

Original Research

Received 10 June 2026

Revised 20 July 2026

Accepted 02 Aug 2026

Natural Resources for Human
Health 2026, Vol. 6 (15s): 1496–1516

KEYWORDS:

Breast cancer, Survival analysis,
Retrospective cohort study,
Kaplan–Meier analysis, Cox
proportional hazards regression,
Socio-demographic factors, Early
detection, Treatment outcomes

A Study on Breast Cancer Patients survival in Dibrugarh District, Assam, India: Breast Cancer Survival

Bornali Khound¹, Dhruba Das², Gayatri Gogoi^{*3}, Porikhit Borpujari⁴

¹M.Phil, Research Scholar, Department of Statistics, Dibrugarh University, Dibrugarh-786004, Assam(India). Email- bornalikhound2019@gmail.com, Orcid ID. 0009-0009-2826-1468

²PhD, Assistant Professor, Department of Statistics, Dibrugarh University, Dibrugarh-786004, Assam(India). Email- dhrubadas16@gmail.com Orcid ID:0000-0001-8546-0193

³Professor, Department of Pathology, Assam Medical College and Hospital, Assam(India). Email: gayatrigogoi303@gmail.com Orcid ID:0000-0001-9845-1835

⁴Technical Officer, Department of Pathology, Assam Medical College and Hospital, Dibrugarh-786002, Assam(India). Email: borpujariporikhit@gmail.com Orcid ID: 0009-0005-8387-4911

Corresponding Author: Dr. Gayatri Gogoi

E-mail address: gayatrigogoi@amch.edu.in

ABSTRACT: Aim: Breast cancer remains a significant health issue globally, particularly in developing countries. This study investigates the survival rates of breast cancer patients in the Dibrugarh district, Assam, India, focusing on socio-demographic, clinical, and treatment factors.

Method: A retrospective cohort study was conducted among patients with histopathologically confirmed breast cancer who were treated between 2017 and 2022 at a tertiary care center in Dibrugarh, Assam. Data were extracted from hospital records, including demographic, reproductive, clinical, and treatment variables of the patients. Survival time was measured from the date of diagnosis to death or the last follow-up. Kaplan–Meier curves, log-rank tests, and Cox proportional hazards regression were applied to identify significant predictors of mortality.

Results: The mean age of participants was 46.27 $\pm$ 29.37 years with 78 deaths occurred. The median survival time was 167.0 [139.0, 195.0] weeks. The study found that breast cancer is the most commonly diagnosed cancer among women in the region, with a high prevalence in older age groups. Risk factors influencing survival include illiteracy (HR:1.684, 95% CI :1.042-2.712, $p<0.05$), older age (HR= 2.354, 95% CI: 1.201-4.613, $p<0.05$), Late Stages (HR=2.367, 95% CI: 1.017-5.502, $p<0.05$) and lack of treatment (HR: 5.063, 95% CI:1.831-14.051, $p<0.05$). The survival rates vary significantly with the stage at diagnosis, with early-stage detection showing a substantially higher survival rate compared to late-stage diagnosis.

Conclusion: The study highlights the urgent need for improved awareness, early detection, and accessible treatment options to enhance survival rates. It also underscores the importance of addressing socio-demographic and lifestyle factors to improve patient outcomes. The findings are intended to inform healthcare policy and targeted interventions to reduce the breast cancer burden in this region.

INTRODUCTION:

Breast cancer (BC) is the most common malignancy among women worldwide and is a leading cause of global cancer incidence. In 2024, an estimated 2.4 million women diagnosed with breast cancer and 694000 deaths globally (Ferlay et al., 2026). In India, about 200,000 new cases of breast cancer were reported in 2020, making it the most prevalent cancer among women in and representing nearly 13.5% of the country's cancer burden. (ICMR, 2017). In the eastern part of India, breast cancer is one of the commonly diagnosed non-communicable diseases among women. Current trends point out that breast cancer is more common in the younger age group and the incidence rates were 68.8% and 64.9% of all women suffering from breast cancer in Guwahati and Dibrugarh, which are two major cities of North-East (NE) India respectively, who were below 50 years of age. (Breast Cancer statistics, 2024). Breast cancer brings unique challenge due to late-stage presentation, limited screening programs, and socio-economic disparities through the women (Sathwara, Balasubramaniam et al., 2017). Breast cancer has surpassed cervical cancer to become a most common cancer among women in India, accounting for nearly 27% of all female cancers (Chaturvedi et al., 2022). The growing of cancer burden affects steadily, particularly in urban and semi-urban regions, primarily due to lifestyle changes, delayed diagnosis, and inadequate awareness (Malvia et al., 2017). In Western countries it is estimated that breast cancer diagnosed among women is above 60 years, in India, the disease occurs a decade earlier, affecting women in their 40s and 50s, thereby contributing to higher disease severity and mortality (Liu et al., 2022).

Socio-demographic factors such as Education, Age, Clinical factor like BMI, Cancer Stages, and lifestyle habits, along with treatment-related determinants like delayed surgery, Chemotherapy and other systemic therapies, play a vital role in survival outcomes in this region (Devi et al., 2021).

Several Indian Studies has reported that the overall 5-year survival rate for breast cancer patients is very high for Stage I patients (95%, stage I), compared to those in advanced stage (21% Stage IV) (Sathwara et al., 2017). Worldwide data has shown that in lower-middle-income countries, especially among rural populations, the impact of cancer is high due to low awareness, limited access to affordable healthcare, and poor disease prognosis. In India, the survival rate for breast cancer patients is lower compared to Western countries due to early onset, delayed diagnosis, and late treatment (Mehrotra et al., 2022). Older Age at diagnosis and the advanced stage of breast cancer are significantly associated with patient survival (Sheikh et al., 2015, Taplin et al., 2004). Dietary factors also have a considerable impact on the development of breast cancer (Holmes et al., 1999). A study conducted in the NE region of India found that the overall survival rate for breast cancer patients is 93.3% at one year and 70.8% at three years, with an average survival time of 38 months. The risk of developing breast cancer is higher among women aged 35–50 than among younger or older women. Non-parametric and semi-parametric analysis of survival data showed that factors like age, educational status, and residence are significantly associated with the risk of mortality from breast cancer (Bhattacharjee S, Deka S, 2022, Dudley et al., 2016).

The NE region of India has the highest cancer incidence rate in the country, facing challenges such as inadequate treatment facilities, low awareness for early detection, and limited access to timely treatment. Factors such as socio-demographic, reproductive, and genetic profiles, along with lifestyle, are strongly associated with breast cancer survival. Given the regional disparities and rising burden of breast cancer in Assam, it is essential to identify the Socio-demographic and treatment related factors on breast cancer survival and severity. This study focuses on breast cancer patients in Dibrugarh District of Assam, India, and examines risk factors such as educational status, age at diagnosis, body mass index (BMI), cancer stage and breastfeeding status with clinical and treatment factors like BMI and cancer stage and various treatments, as well as genetic factors such as family history of cancer on breast cancer survival. These factors influence the risk of breast cancer and are associated with patient survival (Ewertz et al., 1991).

This study applied Survival analysis technique to study the time between an event's start and its occurrence, with the event of interest being death (Kaplan et al., 1958). In the context of breast cancer, survival analysis helps in understanding the likelihood of a patient surviving the disease (Siegel et al., 2018). The main objective of survival analysis is to identify the factors that influence the time to death, enabling healthcare professionals to develop more effective treatment strategies and improve patient outcomes (Clark et al., 1988). Different survival analysis methods, including the Kaplan-Meier method, log-rank test, and Cox proportional hazard regression, are commonly used to analyze survival functions.

Objectives of the Study

The present regional-level study was conducted in a tertiary care centre of Dibrugarh, Assam, India with the following objective:

1. To investigate the survival time and factors affecting survival of breast cancer patients with reference to socio-demographic, reproductive, genetic, lifestyle and treatment related factors.

