
60	2	1	0	2	2	2	1	4	4	0	3

Table 3: Information system for breast cancer dataset.

proposed work is represented in Figure 4.

The approximation sets program

The program is written by MATLAB programing which**:** Computes lower, upper, and accuracy approximations.

Algorithm procedure: Input: Information system (IS) as a decision table.

T=(U,A,D,f)%where $U=x_1,x_2,\ldots,x_m$, $A=a_1,a_2,\ldots,a_n$, $D=h_1,h_2,\ldots h_q$, f is information function

Begin,

Create matrix S ,size m x (n+1), from table T,

$S=\{s_1,s_2,\ldots,s_{m*(n+1)}\}$

if any object $s_x=\emptyset$ then //$(x=_{1,2,\ldots,\ m*(n+1)})$

for every object s_x do replace s_x by -1

if any vector $X=[x_1,\ldots,x_i]$contain -1

then

//i=1,2,…m

delete x_i

end {if} ; end {for}; end {if}

for reduced table T

do compute I

// I=indiscernibility relations IND(A)

if IND(A) contain redundant values

then delete redundant values

end {for}

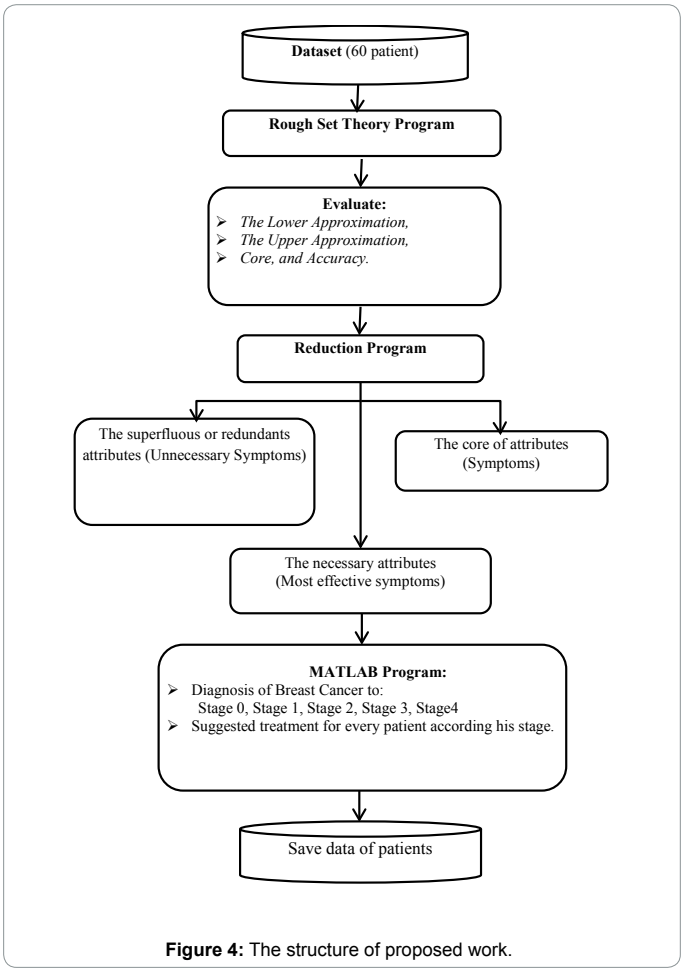

Figure 4: The structure of proposed work.

for T,I compute lower approximation $\underline{A}(X)$

if $x_i \in \underline{A}(X)$ then

create rule and insert it to Rules

end {if}

end {for}

for T,I compute upper approximation A(X)

if $x_i \in A(X)$ then

create rule and insert it to Rules

end {if}; end {for}

end {algorithm}

accuracy=$\underline{A}(X)/ A(X)$

Results of approximation sets: After constructed approximation MATLAB program, we obtained the lower, upper, and accuracy attributes description of each class (stage) results. We see that stages 0, III, and IV are crisp describable in the system, and the remaining stages are roughly describable with the accuracy given in the last column. That is to say that data (symptoms) available from the patients characterize exactly classes 0, III, and IV only, and the remaining classes not are characterized exactly by this data. The accuracy of all stages (from stage 0 to stage IV) are: 100%, 60%, 92%, 100%, and 100% respectively, as shows in Table 4.

