International Journal for Innovative Engineering and Management Research

PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL

www.ijiemr.org

IJIEMR Transactions, online available on 08th Apr 2024. Link :http://www.ijiemr.org/downloads.php?vol=Volume-13&issue=Issue4

10.48047/IJIEMR/V13/ISSUE 04/1

Title Detection of Breast Cancer by Means of Neural Networks

Volume 13, ISSUE 4, Pages: 1-9

Paper Authors **1Dumpeti Moses Manohar, 2Tirumani Prakash, 3Gonnuri Prudhvi, 4Mrs. S. Teena Mrudula**

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Detection of Breast Cancer by Means of Neural Networks

[1]Dumpeti Moses Manohar, [2]Tirumani Prakash, [3]Gonnuri Prudhvi, [4]Mrs. S. Teena Mrudula

[1,2,3,4]Dept. of Electronics and Communication, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, India
dumpetimurali93@gmail.com, tirumaniprakash116@gmail.com
gprudhvichintu@gmail.com, sunkaramrudula@gmail.com

Abstract

In this project, we address the significant and life-threatening impact of breast cancer on women globally, emphasizing the critical role of early detection in enhancing treatment efficacy and survival rates. Harnessing the power of neural networks, we aim to leverage their capacity to analyze intricate patterns and make precise predictions using a diverse range of input data. The focus is on the integration of mammographic images, patient histories, and additional clinical data for breast cancer detection. To ensure data quality and consistency, meticulous collection and preprocessing procedures are implemented, incorporating image enhancement techniques to optimize the clarity and contrast of mammograms. Patient histories and clinical data undergo careful organization and cleaning, targeting the removal of noise and outliers. Within the scope of mammographic images, the extraction of relevant features, including texture, shape, and density characteristics, contributes to a comprehensive approach to breast cancer detection. The ultimate goal is to utilize neural networks as a powerful tool in the fight against this fatal disease, ultimately reducing mortality rates in women.

In the pursuit of early breast cancer detection, we have carefully selected feature extraction methods to capture crucial information. The primary objective of our proposed approach is to attain high sensitivity and specificity, thereby minimizing false negatives and false positives in breast cancer detection. Leveraging the capabilities of neural networks, we anticipate achieving enhanced accuracy compared to conventional detection methods. Notably, our approach has yielded a Peak Signal to Noise Ratio (PSNR) of approximately 27.4%. This underscores the effectiveness of neural networks in early breast cancer detection. The emphasis on data preprocessing and feature extraction is paramount in this approach, demonstrating their pivotal roles in refining the detection process. In conclusion, our research outcomes carry substantial implications for advancing breast cancer screening methods and, ultimately, for saving lives through earlier diagnoses.

Keywords: Breast Cancer, Neural Networks, Machine Learning Algorithms, Mammographic Images, Pre-Processing, Texture, Shape, Density, Feature Extraction, PSN.

Introduction

Breast cancer claims the lives of thousands of females annually, stemming from the unregulated growth of millions of cells within the human body. When this growth becomes abnormal, it leads to the formation of a tumor, causing cells to divide and grow uncontrollably. Tumor cells can invade various systems within the body, disrupting the normal functioning of the digestive, nervous, and circulatory systems. Notably, not every tumor is cancerous, and the classification of cancer depends on the specific type of cell affected, with more than 200 known types of cancers. Focusing on breast cancer, it stands out as the most prevalent type of cancer among females worldwide. According to the National Breast Cancer Foundation, it is the most commonly diagnosed cancer in women and ranks as the second-largest cause of cancer-related deaths among women. Women typically seek medical attention due to mild to severe pain in their breasts, leading to examinations and, often, ultrasound scans. Post-scan, further procedures can be painful, especially when determining the nature of lymph node involvement—whether malignant or benign.

The process becomes crucial in distinguishing between harmful or cancerous (malignant) tumors and harmless ones that can be surgically removed (benign). The subsequent treatment plan hinges on a thorough examination of the tumor's state. This paper addresses the challenging aspects of breast cancer detection, proposing the utilization of data mining algorithms to aid in determining the tumor's state. By leveraging these algorithms, this research aims to enhance the diagnostic process, potentially providing more effective and less painful means of assessing breast cancer. Cancer, identified as a massive global public health issue, has reached alarming proportions. In 2012, the International Agency for Research on Cancer (IARC), a part of the World Health Organization (WHO), reported 8.2 million deaths attributed to cancer. The magnitude of the problem is expected to escalate, with a projected 27 million new cases by 2030. Cancer, medically defined as a malignant neoplasm, encompasses a diverse group of diseases characterized by unregulated cell growth.