MATERIALS AND METHODS

Data collection and follow up: This was a retrospective cohort study conducted at a tertiary care centre of Upper Assam. The period of study was from 2017-2022. The study was carried out for Dibrugarh District, Assam, India. A predesigned, pretested questionnaire was used for the collection of data. Various treatment strategies play a vital role in determining patients' outcomes. However, the effects of treatment-related factors on survival time often vary depending on individual patients' characteristics. Since individual patient characteristics include information about socio-demographic factors (educational status, income, and age) and clinical factors (breastfeeding status, stage of cancer, and family history) that influence health outcomes and the effectiveness of treatments. All these factors and treatment factors were collected from hospital records. The patients were considered censored if they were alive beyond 31 December 2022, died due to other causes, or were lost to follow-up. After collecting the hospital record, a telephonic conversation and home visit survey were conducted to ensure the survival status of the patients, date of death (if she dies due to breast cancer), and continuation of treatment. Also, extension of disease includes stages, surrounding organs/tissues, distant organs, etc. that are collected. Survival days were estimated from the day of diagnosis until death, loss to follow-up, or the end of 2022. The study was approved by the Institutional Ethical Committee for Human of the tertiary care hospital. Informed consent was taken from the study participants. Participants were assured of confidentiality and anonymity.

Participants

Patients were enrolled based on inclusion and exclusion criteria. Pre-enrolment counselling included commentary on the benefits of the study to the participants and its contribution to the general population. The inclusion criteria were

- (1) Age: ≥ 18 years.
- (2) Histo-pathologically confirmed invasive breast carcinoma cases.
- (3) Ability to provide informed consent.
- (4) Available for follow-up assessments.

The Exclusion criteria were:

- (1) Benign / non-cancerous breast tumors or any other neoplasms.
- (2) Unwilling to participate in the study.
- (3) Incomplete basic information

Sample size

To determine the sample size for this study, Yamane's (1967) formula for finite populations was employed

$$n = \frac{N}{1 + Ne^2}$$

Where n is the sample size, N is the total target population ($N = 442$), and e is the margin of error ($e = 0.05$). The formula yielded a required sample size of approximately 209.98, which was rounded to a minimum target of 210 participants. A non-probability purposive sampling technique was subsequently used to select these 210 participants from the broader population. This method was selected because it allows the researcher to intentionally choose individuals who possess specific characteristics, knowledge, or clinical experiences essential to answering the research questions. Participants were chosen based on predetermined inclusion criteria to ensure data quality and relevance to the study's specific focus..

METHODOLOGY:

In this study, we have considered a total of eight socio-demographic, clinical variables and treatment variables as probable risk factors of breast cancer, and time to death from breast cancer. A brief description of the factors is depicted in **Table1**.

Association of the factors with the occurrence of breast cancer

The variable measuring the time to death due to breast cancer is a non-normal continuous variable. All the other variables used in the study are qualitative with two or more categories. As such, to compare it with various factors, the Mann-Whitney U test and Kruskal-Wallis test are used.

Comparison of the Risk Factors with the time to death due to breast cancer

We have considered the time (in weeks) before death from breast cancer as survival time. All the patients under study have either died, been lost to follow-up or are still undergoing treatment. The specific difficulties relating to survival analysis arise largely from the fact that only some individuals have experienced the event and, subsequently, survival times will be unknown for a subset of the study group. For example, in collecting information about the time of death from breast cancer, some individuals may still not have died or may recover, or be lost to follow-up. As such, for patients who were lost to follow-up or still undergoing treatment, the exact times of death for them are unknown. These are right-censored observations or times because the true unobserved event is to the right of the censoring time.

Suppose T denotes the time of death of a patient suffering from breast cancer. Then the survivorship function denoted by $S(t)$ is defined as the probability that a patient survives up to time t , i.e., dies after time t , and is given by,

$$\begin{aligned} S(t) &= P(\text{that a patient dies after time } t) \\ &= P(T > t) \end{aligned} \quad (1.1)$$

When the data contains censored observations, the survivorship function (surviving from death) is estimated as

$$\hat{S}(t) = \frac{\text{Number of patients dies after time } t}{\text{Total number of patients}} \quad (1.2)$$

A simple comparison of the survivorship function can be achieved by drawing graphs of the time of death function. Thus, to compare the impact of a risk factor, the survivorship curves for different levels of the factor are drawn using Equation (1.2). Also, Kaplan-Meier curves give the overall survival rate of Breast cancer patients.

Comparison of the survival times for different risk factors

Suppose $S_1(t)$ and $S_2(t)$ be the survivorship functions under two different levels of a factor. To test the null hypothesis that the two levels of the factor have the same survival time, i.e.,

$$H_0: S_1(t) = S_2(t),$$

To compare the survival times for different levels of the risk factors, we have used the non-parametric Log-Rank test also called the Mental-Heinzen test.

Impact of risk factors on the time to death due to breast cancer

To examine the combined effect of the covariates on the time to death due to breast cancer, Cox's Proportional Hazard Model is used. The model evaluates simultaneously the effect of several covariates on survival, which can be suitably adapted in the present context. The Cox model is a semi-parametric regression model that defines the hazard function $h(t, \mathbf{x})$ by a set of p covariates $\mathbf{x}' = (x_1, x_2, \dots, x_p)$ as

$$h(t, \mathbf{x}) = h_0(t) e^{\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p} = h_0(t) \beta' \mathbf{x} \quad (1.3)$$

Where, t represents the time to death. The term $h_0(t)$ is called the baseline hazard. It corresponds to the value of the hazard if all the x_i are equal to zero, i.e., without any influence of covariates. The coefficients $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ measure

the impact (i.e., the effect size) of covariates. The logarithm of the hazard ratio in (1.4) gives the numerical estimates of the regression coefficients β_i rather than in the shape of $h(t, x)$

$$\log h(t, x) = \log h_0(t) + \beta'x$$

$$\text{Or } \log(HR_0) = \log\left(\frac{\text{Group Hazard}}{\text{Baseline Hazard}}\right) = \sum_{i=1}^k \beta_i x_i \quad (1.4)$$

Which is a linear function of the covariates $x_1, x_2, \dots, x_p$.

RESULTS AND DISCUSSION:

The present 5-year retrospective cohort study (2017-2022) analyzed 210 cases of breast cancer patients, revealing an average age at diagnosis of 46.2 years, with a median age of 45 years and an age range between 20 and 80 years. The median survival time was 167.0 [139.0, 195.0] weeks (Table 2). A total of 78 (37.15%) individuals experienced death due to breast cancer by the end of the study. The overall survival rates for breast cancer participants in our cohort were 92% at 1 year, 68% at 3 years, and 40% at 5 years (Figure 1). Out of 210 breast cancer cases recruited to this study, the majority of patients were Hindus (94.20%), literate (80.0%), and housewives (98.0%). Demographic characteristics show that most patients had a normal BMI (70.4%), with an average weight of 51.6 kg and average BMI of 22.21. The study showed a predominantly non-vegetarian diet (100%) with minimal alcohol consumption (5.3%) and no smoking (0%), with 9.5% occasional tobacco use. Family history of cancer was relatively low (9.1%). Breast cancer patients' stages were distributed as follows: Stage I (22.4%), Stage II (22.8%), Stage III (28.1%), and Stage IV (26.7%), with cancer occurring more or less equally on both breasts (Left: 53.4%, Right: 46.2%). Also, 81.9% of patients were breastfed, and 96.2% were married. Additionally, 30% of patients received only surgery, 56.2% received only chemotherapy, 13.8% received chemotherapy with surgery and other, and 8.09% did not receive any kind of treatment. The variable detail of breast cancer patients are presented in Table 1.

Survival outcomes:

We examined the survival outcomes of 210 patients diagnosed with breast cancer. The overall survival rates for these patients in our study were 93% at one year, 67% at three years, and 40% at five years.

These survival outcomes were linked to several factors, including educational level, age at diagnosis, body mass index (BMI), cancer stage, breastfeeding history, family history of cancer, and types of treatments received. Education level was found to be a key predictor. Individuals with graduate education had survival rates of 98% at one year, 69% at three years, and 49% at five years, while those who were illiterate had survival rates of 82%, 51%, and 20%, respectively (Figure 2(a)). Age also played a role, with patients under 45 years old showing survival rates of 96%, 90%, and 38% at one, three, and five years, whereas patients aged 45 and above had survival rates of 88%, 48%, and 15%, respectively (Figure 2(b)). Individuals with underweight or normal BMI had higher survival rates compared to those who were obese, with survival rates of 69%, 40%, and 5% at one, three, and five years (Figure 2(c)). Cancer stage at diagnosis was strongly associated with survival, with Stage I patients having 96% at one year and 75% at five years, while Stage IV patients had 89% at one year and 5% at five years (Figure 2(d)). Women who had a history of breastfeeding showed better survival rates at 93% and 59% at one and three years, compared to 81% and 40% for those without this history (Figure 2(e)). Lastly, women with a family history of cancer had lower survival rates compared to those without, with 68% survival at five years versus 29% for those with a family history (Figure 2(f)). Lastly, women undergone various treatment has better survival 95% at 1st year and 70% at 3 years) compared to patients who had not taken any kind of treatment (60% and 30%)(Figure2(g)). The detailed survival rate of these factors is shown in Table 2.