Reduction program

According to indiscernibility and reduction condition [21-23] in order to check whether the set of attributes is dependent or not to remove repeated condition attributes in the decision table based on indiscernible and reduction of rough set theory and method.

In this section we will show reducts of attributes, i.e., minimal sets of attributes necessary for the description of these classes (stages). I used a program written in MATLAB to achieve reduction of rough functions which computes the reduct and core sets of attributes.

Reduction algorithm procedure

REDUCT (C,D)

C→the set of all conditional attributes;

D→the set of decision attributes.

R ←{}

do

T←R

$\vee_x \epsilon$ (C–R)

if $\gamma_{R\cup\{x\}}(D)>\gamma_T(D)$

T←R∪{x}

R←T

Until $\gamma_R(D)=\gamma_C(D)$

Return R

Stage Number	Lower Approx.	Upper Approx.	Accuracy
Stage 0	8	8	100%
Stage I	3	5	60%
Stage II	23	25	92%
Stage III	22	22	100%
Stage IV	10	10	100%

Table 4: The lower, upper, and accuracy of description of each stage.

Reduction results: After constructed reduction MATLAB program, we obtained that attributes 1, 2, 6, and 7 are superfluous (redundants), attributes 3, 4, 5, 8, 9, and 10. (i.e., HER2/neu status, Lymph node status, Tumor size, Estrogen Receptor status, Progesterone receptor status, and Metastasis) are the most effective symptoms in predict and accurate diagnosis of breast cancer stages, and the attributes 4, 5, and 10 (i.e., Lymph node status, Tumor size, and Metastasis) are the core of these symptoms. The summary of these results are shown in Table 5.

The results of reduction operations after removing the redundant attributes and merge the same objects are simplified in the decision table as shown in Table 6.

The Decision Rules

All decision rules which can be generated from objects represented in decision table are listed below:

Rule 1: if (TS=0) ∧ (NS=0) ∧ (M=0) ∧ (HER=0) ∧ (ERS=1 ∧ ERS=2) ∧ (PRS=1 ∧ PRS=2) then (Decision=Stage 0)→objects (patients) {3,22,23,33,47,56}

Rule 2: if (TS=1) ∧ (NS=0) ∧ (M=0) ∧ (HER=0) ∧ (ERS=1 ∧ ERS=2) ∧ (PRS=2) then (Decision=Stage 1) → objects (patients) {4,24,34,48}

Rule 3: if (TS=0 ∧ TS=1) ∧ (NS=1) ∧ (M=0) ∧ (HER=0 ∧ HER=1) ∧ (ERS=2 ∧ ERS=3) ∧ (PRS=2 ∧ PRS=3) then (Decision=Stage 2)→ objects (patients) {6,7,19,36,46,58}

Rule 4: if (TS=2) ∧ (NS=0 ∧ NS=1) ∧ (M=0) ∧ (HER=0 ∧ HER=1) ∧ (ERS=2 ∧ ERS=3) ∧ (PRS=2 ∧ PRS=3) then (Decision=Stage2)→objects (patients) {5,20,25,29,35,37,41,48,57}

Rule 5: if (TS=3) ∧ (NS=0) ∧ (M=0) ∧ (HER=0 ∧ HER=1) ∧ (ERS=2 ∧ ERS=3) ∧ (PRS=2 ∧ PRS=3)) then (Decision=Stage 2)→objects (patients) {1,51,52,53}

Rule 6: if (TS=0 ∧ TS=1 ∧ TS=2) ∧ (NS=2) ∧ (M=0) ∧ (HER=0 ∧ HER=1) ∧ (ERS=3 ∧ ERS=4) ∧ (PRS=3 ∧ PRS=4 ∧ PRS=5) then (Decision=Stage 3)∧ objects (patients) {12,40,42,44,50,59,60}

Rule 7: if (TS=3) ∧ (NS=1 ∧ NS=2) ∧ (M=0) ∧ (HER=0 ∧ HER=1) ∧ (ERS=3 ∧ ERS=4) ∧ (PRS=3 ∧ PRS=4 ∧ PRS=5) then (Decision=Stage 3)→objects (patients) {21,26,30,31,38,39,45}

