**THE IMPLEMENTATION OF STOCHASTIC MODEL TO THE STUDY OF BREAST CANCER IN RIVERS STATE OF NIGERIA**

**ABSTRACT**

The Paper applied Stochastic model to the study of breast cancer data in Rivers State by considering Cancer data from the University of Port Harcourt Teaching Hospital, Cancer Registry Unit.

The sample data was transformed to a 3 transition count matrix denoted as M while its corresponding transition probability matrix was denoted by **P**. The stochastic principle was used to determine the ages of women that were susceptible to breast cancer and the number of women who transited from one state of cancer to another and also the number of women who left the system as a result of death. The plot of the raw data against the time showed that as the year went by, the number of breast cancer patient increased.

However, this study investigated the discrepancies that existed in regions among women with breast cancers and also estimated the age bracket of women that would be mostly affected with breast cancer in the nearest future.

This study proffered solution to factors that could reduce the risk of breast cancer in Rivers State and across the globe.

***Keywords: Stochastic Principle, Breast Cancer, Rivers State and Transition Matrix***

**1.0 INTRODUCTION**

Breast cancer is the abnormal proliferation of cells usually uncontrollably beyond their usual boundaries which can go on to invade adjoining body organs **(World Health Organization,2025)**.

Breast cancer is the abnormal, uncontrolled increase of cells originating from the breast tissue. It is more common for breast cancers to emanate from the lining of the milk ducts or the lobules supplying the ducts (**Ganesh et al, 2010).** It is the most common cause of cancer in women and also the leading cause of cancer death among women especially in less developed countries.

However, in more developed countries like the United States, lung cancer is the leading cause of death among women followed by breast cancer **(Farley et al,2018)** It is estimated that over 1 million new cases of breast are diagnosed every year constituting serious health threat to women all over the world4.

Prevalence of breast cancer varies with different regions; hence the Western countries are said to have higher number of cases compared with their African and Asian counterparts **(Obiora & Abu 2019).**

In contrast to the United States with an incidence rate of 101/100,000, globally, American Indian and Alaska Native women have the lowest incidence rate of developing breast cancer at 21/100,000**(Porter,2008)**. Even in Africa, this regional variation still exists, with studies showing that 27% of breast cancers were localized in women from the Northern part of Africa like Egypt while 15% was reported in Sub Sahara region of Africa**(Parkin et al 2010**). In central Africa, incidence rate is 27/100,000 and 39/100,000 in the Southern part of Africa and that of China is 25.9/100,000**(Zeng et al,2010)**

**2.0 RELATED REVIEW OF RELEVANT LITERATURE**

Breast cancer cases are gradually increasing in African countries, probably as a result of gradual adoption of more westernized lifestyle **(Obiora & Abu,2019).** It is believed that breast cancer cases in Africa (Nigeria inclusive) will continue to increase as a result of ageing, increase growth of the population, people adopting unhealthy lifestyles and absence of health-related intervention against breast cancers **(Abu & Obiora,2019, and Briton et al, 2014).**

Late presentation, poor screening and diagnostic modalities, unavailability of quality health care infrastructures, poor health practices and funding among other factors are some of the factors that could lead to poor prognosis, increase morbidity and a low survival outcome in Nigerian women **(Parkin et al, 2010 and Smith & Jones,2022).** It is important to note that not all tumors are cancers, hence there are benign and malignant tumors; cancers are malignant tumors and cancer cell in itself has the capacity to transform and evolve into stages **(John et al,2024).**