By using equation (1.2) several survivorship functions are depicted under various risk factors via, Education status, Age of breast cancer, etc. Figures 2(a), 2(b), 2(c),2(d),2(e),2(f) and 2(g) gives the survival curves for the risk factors included in the study. Also the overall survival is depicted under Figure 1.

The results of Mann-whitney U test, Kruskalwalli's test and p values shows that the breast cancer patients with two factors, Income and Breastfeeding Status are positively associated with the occurrences of breast cancer. Breast cancer patients income (Mann-whitney U=10441.5, p=0.028) has significant difference between high and low income patients as p value is less than 0.05 at 95% confidence interval. Also for Breastfeeding Status (Mann-whitney U=1673.0, p=0.024) has significant

difference between Breastfeeding and non breastfeeding women as p value is less than 0.05 at 95% confidence interval. The result of the test is given in **Table 3**.

Next, to identify the factors that are associated with the survival time of patients from breast cancer using the Log Rank test. The results of the test are as follows:

The mean and median survival times of each covariate, and different groups of Education, Age at diagnosis, BMI, Cancer Stage, Breastfeeding, family history and Treatments had statistically different survival probabilities. This is also evident from Table 5 as the p-values corresponding to these factors are less than 0.05 (at 5% level of significance).

The survival analysis revealed significant differences in mean and median survival times across various factors and groups ($p < 0.05$). Notably, illiterate patients had lower survival probabilities than literate patients (Log-Rank test value: 19.7, $p=0.007$). Additionally, middle- or older-onset age patients (45&Above years) had shorter survival times than younger patients (Log-Rank test value: 12.2, $p=0.005$). Cancer stage at diagnosis was also a crucial factor, with patients diagnosed at Stage III or IV having shorter survival times than those diagnosed at Stage I or II (Log-Rank test value: 16.3, $p=0.001$). Furthermore, overweight or obese patients had shorter survival times than underweight or normal weight patients (Log-Rank test value: 8.8, $p=0.032$). Women who breastfeed had longer survival times than those who did not (Log-Rank test value: 9.93, $p=0.002$). Lastly, cancer treatment also had a significant association with survival time (Log-Rank test value: 22.6, $p=0.00005$). Detailed survival analysis of mean and median survival times of these factors is detailed in **Table 4**

The Cox's Proportional Hazard Model was used to examine the combined effect of these significant factors on survival time. The model included seven significant factors viz., Age at diagnosis, educational status, BMI, breastfeeding status, cancer stage, family history and treatment of cancer. The results of Cox's Proportional Hazard Model showed that p-values for all three tests: Likelihood ratio test, Wald test and Score test are significant, indicating that the model is significant, and hence the omnibus null hypothesis that all β 's are 0 is rejected. So it can be inferred that there was an overall significant relationship between this set of covariates and survival time.

The results showed illiterate patients had a 68.4% increased hazard of death compared to literate patients ($p=0.034$). Older-aged patients (45 & above) had a 134.5% increased hazard of death compared to younger patients ($p=0.012$). Patients diagnosed at Stage IV had a 2.367 times increased hazard of death ($p=0.045$) compared to early stage. No Breastfeeding status had a 84.3% increased hazard of death compared to participants with positive breastfeeding status ($p=0.045$). Patients who underwent complete treatment (surgery with chemotherapy, and others) had substantially better survival than those who underwent incomplete or no treatment ($HR \approx 5.067(1.831, 14.015)$, $p < 0.01$).

Although BMI and family history of cancer were insignificant, as their hazard ratios indicated increased risk. **Table 5** gives detailed outcomes of the proportional hazard model with p-values ($Pr(Z > |z|)$) along with the coefficients (β) and their standard errors (SE), exponentiation of coefficients (e_{β}) and their confidence intervals (CI) and critical values (z_c).

As Breast cancer is one of the most common malignant neoplasm affecting women globally and a leading cause of cancer-related deaths among women. The incidence of breast cancer in India has risen consistently across all regions, with an annual increase of 0.5% to 2% (Agarwal & Ramakant, 2008). There is a high prevalence of malignancy in younger ages, with nearly 48% of patients diagnosed before the age of 50 (Madhab et al., 2018), which contrasts sharply with Western populations, where the peak incidence typically occurs among women in their sixties (Khokhar, 2012). In Northeast India, breast cancer is the most common cancer among females, accounting for 14.5% of cases where incidence rate was also estimated at 27.81 per 100,000, nearly twice the non-Northeast rate of 13.62 (Khokhar, 2012; Rajbongshi et al., 2018). The present 5-year retrospective cohort study of 210 breast cancer cases (2017-2022) provides critical insights into the unique profile of breast cancer patients in Dibrugarh District of Assam, India, particularly in relation to age, socio-demographic, lifestyle, clinical as well as treatment related factors. These findings indicate regional variations in breast cancer trends from both Indian and Global patterns.

This study revealed that the average age of breast cancer diagnosis is 46.2 years, significantly younger than in Western countries, where incidence usually peaks around 60 years [Giaquinto et al., 2022]. These results are corroborated with previous studies from NE India, which reported a mean age at diagnosis is ranging from approximately 44 to 48 years, respectively, reflecting a trend that is consistently a decade younger than the median age of 62 years typically observed in Western

populations (Gogoi et al.,2016 ; Leong et al.,2010). Genetic traits, reproductive patterns, dietary and modern lifestyle habits in the Indian populations may contribute to this early onset [Sharma et al., 2014]. Notably, the younger average age in Northeast India emphasizes the critical need for localized, early screening initiatives.

Furthermore, our study shows a clear link between age at diagnosis and survival rates. Younger patients (Below 45) showed a 95% 1-year and 70% 3-year survival rate, whereas older age patients (45 and above) had reduced survival rates (89% at 1 year and 49% at 3 years) (Table 2). The result also showed that the median survival time of younger patients below 45 is 238 weeks, which is much higher than the median survival time of 154 weeks of older patients(45 and above years) (Table 4). This indicates that the risk of death increases with increasing age. Also, the risk of death for older patients is 2.345 times higher than for younger patients. (Table 5). These results align with the study showing that increased age at diagnosis was associated with decreased survival (Brandt et al., 2015). This trend highlights that younger individuals generally experience fewer co-morbidities and greater tolerance to aggressive treatments [(Mehrotra & Yadav, 2022)]. These findings suggest that age-appropriate screening and management strategies are crucial to optimize outcomes for older age patients.

The socio-demographic analysis of the study highlights that most participants were married Hindu housewives with low incomes. Among them, 80.0% were literate with primary education (43.8%), secondary education (27.2%) and graduates (9.05%) (Table 1). This profile aligns with demographic characteristics in NE India, where women are responsible for household roles, often engaging in informal labour, especially in rural and semi-urban areas [Gajalakshmi]. Notably, literacy among women in this region was generally high compared to other parts of India, potentially attributed to a strong emphasis on female education in states like Assam and Mizoram [Amerian society].

Data confirms that educational attainment is a strong independent predictor of breast cancer survival. Graduates demonstrated higher survival rates at all intervals: 99% at 1 year, 70% at 3 years, and 50% at 5 years. In contrast, the illiterate subgroup had lower survival rates: 81% at 1 year, 50% at 3 years, and 20% at 5 years (Table 2). The study also implies that Illiterate patients have shorter survival time compared to literate patients. The median survival time is 147 weeks for illiterate patients compared to 221 weeks for Graduate patients, indicating illiteracy is possibly associated with shorter survival. (Table 4). Also, the time to death from breast cancer decreased by 68.4% compared to literate patients, if we hold the other factors constant. (Table 5). This highlights the crucial role of education in health outcomes, likely due to improved health literacy, timely healthcare utilization and adherence to treatment protocols (Bradly et al.,2026). Higher literacy levels can assist in seeking medical attention for breast abnormalities, contributing to better health awareness. These findings highlight the importance of educational initiatives targeting breast cancer awareness and self-breast examinations, particularly in populations with lower literacy rates. By addressing education-related inequalities, healthcare access and screening behaviours can be improved, ultimately improving breast cancer outcomes in NE India.