The hallmark of cancer is the uncontrollable division and growth of cells, leading to the formation of malignant tumors that invade nearby parts of the body. This invasive nature allows cancer to spread throughout the body via the lymphatic systems or bloodstream. The diagnosis of cancer involves classifying tumors into two distinct types: malignant and benign. While benign tumors represent abnormal outgrowths that rarely result in a patient's death, some types may increase the risk of developing cancer. In contrast, malignant tumors are more severe, and their timely diagnosis significantly contributes to successful treatment. Predicting and diagnosing cancer play pivotal roles in improving treatment outcomes and can potentially decrease the exorbitant costs associated with medical procedures for cancer patients. Recognizing the types of tumors and their potential for malignancy not only aids in tailoring effective treatment plans but also emphasizes the importance of early detection in managing this global health challenge.

Breast cancer (BC) stands as the most frequently diagnosed cancer globally, and it claims the highest number of lives among women worldwide. Ranking as the second most common cancer for women, excluding skin cancer, BC also exhibits a notably high mortality rate compared to other cancer types. The disease initiates with a rapid and uncontrolled proliferation of a segment of breast tissue, categorized as either benign or malignant based on its potential harm. Broadly, there are two types of BC: in situ and invasive. In situ BC originates in the milk duct without spreading to other organs, even as it grows. Invasive breast cancer, in contrast, is highly aggressive, spreading to nearby organs and causing their destruction. Detecting cancerous cells before they metastasize is crucial, potentially elevating the patient's survival rate to over 97%. Early detection becomes a paramount factor in preventing mortality given the aggressive nature of BC. Within the realm of medical science, a significant challenge involves disease diagnosis through various patient tests. The evaluation of data derived from patients and the decisions made by experts play pivotal roles in the diagnostic process. The accurate diagnosis of BC remains a major challenge in the medical field, and early detection becomes imperative for preventing the potentially fatal consequences of the disease. Clinical diagnosis aids in predicting malignant cases, and timely identification can substantially increase a patient's life expectancy from 56% to 86%.

BC manifests four early signs: microcalcification, mass, architectural distortion, and breast asymmetries. Recognizing these signs and employing early detection strategies are vital in addressing this global health concern and improving the chances of successful treatment outcomes. Various common methods employed for breast cancer diagnosis (BCD) include positron emission tomography (PET), magnetic resonance imaging (MRI), CT scan, X-ray, ultrasound, photoacoustic imaging, tomography, diffuse optical tomography, elastography, electrical impedance tomography, optoacoustic imaging, ophthalmology, and mammogram. The outcomes generated by these methods aim to identify patterns that assist doctors in distinguishing between malignant and benign cases. Despite recent advances in understanding the molecular biology of breast cancer progression and the discovery of new molecular markers, histopathological analysis remains the most widely utilized method for BC diagnosis.

Detection of Breast Cancer

Breast Cancer:

Breast cancer stands as a formidable disease characterized by the unchecked proliferation of cells within the breast. Integral to the breast's anatomy are three primary components: lobules, ducts, and connective tissue. Lobules function as milk-producing glands, while ducts serve as conduits, transporting milk to the nipple. Surrounding and providing structural support to these elements is the connective tissue, composed of fibrous and fatty tissue.

Within the context of breast cancer, cells within the breast undergo abnormal growth, manifesting in various types of the disease contingent upon the affected cell types. Predominantly, breast cancers originate in either the ducts or lobules. Notably, breast cancer is not confined solely to the breast; it can extend beyond its initial site through blood vessels and lymph vessels, a process known as metastasis. This progression underscores the critical importance of comprehending the nature of breast cancer cells and their potential to disseminate throughout the body.