As a result of the resemblance that exists between cancer cells and normal cells from which they originate, they are not often detected by the immune system especially if it is weakened **(Pace & Shulman,2016).** Cancer cells are usually formed as a result of mutations of DNA or RNA. These mutations may occur spontaneously or may follow chronic exposures to nuclear radiation, electromagnetic radiation (X-rays, Gamma rays), viruses, bacteria, parasites, etc. It is thought that invasive cancers arise from series of molecular alterations at the cellular level. It is this alteration that result in breast epithelial cells that assume immortality and uncontrolled growth. There are other risk factors which could make a person more susceptible to developing breast cancer. Increasing age and female sex are known risk factors for the development of breast cancer. Consequently, breast cancer is 100 times more common in women than in men; however, men tend to have poorer prognosis mainly as a result of delay in diagnosis **(World Health Organization, 2020**). A positive family history is another known risk factor including mutation carriers of BRCA 1 and 2. Reproductive and hormonal factor such as late age at first pregnancy, early menarche, late age at cessation of menses, nulliparity have all been shown to be risk factors for the emergence of breast cancer **(Cancer Statistics, 2022 and Kelsey,1996).** This is not unrelated to the effects of increased and unopposed estrogen in these women **(Colditz & Rosner,2000).**

Also, a previous history of breast cancer in one breast is associated with a 3-4-fold increase in occurrence of cancer in the other breast **(Key et al,2002 and Page & Jensen,1994)**. Cancer is a disease in which some of the body’s cells grow uncontrollably and spread to other parts of the body however protective life styles against breast cancer can include consumption of diet rich in grains, fruits and vegetables, reduced saturated fats and low in alcohol **(Joel et al,2023).**

Early breast cancers are usually asymptomatic, and pain is usually not present at first. Breast cancers may present as a lump with breast asymmetry, skin dimpling / changes around the skin of the breast, nipple deviation or related changes, nipple discharge which could be bloody, axillary lump.

Diagnosis and screening for breast cancer could involve any of the following or a combination; -physical examination of both breasts, mammography, ultrasonography, magnetic resonance imaging, biopsy.

Management of breast cancers could include any of or a combination of the following – surgery, radiation therapy, hormone therapy, chemotherapy.

However, this study investigated the discrepancies that existed in regions among women with breast cancers and also estimated the age bracket of women that would be mostly affected with breast cancer in the nearest future and also obtained the state transition probabilities.

The study is limited to the application of Stochastic model on the stages of breast cancer in Rivers State and it is relevant because the model is designed to predict the transition of the stages of breast cancer in Rivers State.

**3.0 METHODOLOGY**

**3.1 AN EXPLANATION OF THE STOCHASTIC MODEL (MARKOV CHAIN)**

Stochastic model like Markov chain had been widely used in different fields to model systems that could be in one or more states with certain probabilities though these probabilities are subject to changes over time. Cancer operates in stages and a model of this nature should be employed in its study.

The study employed the stochastic modeling of breast cancer by describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. Thus, if the current state of a process is known, its past states need not be required to predict its future state. This invariably means that a process’s past and future states are independent of its present state. One usefulness of this property of Markov model is that some parameters are likely to be redundant. Transition matrix and state transition probabilities were developed in the case of this study. The

transition matrix is usually given as:

**The transition probability matrix P is matrix where each element gives the probability of moving from state i to state j in one or more steps.**

**The columns of do not in general sum up to one (1) but the rows of P**

**Each row of P is the probability distribution relating to a transition from state I to state j.**

**The n- step transition matrix may be obtained by multiplying the matrix P by itself n times thus, the probability vectors are said to be state vectors of a Markov Chain.**

**However, means that the system is in state i at nth step. Similarly, is the initial state vector of the Markov chain. Expressing this in another form, we can say that given as the state vector and P as the transition probability matrix then one can write a Markov model for the stages of Breast Cancer in Rivers State as:**

The ROWS represent **NOW or FROM** which is given by

The columns represent **NEXT or TO** which is referred to **( )**

The entry is the conditional probability that NEXT = , given that NOW ; the probability of going FROM state **TO** state is;

The transition matrix P must list all possible states in the state space S. P is a square matrix

because both and take values in the state space S (of size N)

The rows of P should sum to one (1).

This simply states that must take one of the listed values.

**3.1 THE WORKING OF THE MARKOV MODEL**

For the purpose of this study, the raw data was transformed into transition probability matrix. Thus, three states were considered. A random walk is said to exhibit the Markov property if the position of the movement at time (n+1) depends only upon the position of the movement at n.