A significant proportion (70.4%) of study participants had a normal BMI, averaging 22.21 kg/m² (Table 1). These results differ from global findings where breast cancer incidence is often higher among individuals with overweight or obesity, particularly in post-menopausal women (Singh et al.,2011) Notably, breast cancer patients in India tend to have lower BMIs compared to Western populations, potentially due to differing dietary habits and lifestyles.

Minimal alcohol consumption (5.3%) and no smoking history among participants reflected local norms and cultural practices that may influence breast cancer risk profiles. However, tobacco use was reported in 9.5% of participants, primarily through chewing (44.71%), which a concern is given high tobacco consumption rates in certain groups in NE India (World Cancer Research Fund,2018, Sanker et al.,2021,Singal et al.,2023).

BMI also correlated with survival outcomes, with Normal BMI patients showing higher survival rates (90% at 1 year, 65% at 3 years, and 21% at 5 years) compared to obese individuals (Table 2). The present study shows that there is a statistically significant relationship between BMI and breast cancer survival ($p=0.032$) (Table 4). The study also signifies that median survival time is 139 weeks for obese patients compared to 195weeks for normal patients. This suggests that obesity is associated with poorer overall breast cancer survival. On the other hand, the Cox model showed no evidence of significant correlation between the patients' survival time and BMI. These results contradict the studies conducted by Singh *et. al.* (2011) and Blair *et. al.* (2019). Research suggests that higher BMI is generally associated with poorer survival in post-menopausal breast cancer patients due to factors such as co-morbidity like diabetes mellitus, hormonal changes and increased inflammation (Nahmias et al.,2023). The better outcomes observed in underweight and normal-weight patients may reflect a

combination of metabolic, hormonal and treatment-related factors, underscoring the need for personalized nutritional and physical health support in breast cancer care.

Cancer staging data revealed that most patients (22.8%) were diagnosed at Stage II, followed by Stage III (28.1%) and Stage IV (26.7%) (Table 1). This pattern was similar to the national trend, where limited early screening and delayed healthcare access often resulted in late-stage diagnosis. NE India faces unique healthcare access challenges, including rural infrastructure limitations, long travel distances and socio-economic constraints, which may delay diagnosis and treatment initiation [Howlader N, 2019]. Cancer stage at diagnosis is very important for survival. In this study, 1-year survival rates were high across all stages, but 5-year survival rates decreased significantly with advancing stage: Stage I showed a 75% 5-year survival rate, compared to only 5% for Stage IV (Table 2). This finding aligns with data linking early-stage detection to improved survival rates in breast cancer (Dhillon et al., 2018). Our research also demonstrates that women's survival time after being diagnosed with breast cancer is significantly affected by their cancer stage. The average survival time for patients with Stage IV is 2.367 times shorter than that of the reference group (Stage I), when all other parameter held constant (Table 5). This suggests that overall survival declines with increasing clinical stage. Similar finding specify that late-stage breast cancer patients are more likely to experience the event (death). Approximately 60% of breast cancer cases in India are diagnosed at stage III or IV as most women present to the hospital only when there is a noticeable secondary change like a large mass or changes to skin or chest wall, with minor symptoms being ignored (Mehrotra et al., 2022). This condition is now the principal cause of breast cancer-related deaths among Indian women (Thomas et al., 2025, Norris et al., 2025, Srivastava et al., 2025). Stage-wise comparison of estimated survival outcome with national and international research is detailed in **Table 6.1 & 6.2** respectively (Nair 1993, Feleke et al., 2022).

A sustainable portion (81.9%) of participants had breastfeed history (Table 1), which was associated with reduced breast cancer risk. High breastfeeding rates, particularly in NE India, may contribute to potential protective effects that delay cancer onset. Women with a breastfeeding history demonstrated better survival outcomes (92% at 1 year and 58% at 3 years) than those who had not breastfed (78% and 41%) (**Table 2**). Also, the patients with a breastfeeding history had more survival time (212 weeks) than those who had not breastfed (149 weeks). i.e., the factor Breastfeeding is a significant prognostic factor for survival time of Breast Cancer patients ($p=0.015$) (**Table 4**). On the other hand, the Cox model showed 84.3% risk of hazard to non breastfeeding patients. This result is similar to the findings of Pakseresht *et al.* (2009). There is increasing risk of breastfeeding's protective impact, which may be due to hormonal and cellular changes in breast tissue during lactation that lower the likelihood of tumour genesis (Gajalakshmi et al., 2009). This implies that breastfeeding may have post-diagnosis benefits, potentially influencing tumour biology or enhancing immune response.

Women who get cancer and do not have a family history of cancer show better survival outcomes (80% 1-year and 70% 3-year) than those with a family history of cancer (95% 1-year, 70% 3-year and 40% 5-year) (Table 2). Moreover, women without family history have more survival time (206 Weeks) than family History of Cancer (140 Weeks) (Table 4). The findings also demonstrate that family history determines the survival of Breast Cancer patients. Patients who have a breast cancer history in their family are more likely to be at risk of breast cancer as compared to patients who do not have a breast cancer history (Brewer *et al.*, 2017).

Chang *et al.* (2008) performed a similar study on associations between family history of cancer at the time of diagnosis and risk of all-cause mortality after cancer diagnosis, adjusting for established prognostic factors. The hazard ratios for breast cancer mortality were 0.98 (95% CI=0.84-1.15) for having at least one first- or second-degree relative with cancer, and 0.85 (95% CI=0.70-1.02) for having at least one first-degree relative with cancer, compared with having no such family history. Another study by Yusuf *et al.* (2024) found that age and family history had significant effects on the survival time ($p=0.003$), and the risk of death was 3.244 (95% CI: 0.334, 2.020). This suggests that a patient with a family history of cancer is associated with survival outcome (Liu *et al.*, 2021).

The strongest modifiable factor identified was treatment completion. Recent studies focused on breast cancer in Northeast India and the surrounding region (specifically data from Population-Based Cancer Registries, including those in the Northeast) highlight that Patients receiving multimodal treatment combining surgery with chemotherapy and other treatments- significantly improve 5-year survival rates, while incomplete treatment or no treatment are major factors contributing to high mortality. These Key Findings from Studies in Northeast India (e.g., Assam, Meghalaya, Tripura, and Manipur) reported poor survival among patients with disrupted or partial treatment in the north-eastern region [Kumar *et al.* 2024].

al.,2019]. Similar associations have been authenticated in other Indian centres, where economic constraints, long travel distances, and lack of awareness contribute to treatment default [Sharma et al.,2025]. Internationally, large trials and registry data have consistently verified the survival benefit of guideline-based multimodality therapy [Early Breast Cancer Trialists' Collaborative Group,2011].

CONCLUSION AND FUTURE SCOPE OF THE STUDY:

The purpose of the study was to identify factors affecting the time to death among breast cancer patients using survival analysis. For studying the survival affects of 210 numbers of breast cancer patients with Eight numbers of risk factors were collected from a tertiary centre of upper Assam, Dibrugarh. As time to death of breast cancer patients are not normally distributed so non-parametric tests like Mann-Whitney's U test(when only two patients groups are to be compared) and Kendall's (for more than two patient group) were compared. Kaplan Meier survival curves were deployed to visualized and compared patients survival for various groups and their significant difference were tested using Log-rank test. The study identifies that significant risk factors that are associated with occurrence of breast cancer are Educational status, Age at Diagnosis of breast cancer, Family history of cancer, cancer stage and treatment. Furthermore, the combined effect of the significant risk factors for survival time of breast cancer patients is studied using Cox's regression. The overall findings of the analysis are that educational status with illiterate patients, Age at diagnosis(45 and Above),family history, No treatment are the significant risk factor that can impact the survivability of patients with breast cancer. Though the use of non-parametric Kaplan Merrier survival analysis is some cases result a similar or higher estimate of survival rate, the parametric analysis remains more powerful, robust and efficient. The Cox regression model provides better significant results to the survival probability of breast cancer patients. So, future researchers could make use of this model. The findings highlight the importance of early detection, education, better lifestyle and awareness in improving breast cancer survival rates. The study recommended that an improved method for estimating the survival probability and analysis of breast cancer data. Future more, studies also need to assess the level of awareness, treatment and control of these risk factors.

ACKNOWLEDGEMENTS:

The authors acknowledge the technical assistance provided by the Department of Statistics, Dibrugarh University and the Department of Pathology, Assam Medical College & Hospital (AMCH) for their constant support.

AUTHOR CONTRIBUTIONS

Khound B: Conceptualization, Data Curation, Investigation.

Gogoi G : Supervision, Methodology, Writing—Review & Editing.