A multitude of factors influence the development of breast cancer. Inheritance plays a significant role, with the presence of inherited genes from families with a history of breast cancer correlating with an increased susceptibility to the disease. Additionally, the timing of menopause, either occurring at an early age (below 12) or a later age (after 55), as well as exposure to chest radiation during youth, have been associated with heightened risks of breast cancer development. Lifestyle factors, such as alcohol consumption, and the presence of Lobular Carcinoma in Situ (LCIS), further contribute to increased risks. Moreover, gender plays a significant role, with women facing a substantially higher probability of developing breast cancer compared to men. Hormone replacement therapy, while commonly utilized to alleviate menopausal symptoms, poses a potential risk for breast cancer development due to its impact on hormone levels in the body.

The prognosis and treatment of breast cancer are contingent upon various factors, including cancer stage and type. Tailored treatment plans are imperative, with interventions varying based on the timing of cancer detection. Patients may undergo singular treatment modalities or a combination thereof to optimize outcomes.

Local treatment modalities predominantly target the affected breast tissue. Surgical interventions, such as breast-conserving surgery or mastectomy, are common approaches aimed at removing the cancerous lump or the entire breast, respectively. Subsequent to surgery, radiation therapy is often administered to eliminate any residual cancer cells, employing controlled doses of radiation to target specific areas.

Systemic treatment strategies aim to address cancer cells throughout the body. Chemotherapy, utilizing potent anti-cancer drugs, targets and eradicates cancer cells, particularly beneficial in cases where cancer has metastasized beyond the breast. Hormone therapy represents another systemic treatment option, particularly effective for hormone-sensitive breast cancers. By reducing hormone levels, such as estrogen or progesterone, hormone therapy aims to counteract their stimulating effect on cancer cell growth.

So, breast cancer is a multifaceted disease influenced by a myriad of genetic, environmental, and lifestyle factors. Tailored treatment plans, encompassing both local and systemic approaches, are essential for effectively managing the disease and improving patient outcomes. Ongoing research endeavors and heightened awareness efforts are critical in advancing our understanding of breast cancer and refining treatment strategies to mitigate its impact.

Methodology

1. Early detection signs, such as masses and microcalcification clusters, play a crucial role in identifying breast cancer. Microcalcification refers to tiny mineral deposits within breast tissue, resembling small white spots, with potential links to cancer. Masses can vary, encompassing cysts, non-cancerous solid tumors, or potentially cancerous growths.

2. Detecting cancer is challenging due to the subtle appearance and ambiguous margins of abnormalities in normal breast tissues. Automated tools assist radiologists in early breast cancer detection.

3. Following detection, cancer is classified into three categories: Normal, Malignant, and Benign.

4. An adaptive mean filter is employed to remove noise from the image, as it excels among spatial filters in distinguishing fine details from noise.

5. The Adaptive Median Filter performs spatial processing to identify pixels affected

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

by impulse noise, classifying them by comparing each pixel to its surrounding neighbors.

6. The neighborhood size and threshold for comparison are adjustable. Pixels labeled as impulse noise are replaced by the median pixel value of neighbors passing the noise labeling test.

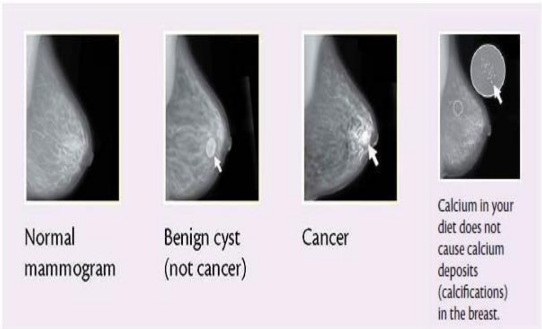
7. Image preprocessing involves converting it into a grayscale image using the rgb2gray() function, applying adaptive mean filtering, and converting the image into unsigned integer 8 using uint8().

8. GMM segmentation is performed on the preprocessed image with two regions, two GMM components, and a maximum of ten iterations. Additionally, k-means segmentation with k=2 is implemented.

