

STUDY ON SURVIVAL OF BREAST CANCER PATIENTS IN ASSOCIATION WITH SOCIO-DEMOGRAPHIC FACTORS OF DIBRUGARH DISTRICT OF ASSAM, INDIA

Bornali Khound¹, Dhruba Das², Gayatri Gogoi^{*3}, Dibyojyoti Bora⁴, Dharmendra Dubey⁵

¹Research Scholar, Department of Statistics, Dibrugarh University.

²Assistant Professor, Department of Statistics, Dibrugarh University.

³Associate Professor, Department of Pathology, Assam Medical College and Hospital, India.

⁴Assistant Professor, Department of Statistics, Cotton University, India.

⁵Assistant professor of Biostatistics, Department of community medicine, Baba kinaram autonomous State medical college chandauli, Uttarpradesh, India.

***Corresponding Author: Gayatri Gogoi^{*3},**

Associate Professor, Department of Pathology, Assam Medical College and Hospital, India

Email: gayatrigogoi303@gmail.com

KEYWORDS

Breast cancer, Dibrugarh district, socio-demographic factors, survival outcome.

ABSTRACT

Background: The survival rate for breast cancer patients in India remains poor compared to Western countries, primarily due to early onset, late diagnosis and delayed treatment. In Dibrugarh district of Assam, India, breast cancer is more prevalent among younger women, with 64.9% of cases, affecting women below 50 years of age according to ICMR cancer registry. Furthermore, very limited data from this region hinders a comprehensive understanding of the disease's burden and trends in this region.

Objectives: To investigate the survival rates of breast cancer patients in Dibrugarh, Assam India and determine the socio-demographic, lifestyle and other factors that influence survival, with a view to identifying key predictors and covariates that associates breast cancer survival in this region.

Materials and Method: A retrospective cohort study was conducted from 2017 to 2022 in a tertiary care centre in Dibrugarh, Assam for 210 breast cancer cases. The data were analysed using Kaplan Meier survival plot(s) for survival rates, Log-Rank test to compare the survival times for different levels of the risk factors and impact of risk factors on the time to death due to breast cancer was analysed using Cox proportional hazard model.

Results: The average age at diagnosis of breast cancer was 45.5 years. The Overall 5-year survival rate of breast cancer patients was 38%. The study signifies the socio-demographic, lifestyle and other factors influencing survival viz. educational status, onset age of breast cancer, BMI, cancer stage, breastfeeding and family history. The survival rates vary significantly with the age at diagnosis, early-stage detection showing a substantially better survival rate compared to late-stage diagnosis. The Cox's proportional Hazard model inferred an overall significant relationship between these set of covariates and survival time.

Discussion and conclusion: The study identified socio-demographic and other factors such as educational status, age at diagnosis, BMI, cancer stage, breastfeeding and family history that influence survival rates. These findings highlight the need for increased awareness and early detection of breast cancer in

India. Regular breast self-examination, clinical breast examination, and mammography can help detect breast cancer at an early stage, improving treatment outcomes and survival rates. It is essential to address the socio-demographic and lifestyle factors that influence breast cancer survival patterns in India.

Introduction:

Breast cancer (BC) is the leading cancer in women worldwide. Every year, 2.3 million breast cancer cases are registered and 6,85,000 dies from breast cancer.¹ Breast cancer accounted for 15.5% of all cancer-related deaths among women worldwide in 2020 and continues to be the most commonly diagnosed form of cancer in females (23.5%). It has now surpassed lung cancer as the leading cause of global cancer incidence in 2020 with 11.7% of all cancer cases. As per the GLOBOCON data (2020) breast cancer accounted for 13.5% (1,78,361) of all cancer cases and 10.6% (90,408) of all deaths in India. In eastern part of India, breast cancer is one of the common non-communicable disease and the most commonly diagnosed cancer in women.² 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.³

The studies from India have reported that 5-year overall survival rate of breast cancer is 95% for stage I patients, 92% for stage II, 70% for stage III, only 21% for stage IV patients.⁴ Previous studies around the world have shown that lower-middle income countries especially in rural populations, the impact of cancer is high, due to low awareness, lack of access to affordable care and poor prognosis. Also the survival rate of patients with breast cancer is poor in India as compared to western countries due to early age at onset, late stage diagnosis and delayed in treatments.⁵ In Western countries, 12 women develop breast cancer at some point in life and estimated that 5% - 10% of the breast cancer cases in women are linked to hereditary factor. Age at diagnosis and stage of breast cancer is associated with survival of breast cancer patients.⁶⁻⁷ Dietary factor has a significant effect on women breast cancer.⁸ In the United States, the effect of some risk and prognostic factors contributing to the survival of the breast cancer patients showed that tumour size, lymph node metastasis and tumour extension have significant effect on breast cancer survival.⁹ A NE Indian study on survival analysis of breast cancer patients revealed that overall survival rate of breast cancer was 93.3% (1 year) and 70.8% (3 year) with average survival time is 38 months. The risk of breast cancer occurrences among women belonging to 35 – 50 years is higher than young and elderly aged women. Non-parametric and semi-parametric study on survival of breast cancer patients shows that the risk factors age, educational status, residence etc. are significantly associated with risk of mortality of breast cancer patients.¹⁰⁻¹¹