Das D: Supervision, Data Curation, Formal Analysis.

Borpujari P : Data Curation and Writing

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

ETHICS STATEMENT

The study was approved by the Institutional Ethical Committee for Human Research of the tertiary care centre (Ref No: AMC/EC/1334). Informed consent was obtained from the study participants. The participants were assured of confidentiality and anonymity.

DATA AVAILABILITY STATEMENT

Data supporting the findings of this study are available from Corresponding author Dr. Gayatri Gogoi (Supervisor) upon reasonable request.

Ethical approval:

The study was approved by the Institutional Ethical Committee for Human research of the tertiary care centre (Ref No: AMC/EC/1334). Informed consents were taken from the study participants. Participants were assured of confidentiality and anonymity.

Patient declaration of consent statement:

Informed consent was obtained from each participant before enrolment in the study. Patients' anonymity and confidentiality was strictly maintained

REFERENCES:

- [1] Ferlay J EM, Lam F, Laversanne M, Colombet M, Mery L et al. Global Cancer Observatory: Cancer Today [website]. International Agency for Research on Cancer; 2026 (<https://gco.iarc.fr/today>)
- [2] National Cancer Registry Programme NCRP. Three-year report of the population based cancer registries 2012–2014. Bengaluru: National Cancer Registry Programme, Indian Council of Medical Research (ICMR); 2017.
- [3] *Breast Cancer statistics from Kolkata, Guvahati* (https://www.breastcancerindia.net/statistics/stat_kolkata.html/ accessed on 06-11-2024)
- [4] Sathwara, J. A., Balasubramaniam, G., Bobdey, S. C., Jain, A., & Saoba, S. (2017). Sociodemographic factors and late-stage diagnosis of breast cancer in India: A hospital-based study. *Indian Journal of Medical and Paediatric Oncology*, 38(03), 277-281.
- [5] Chaturvedi, M., Sathishkumar, K., Lakshminarayana, S. K., Nath, A., Das, P., & Mathur, P. (2022). Women cancers in India: Incidence, trends and their clinical extent from the National Cancer Registry Programme. *Cancer Epidemiology*, 80, 102248.
- [6] Malvia, S., Bagadi, S. A., Dubey, U. S., & Saxena, S. (2017). Epidemiology of breast cancer in Indian women. *Asia-Pacific Journal of Clinical Oncology*, 13(4), 289-295.
- [7] Liu, H., Shi, S., Gao, J., Guo, J., Li, M., & Wang, L. (2022). Analysis of risk factors associated with breast cancer in women: a systematic review and meta-analysis. *Translational Cancer Research*, 11(5), 1344.
- [8] Shanker, N., Mathur, P., Das, P., Sathishkumar, K., Shalini, A. M., & Chaturvedi, M. (2021). Cancer scenario in North-East India & need for an appropriate research agenda. *The Indian journal of medical research*, 154(1), 27.
- [9] Mehrotra, R., & Yadav, K. (2022). Breast cancer in India: Present scenario and the challenges ahead. *World Journal of Clinical Oncology*, 13(3), 209.
- [10] Sheikh, A., Hussain, S. A., Ghorri, Q., Naeem, N., Fazil, A., Giri, S., ... & Al Tamimi, D. M. (2015). The spectrum of genetic mutations in breast cancer. *Asian Pac J Cancer Prev*, 16(6), 2177-2185.
- [11] Taplin, S. H., Ichikawa, L., Yood, M. U., Manos, M. M., Geiger, A. M., Weinmann, S., ... & Barlow, W. E. (2004). Reason for late-stage breast cancer: absence of screening or detection, or breakdown in follow-up?. *Journal of the National Cancer Institute*, 96(20), 1518-1527.
- [12] Holmes, M. D., Hunter, D. J., Colditz, G. A., Stampfer, M. J., Hankinson, S. E., Speizer, F. E., ... & Willett, W. C. (1999). Association of dietary intake of fat and fatty acids with risk of breast cancer. *Jama*, 281(10), 914-920.
- [13] Bhattacharjee S, Deka S.(2022), Analysis of breast cancer patients of NE India during 2016-2019. *JP J. Biostat.* 2022;19:1-3.
- [14] Dudley, W. N., Wickham, R., & Coombs, N. (2016). An introduction to survival statistics: Kaplan-Meier analysis. *Journal of the advanced practitioner in oncology*, 7(1), 91.
- [15] Ewertz, M., Gillanders, S., Meyer, L., & Zedeler, K. (1991). Survival of breast cancer patients in relation to factors which affect the risk of developing breast cancer. *International journal of cancer*, 49(4), 526-530.
- [16] Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American*

statistical association, 53(282), 457-481.

- [17] Siegel, R. L., Miller, K. D., & Jemal, A. (2020). Cancer statistics, 2020. *CA: a cancer journal for clinicians*, 70(1), 7-30.
- [18] Clark, G. M., & McGuire, W. L. (1988, April). Steroid receptors and other prognostic factors in primary breast cancer. In *Seminars in oncology* (Vol. 15, No. 2 Suppl 1, pp. 20-25).
- [19] Agarwal, G., & Ramakant, P. (2008). Breast cancer care in India: the current scenario and the challenges for the future. *Breast care*, 3(1), 21-27.
- [20] Madhav, M. R., Nayagam, S. G., Biyani, K., Pandey, V., Kamal, D. G., Sabarimurugan, S., ... & Jayaraj, R. (2018). Epidemiologic analysis of breast cancer incidence, prevalence, and mortality in India: Protocol for a systematic review and meta-analyses. *Medicine*, 97(52), e13680.
- [21] Khokhar, A. (2012). Breast cancer in India: where do we stand and where do we go. *Asian Pac J Cancer Prev*, 13(10), 4861-4866.
- [22] Rajbongshi, N., Nath, D. C., & Mahanta, L. B. (2018). Estimating risk of breast cancer occurrences at different ages: application of survival techniques. *Asian Pacific journal of cancer prevention: APJCP*, 19(11), 3033.
- [23] Giaquinto, A. N., Sung, H., Miller, K. D., Kramer, J. L., Newman, L. A., Minihan, A., ... & Siegel, R. L. (2022). Breast cancer statistics, 2022. *CA: a cancer journal for clinicians*, 72(6), 524-541.
- [24] Gogoi, G., Borgohain, M., Saikia, P., & Fazal, S. (2016). Profile of molecular subtypes of breast cancer with special reference to triple negative: A study from Northeast India. *Clinical Cancer Investigation Journal*, 5(5-2016), 374-383.
- [25] Leong, S. P., Shen, Z. Z., Liu, T. J., Agarwal, G., Tajima, T., Paik, N. S., ... & Foulkes, W. D. (2010). Is breast cancer the same disease in Asian and Western countries?. *World journal of surgery*, 34(10), 2308-2324.
- [26] Sharma, D. (2014). Triple negative breast cancer in people of North East India: critical insights gained at a regional cancer centre. *Asian Pacific Journal of Cancer Prevention*.
- [27] Brandt, J., Garne, J. P., Tengrup, I., & Manjer, J. (2015). Age at diagnosis in relation to survival following breast cancer: a cohort study. *World journal of surgical oncology*, 13(1), 33.
- [28] Mehrotra, R., & Yadav, K. (2022). Breast cancer in India: Present scenario and the challenges ahead. *World Journal of Clinical Oncology*, 13(3), 209.
- [29] Gajalakshmi, C. K., Shanta, V., Swaminathan, R., Sankaranarayanan, R., & Black, R. J. (1997). A population-based survival study on female breast cancer in Madras, India. *British journal of cancer*, 75(5), 771-775.
- [30] Macdonald, S., Oncology, R., & General, M. (2022). Breast cancer facts & figures 2022–2024. *Am Cancer Soc*, 10, 201006001.
- [31] Bradley, E. M., Woodhouse, P., Moses, K. A., Salwei, M. E., Sexton, K. W., Kripalani, S., ... & Idrees, K. (2026). The Role of Health Literacy in Breast Cancer Treatment and Outcomes. *Journal of the National Comprehensive Cancer Network*, 24(5), 179-185.
- [32] Singh, P., Kapil, U., Shukla, N. K., Deo, S. V. S., & Dwivedi, S. N. (2011). Association of overweight and obesity with breast cancer in India. *Indian journal of community medicine: official publication of Indian Association of Preventive & Social Medicine*, 36(4), 259.
- [33] Blair, C. K., Wiggins, C. L., Nibbe, A. M., Storlie, C. B., Prossnitz, E. R., Royce, M., ... & Hill, D. A. (2019). Obesity and survival among a cohort of breast cancer patients is partially mediated by tumor characteristics. *NPJ breast cancer*, 5(1), 33.
- [34] Nahmias-Blank, D., Maimon, O., Meirovitz, A., Sheva, K., Peretz-Yablonski, T., & Elkin, M. (2023, November). Excess body weight and postmenopausal breast cancer: emerging molecular mechanisms and perspectives. In *Seminars in Cancer Biology* (Vol. 96, pp. 26-35). Academic Press.
- [35] Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, et. al., (eds). SEER Cancer Statistics Review, 1975-

2017, National Cancer Institute. Bethesda, MD based on November 2019 SEER data submission, posted to the SEER web site, April 2020. (https://seer.cancer.gov/csr/1975_2017/ accessed on 16-11-2024).