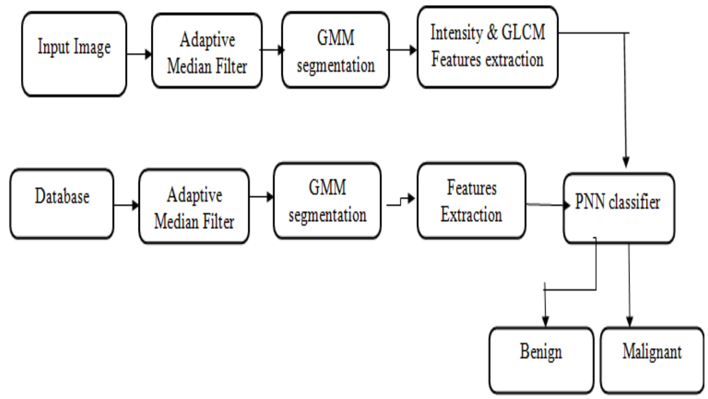
9. Finally, the HMRF-EM (Hidden Markov Random Field Model) and its Expectation-Maximization Algorithm are incorporated into the methodology, aiming to optimize breast cancer detection and segmentation in medical imaging.

10. Difference between Normal, Benign and Malignant:

The below figure gives the detailed information about the normal, benign (not cancerous-starting) and malignant:



Block Diagram



Evaluation Metrics:

Adaptive Median Filter:

The adaptive median filter is utilized for spatial processing to diminish noise in an image. In this filtering process, each pixel in the image is compared to its surrounding pixels. If the value of a pixel significantly differs from the majority of its surrounding pixels, it is identified as noise. Subsequently, the filtering algorithm replaces the noise pixel with the median value derived from the surrounding pixels. This iterative process continues until all noise pixels in the image have been effectively removed.

The adaptive median filter addresses the limitations encountered by the standard median filter. Its primary advantage lies in the variable size of the kernel surrounding the corrupted image, leading to a more favorable output result. Another significant benefit of the adaptive filter is its departure from the median filter's approach of replacing all pixel values with the median value.

The functioning of the adaptive filter involves a two-step process. In the first step, it determines the median value for the kernel, while in the second step, it assesses whether the current pixel value represents impulse (salt and pepper noise). If the pixel value is deemed corrupted, it undergoes a change by adopting the median value; otherwise, the grayscale pixel's value remains unchanged. Ensuring that only pixels with impulse noise are altered, all other pixel values are preserved through this process.

2. GMM SEGMENTATION:

In recent years, much attention has been devoted to the standard Gaussian mixture model (GMM) as a well-known method for image segmentation. This model assumes a common prior distribution that independently generates pixel labels.

An advantage of the standard GMM lies in its requirement for a small number of parameters for learning. Additionally, these parameters can be efficiently estimated by adopting the EM algorithm to maximize the log-likelihood function.

In this technique, the standard GMM serves as a robust approach to image segmentation, leveraging its advantages in parameter efficiency and effective parameter estimation through the EM algorithm.

Let $\boldsymbol{Xi}$; $\boldsymbol{i=1,2,3 \ldots N}$; denote the observation at the i-th pixel of an image. Labels are denoted by Ω1,Ω1,...,ΩK. Consider the problem of estimating the posterior probability of xi belonging to label Ωj. If we assume that xi is drawn independently from the distribution, then the standard GMM, assumes that the density function at an observation xi is given by:

$$f(x_i|\pi,\theta) = \sum_{j=1}^{K} \pi_j \emptyset(x_i|\emptyset_j)$$

The graphical representation of a Gaussian mixture model for a set of N pixel xi is shown in Figure 2.3. Where Π = {π1, π2, ..., πK}, and πj is the prior distribution of the pixel xi belonging to the label Ωj, which satisfies the constraints:

$$0 \le \pi_j \le 1 \; and \; \sum_{j=1}^{K} \pi_j = 1$$

Each Gaussian distribution $\emptyset(\boldsymbol{X_i}\,|\boldsymbol{\theta_j})$ is called a component of the mixture. For the case of a single real-valued variable xi, the Gaussian distribution has its own mean μj and covariance σj and is defined by:

$$\emptyset(x_i|\emptyset_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} exp\left(-\frac{(x_i-\mu_j)^2}{2\sigma_j^2}\right)$$

where $\boldsymbol{\theta j} = \{\boldsymbol{\mu j},\ \boldsymbol{\sigma j}\}$ The observation xi in Eq.(2.21) is modelled as statistically independent. And the joint conditional density of the data set X =(x1, x2, ..., xN) can be modeled as:

$$p(X|\pi,\emptyset) = \prod_{i=1}^{N} f(x_i|\pi,\emptyset) = \prod_{i=1}^{N}\left[\sum_{j=1}^{K} \pi_j \emptyset(x_i|\emptyset_j)\right]$$

Given the joint conditional density from Eq.(2.24), the log-likelihood function of the standard GMM is given by:

$$L(\emptyset,\pi|X) = \sum_{i=1}^{N} log\left\{\sum_{j=1}^{K} \pi_j \emptyset(x_i|\emptyset_j)\right\}$$

Where Θ = {Θj}; j=1,2,...,K. As can be seen from the likelihood function.