The NE region has the highest incidence of cancer in India burdened by issues not only inadequate cancer treatment facilities but lack of awareness for early detection and treatment. Moreover, factors like socio-demographic, reproductive and genetic profile along with lifestyle are strongly associated with breast cancer survival. In the present study, educational status, onset age of diagnosis, Body Mass Index (BMI), stage of cancer and breastfeeding are the risk factors discussed for breast cancer patients of Dibrugarh District of Assam, India that are strongly associated with the survival of breast cancer patients. The study also highlights some clinical factors like age at first delivery, delivery mode, BMI and cancer stage along with genetic factor like family history of cancer. These factors influence the risk of breast cancer and indicate an association between breast cancer and patient's survival.¹²

Survival analysis is a statistical method used to examine the time-to-event data, where the event of interest is death.¹³ In the context of breast cancer, survival analysis helps researchers and clinicians understand the probability of survival for patients suffering from the disease.¹⁴ The primary goal of survival analysis is to identify the factors that influence the time-

to-death, enabling healthcare professionals to develop effective treatment strategies and improve patient outcomes.¹⁵ Additionally, predictive models can be developed to forecast individual patient outcomes, enabling personalized medicine approaches.¹⁶ Different survival methods like Kaplan-Meier method of Log Rank test and Cox Proportional Hazard regression are commonly used to analyze survival function.

Objectives of the study:

Even though various studies have been undertaken in India and worldwide on breast cancer survival but in regional level, research has been limited. The present regional level study was conducted in a tertiary care of Dibrugarh, Assam, India with the following objectives:

- 1) To investigate the survival rates of breast cancer patients with reference to socio-demographic, reproductive, genetic and lifestyle factors.
- 2) To compare survival time among the different sub-groups of breast cancer patients.
- 3) To determine the factors and/or covariates that affects the survival function.

Materials and methods:

i. Study design

This was a retrospective cohort study conducted at a tertiary care center of Upper Assam. The period of study was from 2017-2022. The study was carried out in a tertiary care centre of Upper Assam region. The patients were considered as censored if they alive beyond 2022 died due to other cause or loss to follow up.

ii. Inclusion and exclusion criteria

Inclusion criteria:

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

Exclusion criteria:

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

iii. Enrolment

Participants 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.

iv. Sample size

A total of 210 adult female individuals who fulfilled the inclusion criteria participated in the present study. It was a hospital-based study and purposive sampling technique was used for determination of sample size meeting with the feasibility to conduct the study.

v. Data collection and follow up

Data had been retrospectively collected from 210 breast cancer participants through telephonic conversation; while some participants visited the tertiary care centre for physical follow-up to ensure survival. A total of 14 variables were assessed which included socio-demographic variables (educational status, employment, income, marital and religion statuses), clinical variables (mode of delivery, breast feeding status, cancer breast, stage of cancer and family history) and status of substance consumption (alcohol, smoking and use of tobacco)

vi. Survival analysis

In the present study, survival time was defined as time to death due to breast cancer. From the right-censored observations, the Kaplan-Meier test was used to estimate the survivorship function among the study population at 1, 3 and 5 years duration.

vii. Association of the factors with the occurrence of breast cancer

The variable measuring the times 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 - walis test are used.

vii. Comparison of Survival Times

Log-Rank test (Mental-Heinzen test) was used to compare survival times for different risk factors.

viii. Analysis of combined effect of covariates on survival

Examination of combined effect of the covariates on the time to death due to breast cancer was done using Cox's Proportional Hazard Model. The model simultaneously evaluates the effect of several covariates on survival.

ix. Ethical approval

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.

x. Statistical analysis

The statistical tests performed in the present study were: Mann-whitney *U* test/Kruskal walis test, Kaplan-Meier survival plots, Log-Rank test (Mental-Heinzen test) and Cox's Proportional Hazard Model. P-values less than or equal to 0.05 was considered as statistical significance. Statistical analysis was performed with Microsoft Excel, IBM SPSS software (version 21) and R (version 4.3.3).