- [36] Dhillon, P. K., Mathur, P., Nandakumar, A., Fitzmaurice, C., Kumar, G. A., Mehrotra, R., ... & Dandona, L. (2018). The burden of cancers and their variations across the states of India: the Global Burden of Disease Study 1990–2016. *The Lancet Oncology*, 19(10), 1289-1306.
- [37] Mehrotra, R., & Yadav, K. (2022). Breast cancer in India: Present scenario and the challenges ahead. *World Journal of Clinical Oncology*, 13(3), 209.
- [38] Thomas, N. S., & Norris, E. (2025). Understanding the delays in diagnosis of breast cancer in India using blogs and YouTube videos. *BMC Public Health*, 25(1), 2792.
- [39] Thomas, N. S., & Norris, E. (2025). Understanding the delays in diagnosis of breast cancer in India using blogs and YouTube videos. *BMC Public Health*, 25(1), 2792.
- [40] Srivastava, T. P., Goel, I., Gogia, A., Parshad, R., Monga, S., Talukdar, J., ... & Karmakar, S. (2025). Comprehensive overview of breast cancer in India. *JCO Global Oncology*, 11, e2500083.
- [41] Nair, M. K., Sankaranarayanan, R., Nair, K. S., Amma, N. S., Varghese, C., Padmakumari, G., & Cherian, T. (1993). Overall survival from breast cancer in Kerala, India, in relation to menstrual, reproductive, and clinical factors. *Cancer*, 71(5), 1791-1796.
- [42] Feleke, B., Tesfaw, L. M., & Mitku, A. A. (2022). Survival analysis of women breast cancer patients in Northwest Amhara, Ethiopia. *Frontiers in Oncology*, 12, 1041245.
- [43] Pakseresht, S., Ingle, G. K., Bahadur, A. K., Ramteke, V. K., Singh, M. M., Garg, S., & Agarwal, P. N. (2009). Risk factors with breast cancer among women in Delhi. *Indian journal of cancer*, 46(2), 132-138.
- [44] Brewer, H. R., Jones, M. E., Schoemaker, M. J., Ashworth, A., & Swerdlow, A. J. (2017). Family history and risk of breast cancer: an analysis accounting for family structure. *Breast cancer research and treatment*, 165(1), 193-200.
- [45] Chang, E. T., Milne, R. L., Phillips, K. A., Figueiredo, J. C., Sangaramoorthy, M., Keegan, T. H., ... & John, E. M. (2009). Family history of breast cancer and all-cause mortality after breast cancer diagnosis in the Breast Cancer Family Registry. *Breast cancer research and treatment*, 117(1), 167-176.
- [46] Liu, L., Hao, X., Song, Z., Zhi, X., Zhang, S., & Zhang, J. (2021). Correlation between family history and characteristics of breast cancer. *Scientific reports*, 11(1), 6360.
- [47] Kumar, A., Bhagabaty, S. M., Tripathy, J. P., Selvaraj, K., Purkayastha, J., & Singh, R. (2019). Delays in diagnosis and treatment of breast cancer and the pathways of care: a mixed methods study from a tertiary cancer centre in North East India. *Asian Pacific journal of cancer prevention: APJCP*, 20(12), 3711.
- [48] Sharma, D. K., & Saripilli, R. (2025). Recent strategies in diagnosis, screening, prevention, and treatment of breast cancer in young women. *Discover Oncology*, 16(1), 1532.
- [49] Early Breast Cancer Trialists' Collaborative Group. (2011). Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10 801 women in 17 randomised trials. *The Lancet*, 378(9804), 1707-1716.

Tables and Figures Legends:

Table 1: Variable Description of n=210 breast cancer Participants

Table 2: Five Year Overall Survival rates of n=210 BC patients

Table 3: Association of Risk factors with the occurrence of breast cancer

Table 4: Mean and Median Survival time (Weeks) of breast Cancer Patients by socio-demographic and treatment related risk factors

Table 5: Impact of Various Factors on the Survival Time.

Table 6.1: Stage wise comparison of estimated Survival results with National Figure.

Table 6.2: Stage wise comparison of estimated survival results with International Figure

Figure 1: Kaplan- Meier curve for Overall 5- Year Survival

Figure 2: Kaplan-Meier curve for survival function of breast cancer patients in association with socio-demographic and other studies factors: (a) Education status (b) Age (c) BMI (d) Breastfeeding (e) Stage of breast cancer at diagnosis (f) Family history of cancer (g) Treatment

Table 1: Variable Description of n=210 breast cancer Participants

Variable	Category(Statistic)	N(%)or Mean $\pm$ SD
Age(Years)	Below 45	46.27 $\pm$ 29.37
	45 and Above	
Education	Illiterate	42(20.0%)
	Primary	92(43.8%)
	Secondary	57(27.2%)
	Graduate	19(9.05%)
Income	Lower	189(90.0%)
	Middle	16(7.6%)
	Higher	5(2.4%)
BMI	Underweight	14(6.7%)
	Normal	148(70.4%)
	Overweight	35(16.7%)
	Obese	13(6.2%)
Breastfeeding	Yes	172(81.9%)
	No	38(18.1%)
Cancer Stages	First (I)	47(22.4%)
	Second (II)	48(22.8%)
	Third (III)	59(28.1%)
	Fourth (IV)	47(26.7%)
Family History	Yes	19(9.1%)
	No	191(90.9%)
Breast Cancer treatments	Only Surgery	63(30.0%)
	Only Chemotherapy	101(56.2%)
	Chemotherapy with Surgery& other	29(13.8%)
	No Treatment	17(8.09%)

Table 2. Five Year Overall Survival rates of n=210 BC patients

Factor	Sub groups	Survival rate (%)		
		1 year	3 year	5 year
Education	Illiterate	81	50	20
	Primary	92	58	-
	Secondary	90	62	40
	Graduate	99	70	50
Age	Below 45	96	90	38
	45 & Above	88	48	15
BMI	Underweight	90	71	-
	Normal	90	65	21
	Overweight	90	41	20
	Obese	69	40	5
Cancer Stage	Stage I	96	75	75
	Stage II	95	70	60
	Stage III	90	60	22
	Stage IV	89	31	5
Breastfeeding	Women with children (Yes)	93	59	21
	Women with no Children (No)	81	40	-
Family History	Yes	68	29	-
	No	90	55	21
Breast Cancer Treatments	No Treatment	60	30	-
	Only Surgery	90	70	15
	Only Chemotherapy	92	90	40
	Chemotherapy with Surgery and other treatments	95	70	42
Overall Survival		92	68	40

Table 3: Association of Risk factors with the occurrence of breast cancer

Risk Factors	Status	Mean Survival Time(MST) in Weeks	SD@ of MST	Median Survival Time in Weeks	Test Statistic	p-value
1.Education Status	Primary	109.8	66.7	111.0	Kruskalwallis Test $U=2.221$	0.528
	Secondary	101.5	63.0	100.0		
	Graduate	121.1	63.1	127.0		
	Illiterate	115.2	67.2	114.0		
2. Age	Below 45	106.7	66.6	97.5	Mann-Whitney Test $U=4948.5$	0.733
	45 & above	109.7	64.7	111.0		
3.Income	Low	110.5	65.8	108.0	Mann-Whitney $U=10441.5$	0.028*
	High	120.0	23.07	114.0		
4.Cancer Stage	Stage I	105.4	63.3	96.0	Kruskalwallis Test= 4.903	0.179
	Stage II	105.0	60.2	100.0		
	Stage III	107.2	65.8	109.0		
	Stage IV	116.0	71.3	114.0		
5.Breastfeeding Status	Yes	107.9	64.8	104.0	Mann-Whitney $U=1673$	0.024*
	No	143.9	62.0	142.0		
6.Family History of cancer	Yes	95.3	66.4	84.0	Mann-Whitney $U=1963.5$	0.851
	No	109.9	65.2	108.0		
7. BMI	Underweight	76.1	60.0	54.2	Kruskalwallis Test= 0.931	0.818
	Normal	111.0	65.6	111.0		
	Overweight	109.4	61.9	104.0		