Rule 8: if (TS=4) ∧ (NS=0 ∧ NS=1 ∧ NS=2 ∧ NS=3) ∧ (M=0) ∧ (HER=0 ∧ HER=1) ∧ (ERS=3 ∧ ERS=4) ∧ (PRS=3 ∧ PRS=4 ∧ PRS=5) then (Decision=Stage 3)→objects (patients) {54}

Rule 9: if (TS=0 ∧ TS=1 ∧ TS=2) ∧ (NS=3) ∧ (M=0) ∧ (HER=0 ∧ HER=1) ∧ (ERS=3 ∧ ERS=4) ∧ (PRS=3 ∧ PRS=4 ∧ PRS=5) then (Decision=Stage 3)→objects (patients) {10,11,15,16,27,28}

Rule 10: if (TS=0 ∧ TS=1 ∧ TS=2 ∧ TS=3 ∧ TS=4) ∧ (NS=0 ∧NS=1 ∧ NS=2 ∧ NS=3) ∧ (M=1) ∧ (HER=1) ∧ (ERS=3 ∧ ERS=4) ∧ (PRS=3 ∧ PRS=4 ∧ PRS=5) then (Decision=Stage4) →objects (patients) {2,8,9,13,14,17,18,32,49,55}

Breast Cancer Treatment

Breast cancer treatment [1,24-27] involves some combination

Description	Attrib. No.	Attributes (Symptoms) Name
The necessary attributes (Most effective symptoms) in predict and accurate diagnosis of breast cancer	3,4, 5, 8,9,10	*HER2/neu StatusLymph Node Status, Tumor Size, Estrogen Receptor Status,Progesterone Receptor Status , and Metastasis.*
The core of attributes (Symptoms)	4, 5,10	*Lymph Node Status,Tumor Size, and Metastasis.*
The superfluous or redundants attributes (Unnecessary)	1, 2, 6 ,7	*Age Group, Family History,Histological Type,Histological Grade.*

Table 5: Summary of reduction operations.

Objects	Condition Attribution						Decision (Stage)
	HER	LNS	TS	ERS	PRS	M	
1	0	0	3	2	3	0	2
2	1	1	4	3	5	1	4
3,56	0	0	0	1	1	0	0
4	0	0	1	2	2	0	1
5,41	0	0	2	3	2	0	2
6,36	0	1	1	2	2	0	2
7	1	1	1	3	3	0	2
8	1	2	4	4	5	1	4
9	1	2	4	4	4	1	4
10	0	3	1	3	5	0	3
11	1	3	1	4	4	0	3
12	0	2	1	3	2	0	2
13	1	0	3	4	5	1	4
14	1	1	4	4	5	1	4
15	1	3	2	4	3	0	3
16	1	3	3	3	4	0	3
17	1	3	4	4	4	1	4
18	1	3	3	4	5	1	4
19	1	1	1	4	5	0	2
20	1	1	2	4	3	0	2
21	1	1	3	4	5	0	3
22,47	0	0	0	1	2	0	0
23	0	0	0	2	1	0	0
24,34	0	0	1	1	2	0	1
25,29,35,37	0	1	2	2	3	0	2
26	0	1	3	2	2	0	3
27	1	3	2	4	4	0	3
28	1	3	3	3	3	0	3
30,38	0	2	3	3	4	0	3
31,39	0	2	3	4	3	0	3
32	1	0	4	4	3	1	4
33	0	0	0	2	2	0	0
40	0	2	1	3	4	0	3
42	1	2	2	4	4	0	3
43	0	0	1	3	3	0	2
44	1	2	1	3	4	0	3
45	1	2	3	3	3	0	3
46	0	1	1	2	3	0	2
48	0	0	2	2	2	0	2
49	1	2	3	4	5	1	4
50	1	2	2	3	4	0	3
51,53	0	0	3	2	2	0	2
52	0	0	3	3	3	0	2
54	1	1	4	4	5	0	3
55	1	3	1	4	5	1	4
57	0	1	2	2	2	0	2
58	0	1	1	3	3	0	2
59	0	2	2	3	4	0	3
60	0	2	2	4	4	0	3

Table 6: The decision table after reduction algorithm.

of surgery, radiation, chemotherapy, hormone therapy, and targeted therapy. The treatment that's best for one woman may not be best for another.