Let denote the position of the random movement at time n, then the equation becomes …………………………………………………………..3.1

Where is independent of so that the state at time depends only upon the state of at time n.

This implies that each for is a probability row vector describing every possible transition from state i to any other existing N possible states in the process.

Thus, for every I, …………3.2

Generally, a random movement exists in N possible states in the system hence will depend on the whole sequence of random variables starting with the initial **value10** yields

………………………………3.3

**3.3 ASSUMPTIONS OF THE MARKOV MODEL**

. for all in the matrix

. All the elements on a specific line of the transition probability matrix must be positive for the matrix to be closer to one (1).

. Transition probability relates to the finite state in different states with the property **that next state or** the state of the system after it goes through the transition depends only on current state and is independent of any other prior states

. Transition probability matrix attains a stationary distribution that is, the probability distribution that the matrix remains unchanged after applying the transitions defined by the matrix. Thus, if there is a stationary distribution ; then hence that distribution is considered invariant and corresponds to the eigenvalue of one (1).

**3.4 MODEL VALIDITY**

To test for the validity of this model, the graph of the **actual breast cancer data** from 2016 to 2024 were compared with the graph of **expected cancer cases.**

**3.5 MODEL SUITABILITY**

We checked for the Markovian suitability using the hypothesis of goodness of fit. To achieve this, the actual breast cancer data from 2016 to 2024 was compared with the expected breast cancer data using the Markovian model. The Chi- square statistic of 2.7806 was compared with the statistical significance of 9.488 on the critical region of 0.05 with 4 degrees of freedom.

With the acceptance of the hypothesis, we deduce that the actual transition probability matrix from 2016 – 2024 is fitted with expected transition probability prepared using Markov method.

The statistical test of independence was used to check whether the changes are dependent or independent

**3.6 MODEL FORMATION**

In this research, the ages of breast cancer patients were grouped into stages (states). Thus, we considered 0- 44(state 1) 45-58(state 2) and 58 & above (state 3).

The model is presented below;

3

2

1

The horizontal arrow signifies transition from one state to the other while the vertical arrow signifies exiting or leaving the system(death).

**4.0 RESULT ANALYSIS ONE**

The data used for this work were collected from University of Port Harcourt Teaching Hospital Cancer Registry Unit. A total observation of 568 was recorded within the period under consideration.

The collected data were transformed to a 3 x 3 matrix called the state transition matrix and probability transition matrix represented by M and P respectively.

Three age brackets were considered as below:

**TABLE 1 State and age - wise scenario of Breast Cancer**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| NUMBER OF STATES | 0 - 44 | 45-58 | 58 & ABOVE | TOTAL |
| 1 | 38 | 86 | 75 | 199 |
| 2 | 37 | 63 | 58 | 158 |
| 3 | 48 | 97 | 66 | 211 |
| TOTAL | 123 | 227 | 189 | 568 |

We generated a simultaneous equation to determine the value of which represented the state that would be susceptible to breast cancer in the future. Hence linear equations representing the limiting probabilities will be generated. The equations would be of the form

Hence

This gave rise to three distinct linear equations as shown below:

From eqns. 1, 2 and 3 we have;

Rearranging and subtracting eqn. (6) from eqn. (5)

From eqn. (4) we have;

Recall that

By replacement we have;

Taking L.C.M, we have;

From eqn. (7) hence

From the values of we deuced that at the long run, the number of People with breast cancer at state 2 (45-58) will equal the number of People at state 3 (55 & above).

**4.1 RESULT**

**TABLE 2** **People with breast cancer at state 2**

|  |  |  |  |
| --- | --- | --- | --- |
| **Observed Frequencies** | | | **ROW TOTAL** |
| **23**  **22**  **21** | **87**  **133**  **124** | **45**  **60**  **53** | **155**  **215**  **198** |
| **66** | **344** | **158** | **568** |

**The expected frequencies were obtained using:**

=

Thus, with the degree of freedom (df) of (r-1) (c-1), we have 4 degrees of freedom from our contingency table above. Hence with a level of significance of 0.05 and 4 degree of freedom we have a critical value of 9.488.