Table 4: Mean and Median Survival time (Weeks) of breast Cancer Patients by socio-demographic and treatment related risk factors

Variable	Category(Statistic)	Mean Survival Time(Weeks)[95%CI]	Median Survival(weeks)[95%CI]	Log-rank value	p-value
Age(Years)	Below 45	205.5[183.5,227.4]	238.0 [176.1, 299.8]	12.204	0.0005*
	45 and Above	155.8[140.6,171.0]	154.0 [143.9, 164.1]		
Education	Illiterate	132.7[107.7,157.7]	147.0 [86.0, 207.9]	12.262	0.007*
	Primary	188.6[166.5,210.6]	206.0[174.1, 237.8]		
	Secondary	170.8[141.3,194.4]	167.0[113.7, 220.2]		
	Graduate	189.0[154.4,223.4]	221.0 [190.3, 257.8]		
Income	Lower	173.4[158.7,188.0]	177.0 [144.9, 209.0]	0.577	0.757
	Middle	159.0[125.2,192.8]	162.0 [140.3, 183.6]		
	Higher	172.5[103.1,241.8]	150.0 [91.6, 208.3]		
BMI	Underweight	151.7[124.0,179.3]	- [-,-]	8.838	0.032*
	Normal	180.2[165.0,195.4]	195.0 [156.6, 233.4]		
	Overweight	146.9[116.8,177.0]	143.0 [98.3, 187.6]		
	Obese	126.8[82.4,171.2]	139.0 [69.5, 208.4]		
Breastfeeding	Yes	182.7[167.2,198.2]	195.0 [152.3, 237.7]	9.931	0.002*
	No	132.3[109.5,155.0]	140.0 [88.7, 191.2]		
Cancer Stages	First (I)	204.4[175.9,232.7]	212.0 [-,-]	16.352	0.001*
	Second (II)	203.8[176.7,230.8]	- [-,-]		
	Third (III)	170.3[147.7,193.0]	171.0 [143.7, 198.3]		
	Fourth (IV)	139.4[118.2,160.4]	140.0 [99.2, 180.8]		
Family History	Yes	118.8[86.6,151.0]	100.0 [28.2, 171.7]	9.186	0.002*
	No	177.2[163.1,191.4]	180.0 [145.0, 214.6]		
Breast Cancer treatments	Only Surgery	211.0[187.7,234.2]	238.0 [181.5, 294.5]	22.69	0.00005*
	Only Chemotherapy	163.0[145.8,180.2]	159.0 [145.3, 172.6]		
	Chemotherapy with Surgery and other	157.8[124.7,190.9]	157.0[129.2, 184.8]		
	No Treatment	102.2[69.1,135.2]	97.0 [20.2, 173.8]		
Overall Survival		171.1[157.8,184.5]	167.0 [139.0, 195.0]		

Table 5: Impact of Various Factors on the Survival Time.

Factors	B	$SE\ of\ \beta$	E^β	$CI\ of\ e^\beta$	Z	$P(Z > z)$
Education status(Graduate)	<i>Reference</i>					
Education status(Primary)	-0.212	0.617	0.080	[0.241, 2.711]	-0.345	0.730
Education status(Secondary)	0.065	0.300	1.067	[0.592,1.922]	0.217	0.528
Education status(Illiterate)	0.521	0.245	1.684	[1.042, 2.712]	2.121	0.034*
Age (Below 45)	<i>Reference</i>					
Age(45& Above)	0.856	0.343	2.354	[1.201, 4.613]	2.495	0.012*
BMI(Underweight)	<i>Reference</i>					
BMI(Normal)	-0.126	0.759	0.882	[0.199,3.898]	-0.167	0.869
BMI(Overweight)	0.336	0.780	1.399	[0.303,6.459]	0.431	0.667
BMI(Obese)	0.708	0.839	2.029	[0.391, 10.513]	0.844	0.398
Cancer stage(I)	<i>Reference</i>					
Cancer stage(II)	0.485	0.505	1.625	[0.604,4.369]	0.962	0.335
Cancer stage(III)	0.732	0.453	2.080	[0.856, 5.057]	1.617	0.105
Cancer stage(IV)	0.862	0.430	2.367	[1.017, 5.502]	2.000	0.045*
Breastfeeding(No)	<i>Reference</i>					
Breastfeeding (Yes)	0.612	0.306	1.843	[1.011, 3.361]	1.997	0.045*
Family History(No)	<i>Reference</i>					
Family History(Yes)	-0.334	0.379	0.719	[0.341, 1.506]	-0.879	0.379
Treatment(Surgery)	<i>Reference</i>					
Treatment(Chemotherapy)	0.695	0.364	2.00	[0.980, 4.096]	1.907	0.056
Treatment(Chemotherapy with Surgery and others)	0.749	0.487	2.11	[0.812,5.501]	1.534	0.125
Treatment (No treatment)	1.623	0.519	5.06	[1.831,14.015]	3.126	0.001*
Likelihood ratio test= 51.72 on 15 df, p=6e-06 Wald test = 49.45 on 15 df, p=1e-05 Score (log rank) test = 56.57 on 15 df, p=1e-06						

*= Significant at 95% CI

Table 6.1: Stage wise comparison of estimated Survival results with National Figure.

Factor	Median Survival Time		p-value	
	Estimated from present study (Weeks)	National Reference (weeks)	Estimated from present study	National Reference
Stage I	212	-	0.000	<0.01
Stage II	-	-		
Stage III	171	130.36		
Stage IV	140	78.27		

Table 6.2: Stage wise comparison of estimated survival results with International Figure.

Factor	Median survival time		Hazard ratio		p-value	
	Estimated from present study (Weeks)	International Reference (Months)	Estimated from present study	International Reference	Estimated from present study	International Reference
Stage I	212	-	-	-	-	-
Stage II	-	-	1.412	3.973	0.491	0.195
Stage III	171	156.5	1.646	14.997	0.256	0.008
Stage IV	140	147.8	2.270	11.936	0.036	0.016

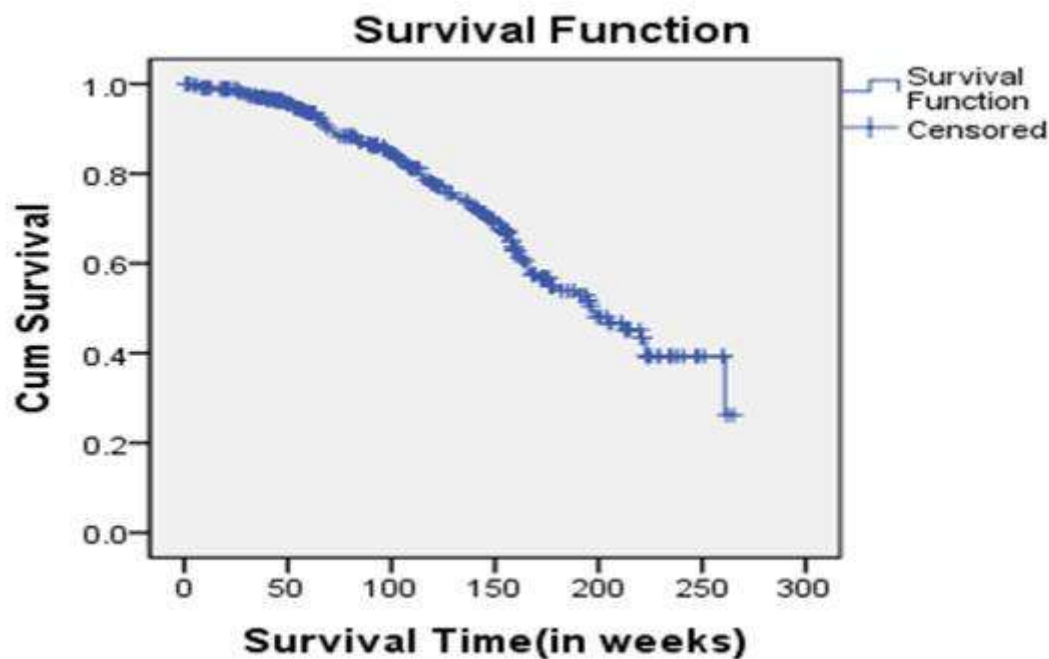


Figure 1: Kaplan-Meier curve for overall 5-Year Survival.