3. Intensity and GLCM Features Extraction:

The Gray-Level Co-occurrence Matrix (GLCM) is a texture analysis technique in digital image processing that evaluates the relationship between neighboring pixels based on gray intensity, distance, and angle. It extracts statistical texture parameters such as entropy, inverse difference moment, angular second moment, and correlation by quantifying how frequently pairs of pixels with specific values and spatial relationships occur in an image. GLCM is crucial for feature extraction before image classification, providing insights into texture characteristics like contrast, homogeneity, energy, correlation, and entropy. It categorizes statistics into first, second, and higher-order statistics based on intensity combinations, with GLCM focusing on second-order statistical texture features. While higher-order textures are theoretically possible, they are rarely implemented due to computational complexity. GLCM matrices, with rows and columns equal to the number of gray levels in the image, contain relative frequency or probability values for pixel pairs at specified distances and angles. To manage dimensionality, the number of gray levels is often reduced. Overall, GLCM offers valuable insights into texture analysis and is widely utilized across various applications in image processing.

The extraction features of image include Angular Second Moment, Inverse Different Moment**,** Entropy, Correlation,

4. PNN CLASSIFIER:

The Probabilistic Neural Network (PNN), first proposed by Specht in 1990, is a classifier designed to map input patterns into different class levels. It can be adapted for a more general function approximation. The network is structured as a multilayer feed-forward network, comprising the input layer, pattern layer, summation layer, and the output layer. Specifically tailored for classification problems, Probabilistic Neural Networks (PNN) function as a type of radial basis network.

The PNN, rooted in Bayes theory and developed by Specht, estimates the probability of a sample belonging to a learned category. Spread plays a crucial role, and if it's near zero, the network acts as a nearest neighbor classifier. The PNN consists of four layers: an input layer, a pattern layer, a summation layer, and a decision layer. Most of the terms in the PNN are calculated from the training data, contributing to its effectiveness in classification tasks.

5. Hidden Markov Random Field Model and its Expectation- Maximization Algorithm:

The concept of a hidden Markov random field model is rooted in hidden Markov models (HMM). HMMs are stochastic processes generated by a Markov chain, wherein the state sequence is not observed directly but rather through a series of observations. Each observation is considered a stochastic function of the state sequence. The underlying Markov chain undergoes state transitions based on a $l \times l$ transition probability matrix, where l is the number of states. HMMs have found successful applications in speech recognition and handwritten script recognition.

Original HMMs were designed for 1D Markov chains with first-order neighborhood systems, making them unsuitable for direct use in 2D/3D problems like image segmentation. In response to this limitation, we explore a special case of an HMM, where the underlying stochastic process is a Markov random field (MRF) instead of a Markov chain. This adaptation allows for the model's

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

applicability to multidimensional problems. We term this special case a hidden Markov random field (HMRF) model.

Simulation and Results

For Person 1:

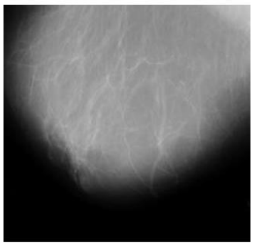


Fig-1: Mammograph input image of person1

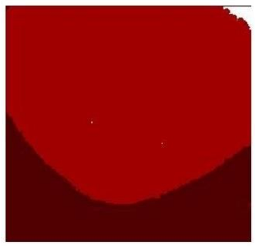


Fig-2: Output Segmented Image represented normal (3) Breast

For Person 2:

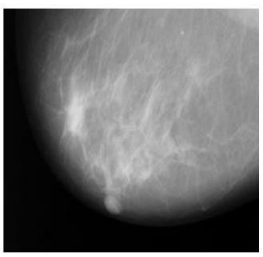


Fig-3: Mammograph input image Of person2

Fig-4: Output Segmented Image Represented Benign (1) Breast i.e., starting stage of cancer compared with fig in Sec III