Sequence of breast cancer treatment

The common treatment sequence (pathway):

First: Surgery to remove the breast cancer and to reconstruct the breast happen during the same operation, and Axillary lymph node removal (if there).

Second: Radiation therapy usually follows surgery.

Third: Chemotherapy almost always recommended.

Fourth: Hormonal therapy if hormone receptor-positive (ER-positive or PR-positive).

Fives: Targeted therapy if HER2-positive.

There are many exceptions to this sequence, however, other treatments given. There are also many other variations in sequence and timing [24-27].

Treatment of breast cancer by stage

Breast cancer treatments [24-27] for all stages are summarized in Table 7.

Treatment during pregnancy

Breast cancer during the pregnancy [25,28] can be treated by using certain chemo drugs during the second and third trimesters (the fourth to ninth months), because chemo drugs does not increase the risk of birth defects. The safety of chemo during the first trimester (the first 3 months) of pregnancy has not been studied.

Both hormone therapy and targeted therapy can affect the fetus and should not be started until after the baby is born. Treatment of pregnancy is summarized in Figure 5.

Using MATLAB for Diagnosis and Treatment of Breast Cancer

We suggest MATLAB [7,9] program to diagnosis and treatment of breast cancer patients. This program is building according the decision rules are generated from decision table (section 4.7), and dependent on the information's taken from patients records too. The system is classifying the patients of breast cancer to: Stage 0, Stage 1, Stage 2, Stage 3, and Stage 4. In addition the system determines the appropriate treatment for each stage depending on the opinion of medical experts, researchers, and so taking into account the health status of each patient.

The front user interface GUI [28] is designed in a user friendly manner to help who use the system without any hassles. Figure 6 contains: Patient ID, Patient Name, HER2/neu Status, Lymph Node Status, Tumor Size, Estrogen Receptor Status, Progesterone Receptor Status, and Metastasis.

In report screen the patients are diagnosed to stages such (stage 0, stage 1, stage 2, stage 3, or stage 4) by matching his data with the entire database, the information's about the patients, cancer stage, and the suggested treatments all are saved in data base, and then display in Figure 7.

Conclusion

Predicting and diagnosis of breast cancer early is very important to improve the treatment quality and survivability rate of patients.

In this paper we presented a Rough Set Theory (RST) as an efficient technique for predicting and diagnosing breast cancer in early stages, where RST technique proved that it is more efficient in the field of computational biology because of the effective classification and high diagnostic capability.

In addition, People avoid detection of breast cancer because of the high cost of the tests, this leads to a diagnosed in late stages. Therefore, we have proposed this system to predict and diagnose breast cancer easy and effective way, and selects the appropriate treatment for each patient

TREATMENT OPTIONS	STAGE 0	STAGE I	STAGE II	STAGE III	STAGE IV
The Breast	None	Mastectomy OR Lumpectomy plus Radiation	Mastectomy OR Lumpectomy plus Radiation	Mastectomy OR Lumpectomy plus Radiation	Surgery, Radiation, or both may be used, depending on many individual factors
The Lymph Nodes	None	Axillary lymph node biopsy	Axillary lymph node biopsy	Axillary lymph node removal AND radiation	Enlarged lymph nodes may be treated if they are producing uncomfortable symptoms
Chemotherapy	None	None	Commonly recommended	Almost always recommended	Almost always recommended
Hormonal Therapy	None	If hormone-receptor-positive	If hormone-receptor-positive	If hormone-receptor-positive	If hormone-receptor-positive
Targeted Therapy	None	If HER2-positive.	If HER2-positive.	If HER2-positive.	If HER2-positive.
Other Parts of the Body	None	None	None	None	Radiation most commonly used
Other	Check up	Doctor decide	Doctor decide	Doctor decide	Doctor decide

Table 7: Treatment of breast cancer by stage.