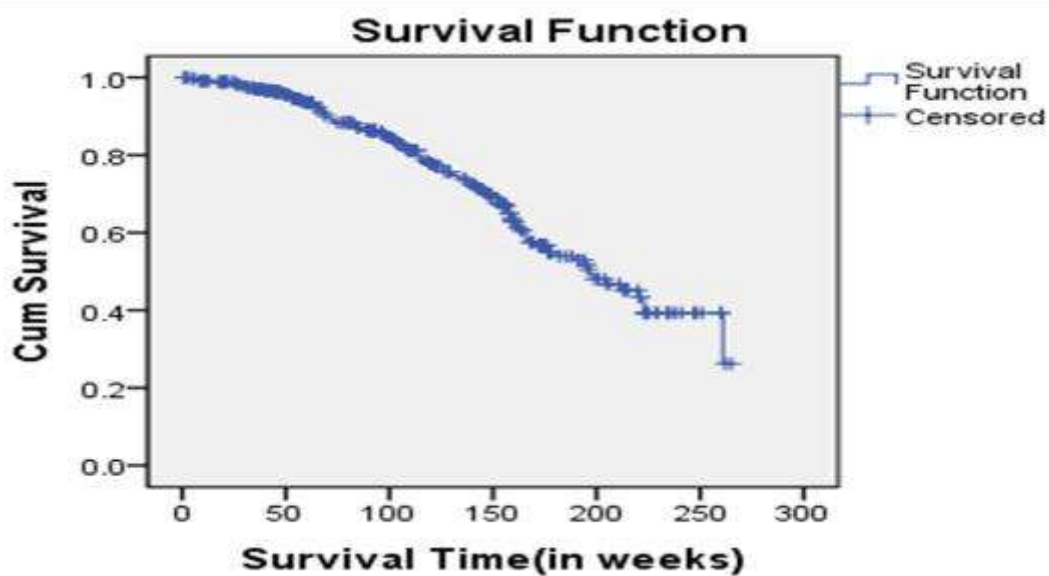


Figure 2: Kaplan-Meier curve for survival function of breast cancer patients in association with socio-demographic and other studies factors: (a) Education status (b) Age (c) BMI (d) Breastfeeding (e) Stage of breast cancer at diagnosis (f) Family history of cancer (g) Treatment

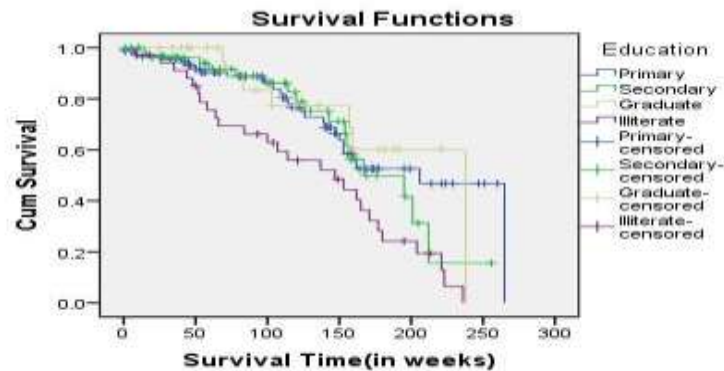


Figure 2(a)



Figure 2(b)

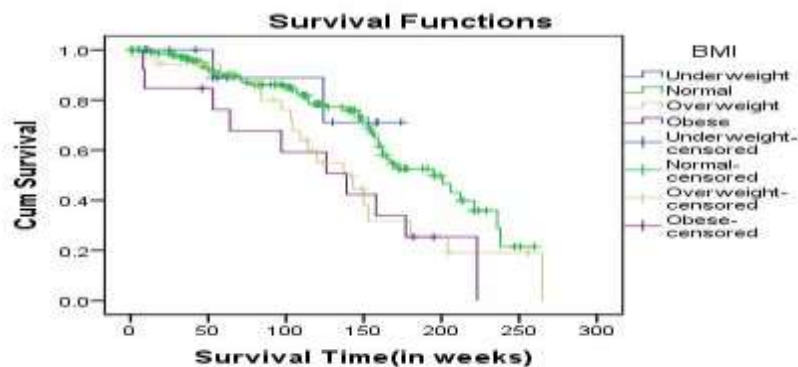


Figure 2(c)

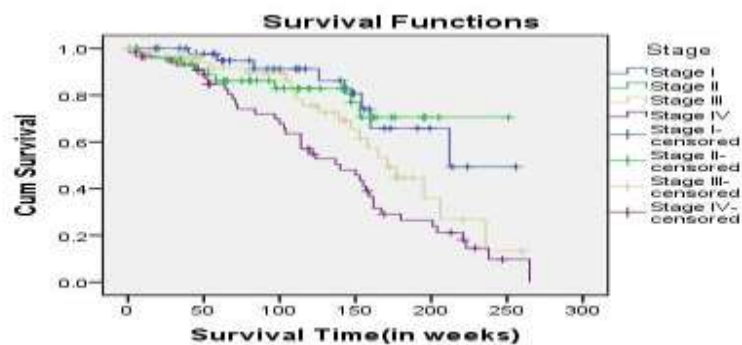


Figure 2(d)

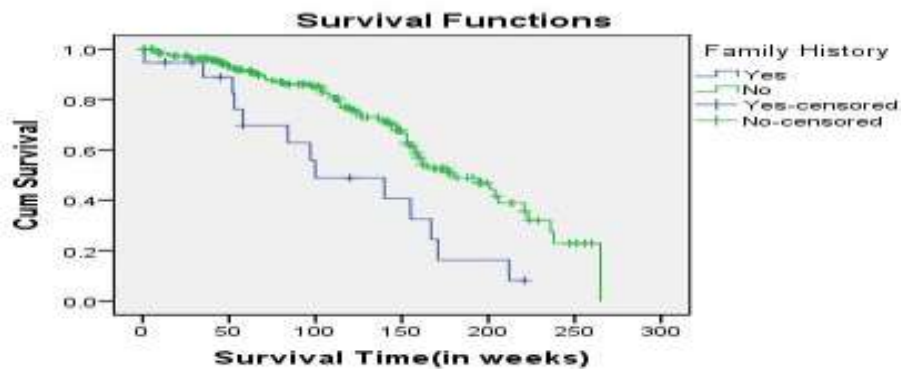


Figure 2(f)

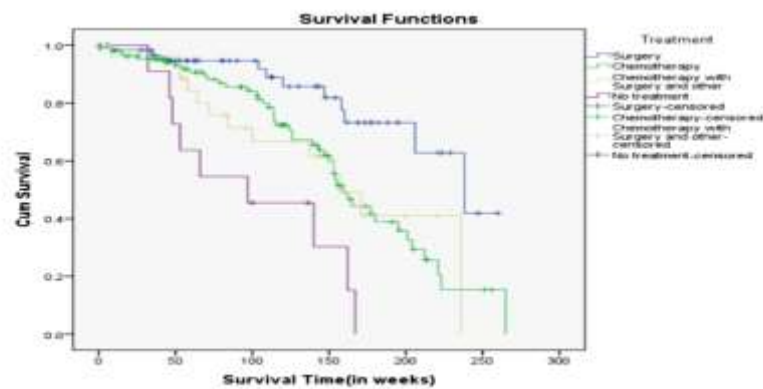


Figure 2(g)

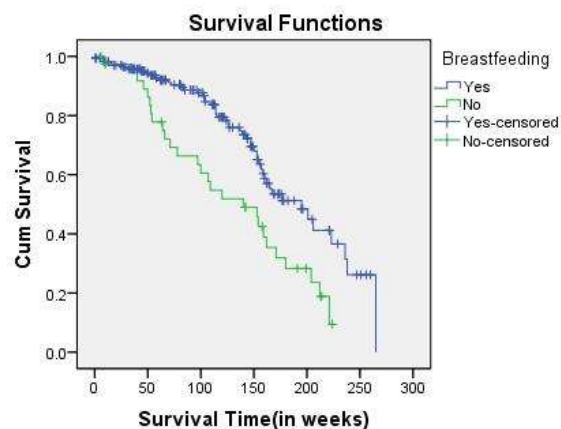


Figure 2(e)