For Person 3:

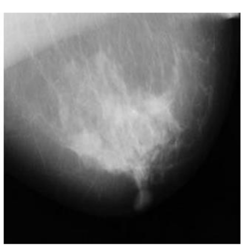


Fig-5: Mammograph input image person3

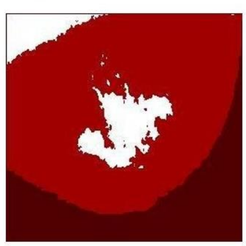


Fig-6: Output Segmented Image of represented malignant (2) Breast

For Person 4:

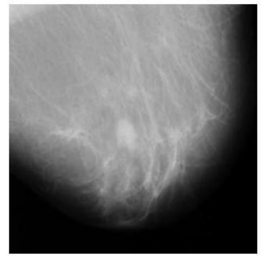


Fig-7: Mammograph input image Of person4

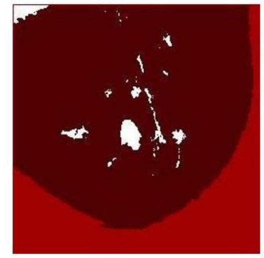


Fig-8: Output Segmented Image Represented Benign (1) Breast

For Person 5:

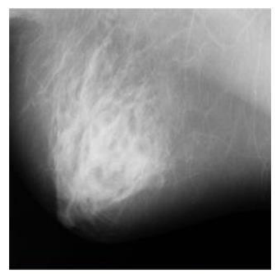


Fig9: Mammograph input image Of person5

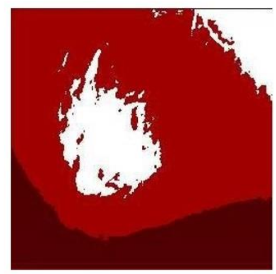


Fig-10: Output Segmented Image represented malignant (2) Breast

Conclusion

In conclusion, the fusion of neural networks with advanced image processing techniques such as the Adaptive Median Filter and GMM Segmentation, bolstered by a robust classifier, presents a groundbreaking approach to breast cancer detection. By enhancing the quality of mammographic images, isolating regions of interest, and discerning intricate patterns, this comprehensive system offers improved sensitivity and specificity, leading to reduced false positives and false negatives. The potential for earlier diagnosis afforded by this methodology underscores its significance in improving patient outcomes and underscores the pivotal role of artificial intelligence in revolutionizing healthcare. However, the efficacy of such systems hinges on the quality and quantity of data and continuous optimization, necessitating ongoing research and validation to ensure reliability in real-world clinical applications.

References

1. Bernard WS, Christopher PW. World Cancer Report 2014. International Agency for Research on Cancer. WHO Press. World Health Organization; Switzerland. 2014.

2. Bray F, Jemal A, Grey N, Ferlay J, Forman D. Global cancer transitions according to the human development index (2008-2030): A population base study. The Lancet Oncology. 2012;13(8):790-801.

3.Weinberg RA, editor. One Renegade Cell: How Cancer Begins. Basic Books; New York, USA. 1999.

4.Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. Cell. 96;87(2):159-170.

5.Steyeberg E. Clinical Prediction Models: A Practical Approach to Development, Validation and Updating. Springer Science & Business Media; New York, USA. 2008.

6.Siegel R,Ward E, Brawley O, Jemal A. Cancer statistics. CA:A Cancer Journal for Clinicians.2011;61(4):212-236.

7.Miller KD, Siegel RL, Lin CC, Mariotto AB, Kramer JL, Rowland JH, JA. Cancer treatment and survivorship statistics. CA: A Cancer Journal for Clinicians. 2016;66(4):271-289.

8.Love SM, Barsky SH. Breast-duct endoscopy to study stages of cancerous breast disease. The Lancet. 1996;348(9033):997-999.

9.HortobagyiGN. Treatment of breast cancer New England Journal of Medicine. 1998;339(14):974-984.

10.Osborne C, Ostir GV, Du X, Peek MK, Goodwin JS. . Breast Cancer Research and Treatment. 2005;93(1):41-47.