Since the Chi -square statistic (2.7806) is less than the critical value (9.488), then the difference between the observed and the expected distribution is statistically not significant. Since the Chi-square statistic (2.7806) is less than the critical value (9.488), then the difference between the observed and expected distribution is statistically not significant hence the breast cancer data does not permit me to reject the null hypothesis that the variables are unrelated and does not also provide support for the alternative hypothesis that the variables are related.

**4. 2 CHARTS OF CANCERS CASES WITHIN THE 9 YEARS PERIOD**

**FIGURE 1- BREAST CANCER INCIDENCE IN RIVERS STATE**

NUMBER OF PERIODS

The graph in figure I, above showed that breast cancer invasion increased yearly throughout the nine years of investigation and rose to an unprecedented level during the Covid periods.

We also considered number of Periods (years) against breast cancer’s states (where series 1,2 and 3 referred to state 1, 2 & 3 respectively).

**FIGURE 2 - TIME PLOT OF BREAST CANCER DATA**

The plot agreed that the age 45-58 (state 2) recorded the highest breast Cancer rate or prevalence.

**FIGURE 3- NUMBER OF WOMEN WITH BREAST CANCER ACCORDING TO THEIR AGE**

**FIGURE 4- LINE GRAPH OF THE OBSERVED FREQUENCIES AGAINST NUMBER OF PERIODS**



**FIGURE 5- THE BAR CHART OF THE OBSERVED FREQUENCIES AGAINST THE NUMBER OF PERIODS**



**FIGURE 6 -THE LINE GRAPH OF THE EXPECTED FREQUENCIES OF BREAST CANCER AGAINST NUMBER OF PERIODS**



**FIGURE 7– THE BAR CHART OF THE EXPECTED FREQUENCIES AGAINST THE NUMBER OF BREAST CANCES CASES**



4.3 **DISCUSSION**

A six-year probability matrix was randomly selected in order to determine the state transition matrices of breast cancer prevalent in Rivers State. This was taken to be from **state one to state two and from state two to state three and from state three remaining at state 3(absorbing state). From the analysis it was observed that** in the year 2030 that is, P6 the transition probability from state 1 to state 2 was 0.6146 and from state 2 to state 3 was 0.2772. Also, the transition probability from state 2 to state 3 was 0.2772, this was the same probability in transiting from state 3 to state 3(0.2772). Similarly, in the year 2048, that is, P24, the transition probability from state 1 to state 2 would be 0.5265 and from state 2 to state 3 would be 0.2315 and from state 2 to state 2 would be 0.5201 while from state 2 to state 3 would be 0.2287, from state 3 to state 1 would be 0.0928, from state 3 to state 2 would be 0.5261 and lastly from state 3 to state 3 would be 0.2313. The surprising finding here is that number of breast cancer patients in state 3 would equal the number in state 2 as was indicated in the analysis;

**This analysis sends a serious warning to the People of Rivers State and Nigeria as a Country.**

The bar charts of the observed and expected frequencies supported the claim that the difference between the observed and expected distributions **were not statistically significant leading to the non-rejection of the null hypothesis that variables are unrelated.**

**4.4 CONCLUSION**

Breast cancer is the most common cancer in women and also the leading cause of cancer death among women. Breast cancer invasion increased yearly throughout the nine years of study and if urgent and immediate attention is not taken, most women in Rivers State might be living with cancer.

Women between the ages of 45-58 (state 2) suffered most from breast cancer, the analysis revealed that by the year 2048, women between 58 years and above would equal the number of women aged 45-58 living with cancer.