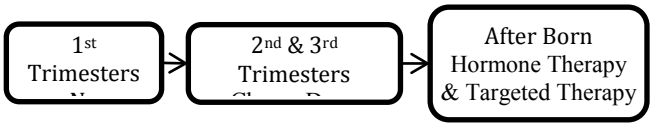

Figure 5: Treatment during pregnancy.

Figure 6: User interface GUI.

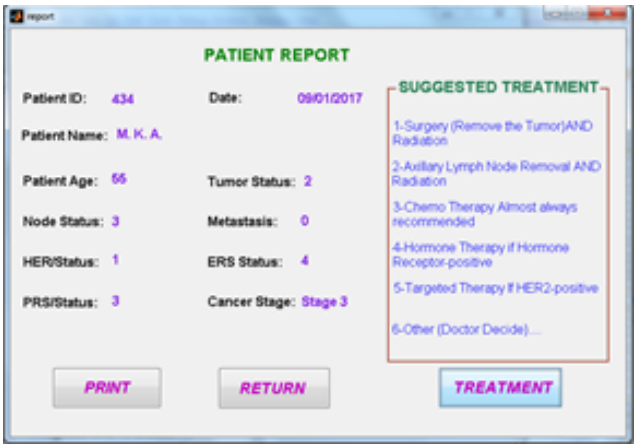

Figure 7: Report screen and suggested treatment.

according to her/his stage. This system also stores patient records in the database to be used as a source in hospitals to help doctors in follow-up the medical history and treatments for any patient.

References

1. http://www.nationalbreastcancer.org/about-breast-cancer
2. http://www.cancer.net/cancer-types/breast-cancer
3. http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-breast-cancer-types
4. http://www.cancer.gov/cancertopics/types/breast
5. http://www.ccrcal.org
6. http://stedmansonline.com/webFiles/Dict-Stedmans28/APP21.pdf
7. Carter CL, Allen C, Henson DE (1989) Relation of Tumor Size, Lymph Node Status, and Survival in 24,740 Breast Cancer Cases. Cancer 63: 181-187.
8. Abdelghani B, Erhan G (2005) Predicting Breast Cancer Survivability using Data Mining Techniques. Department of Computer Science, The George Washington University. pp: 1-4.
9. http://www.breastcancer.org/risk/factors/
10. http://www.breastcancer.org/symptoms/
11. http://www.ombc.ca.gov/forms_pubs/breast_cancer.pdf
12. Silvia R, Germano LT (2009) Rough Set Theory – Fundamental Concepts, Principals, Data Extraction, and Applications. Federal University of Rondonia, Itajuba Federal University, Brazil.
13. Pawlak Z, Grazymala-Busse JW, Slowinski R, Ziarko W (1995) Rough sets. Communications of the ACM 38: 88-95.
14. Lech P, Shusaku T, Tsau YL (2000) Rough Set Methods and Applications.
15. Hirano S, Tsumoto S (2005) Rough representation of a region of interest in medical images. Int J Approx Reasoning 40: 2334.
16. Pawlak **Z** (1984) Rough classification. Int J Man-Machine Stud 20: 469-483.
17. Pawlak **Z** (2002) Rough set theory and its applications. J Telecommun and Inform Technol 44.
18. Jan GB, Hung SN, Sinh HNPS, Jakub W (2003) Rough Set Algorithms in Classification Problem. 2: 2-97.
19. Stankov JE, Bargallo R, Namendys-Silva AS, Ramirez MT, Stankova-Ninova K, et al. (2012) Prognostic Factors and Recurrence in Breast Cancer. Int Sch Res Notices 2012:7.
20. http://www.cancer.org/cancer/breast-cancer/treatment.html
21. http://www.breastcancer.org/treatment/planning/cancer_stage/
22. http://www.cancer.org/cancer/breast-cancer/treatment/treatment-of-breast-cancer-by-stage.html
23. http://www.nationalbreastcancer.org/breast-cancer-stage-0-and-stage-1
24. http://www.cancer.org/cancer/breast-cancer/treatment/treating-breast-cancer-during-pregnancy.html
25. www.mathworks.com
26. Walter G (2015) Learning MATLAB: A Problem Solving Approach. ETH Zürich.
27. Fitzpatrick JM (2013) Computer Programming with MATLAB.
28. https://www.mathworks.com/discovery/matlab-gui.html