**4.5 RECOMMENDATIONS**

We recommend protective life styles against breast cancer such as consumption of diet rich in grains, fruits and vegetables, reduced saturated fats and low in alcohol. Also, women should be guided on the choice of contraceptives, clamping down on contraceptives that pose higher risks for occurrence of breast cancer like the oral contraceptive pills.

Also, early detection and screening is highly recommended through less expensive modalities like physical examination, mammography and ultrasonography.

**Disclaimer (Artificial intelligence)**

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

1.

2.

3.

**REFERENCES**

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018). *Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*. CA Cancer J Clin, 68 (6):394-424

Brinton LA, Figueroa JD, Awuah B, et al (2014). Breast Cancer in Sub-Sahara Africa; Opportunities for prevention. Breast Cancer Research Treat, 144(3):467-478.

Cancer-its Various Types Along With Causes, Symptoms, Treatments and Stages, in Cancer Info Guide. 2009. [http://www.cancer-info-guide.com/[Ref](http://www.cancer-info-guide.com/%5bRef) list]

Colditz GA, Rosner B. (2000**).** *Cumulative Risk of Breast Cancer to age 70 years According to Risk Factor Status Data from the Nurses’ Health Study*. Am J Epidemiol; 152 (10): 950-64.

Elinor Washbrook (2006): Risk Factors and Epidemiology of Breast Cancer; Women’s Health Medicine, Volume 3, Issue 1, pages 8 -14

Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM (2010). *Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008.* Int J Cancer, 127;2893-2917.

Ganesh N.S, Rahul D, Jyotsana S, Piush S, K.K Sharma (2010). *Various Types and Management of Breast Cancer: An Overview.* Journal of Advanced Pharmaceutical Technology & Research, 1(2): 109-126

Joel S. Brown, Sarah R. Amend, Robert H. Austin, Robert A. Gatenby, Emma U. Hammarlund & Kenneth J. Pienta: *Updating the Definition of Cancer; American Association for Cancer Research* 2023 July 6,2(11):1142-1147

John F.Speer, Victor E. Petrosky, Michael W. Retsky Robert h. Wardwell (2024): *A Stochastic Numerical Model of Breast Cancer Growth that Simulates Clinical Data;* Cancer Research 2024

44:4124-4130

Key T, Appleby P, Barnes I, Reeves G. (2002). *Endogenous Sex Hormones and Breast Cancer in Postmenopausal Women: Reanalysis of Nine Prospective Studies*. J Natl Cancer Inst; 94(8):606-16.

Obiora C.C, Abu E.K (2019). *Breast Cancer in Rivers State, Nigeria: Ten-Year Review of the Port Harcourt Cancer Registry.* SA Journal of Oncology, Porter P (2008). ‘Westernizing’ women’s risk? *Breast cancer in lower income countries*. N Engl J Med., 358(3):213-216

Pace LE, Shulman LN (2016). Breast Cancer in Sub-Sahara Africa: Challenges and Opportunities to Reduce Mortality. Oncologist; 21(6):739-744.

Page DL, Jensen RA (1994). *Evaluation and Management of High Risk and Premalignant Lesions of the Breast.* World J Surg; 18(1):32-8.

Parkin D, Ferlay J, Hamdi-Cherif M et al. Cancer in Africa: *Epidemiology and Prevention. International Agency for Research on Cancer* (IARC Sci Publ. Sci.), 153:1- 414

Porter P.(2008): Westernizing Women’s Risk? Breast Cancer in lower income countries: N Engl. J Med., 358(3):213-216

Smith, J. and Johnson, P. (2022): *The Impact of AI on Medical Data Analysis: A Case Study of IBM Watson Health.* Journal of Health Informatics*,4*5,145-156

World Health Organization (2022): Cancer Statistics

World Health Organization. Cancer (2024). <https://www.who.int/health-topics/cancer#tab=tab_1>

Zeng H, Zheng R, Zhang S, Zou X, Chen W (2014). *Female Breast Cancer Statistics of 2010 in China: Estimates Based on Data From 145 Population-Based Cancer Registries*. J Thoracic Dis., 6(5):466-470