Results:

The present 5-year retrospective cohort study (2017-2022) analyzed 210 breast cancer cases, revealing an average age at diagnosis of 44.5 years, with a median age of 45 years. The majority of Breast cancer (BC) patients were Hindus (94.2%), literate (80.0%), and housewives (98.0%). Most BC patients had a normal BMI (70.4%) with an average weight of 52.5 kg and average BMI of 22.20. The study showed a minimal alcohol consumption (5.3%) and 9.5% occasional tobacco use. Family history of cancer was very rare (9.1%). Breast cancer stages were distributed as follows: Stage I (22.4%), Stage II (22.8%), Stage III (28.1%), and Stage IV (25.7%), with BC occurring equally on both breasts (Left: 53.4%, Right: 46.2%). Additionally, 81.9% of participant breastfeed and 96.2% were married. Detailed socio-demographic and clinical characteristics along with status of substance consumption are shown in Table1.

Survival outcomes:

We analyzed the survival outcomes of 210 cases of Breast Cancer (BC) patients. 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).

The survival outcomes were also associated with educational status, age at diagnosis, BMI, stage of cancer and breast-feeding status and family history. Education level emerged as a critical predictor, with graduate individuals showing 99% in 1-year, 70% in 3-year, and 50% in 5-year survival rates, contrasting with illiterate individuals' rates of 81%, 50%, and 20%, respectively (Figure2(a)). Age also impacted survival, with the Below 45 age group exhibiting 95% 1-year, 89% 3-year and 40% 5-year survival rates, whereas the 45 and above age group had 89% 1-year, 45% 3-year and 15% 5-year survival rates (Figure2(b)). Additionally, Underweight or normal individuals had higher survival rates compared to obese individuals (70% 1-year survival, 41% 3-year survival and 5% 5-year survival (Figure 2(c)). Cancer stage at diagnosis was strongly correlated with survival, ranging from 95% 1-year and 75% 5-year survival rates for Stage I to 90% 1-year and 5% 5-year rates for Stage IV (Figure2(d)). Furthermore, women with a breastfeeding history demonstrated better survival rates (92% at 1 year and 58% at 3 years) compared to those without (80% and 41%)(Figure2(e)). Lastly, women with family history of cancer have lower survival rates as compared to without family

history (70% 3year survival and 30% 3 year survival)(Figure2(f)). The detailed survival rate of these factors is shown in Table 2.

The results of Mann-whitney U test, Kruskalwalli's test and p values shows that the breast cancer patients with three factors , namely Caste, Employment Status and Breastfeeding Status are positively associated with the occurrences of breast cancer. In case of Caste (Mann-Whitney U=1534.0, p=0.048) there is a significant difference between Reserved and Unreserved category as p value is less than 0.05. Similarly for breast cancer patients with Employment status (Mann-whitney U=10441.5, p=0.028) has significant difference between Professional and Housewife women 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 results of the test is given in Table 3

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 covariates, and different groups of Education, Age at diagnosis, BMI, Cancer Stage, Breastfeeding and family history had a statistically difference in 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, $p=0.0002$). Additionally, middle or older onset age patients (45& years) had shorter survival times than younger patients (Log-Rank test value: 12.2, $p=0.0004$). 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, family history with cancer has also significant association with survival time (Log-Rank test value: 7.5, $p=0.006$). Detailed survival analysis in 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 six significant factors viz. educational status, age at diagnosis, BMI, cancer stage at diagnosis, breastfeeding status and family history 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 these set of covariates and survival time.

The results showed illiterate patients had 68.4% increased hazard of death compared to literate patients ($p=0.034$). Middle or older age patients (45 & above) had a 130.9% increased hazard of death compared to younger patients ($p=0.009$). Patients diagnosed at Stage IV had a 2.210 times increased hazard of death ($p=0.031$) compared to early stage. No Breastfeeding status had 73.1% increased hazard of death compared to participants with positive breastfeeding status ($p=0.046$). Although BMI and family history with cancer were insignificant as their hazard ratios indicated increased risk. Table 5 gives detailed outcomes of proportional hazard model with p -values ($Pr(Z>|z|)$) along with the coefficients (β) and its standard errors (SE), exponentiation of coefficients (e^{β}) and its confidence intervals (CI) and critical values (z).

Discussion:

Breast cancer is one of the most common malignancies affecting women globally and a leading cause of cancer-related deaths among women. In India, breast cancer incidence has been steadily rising, with a particularly noticeable trend of diagnosis at younger ages compared to Western populations.¹⁷⁻¹⁸ In NE (NE) India, breast is the leading site of cancer (14.5%) among the females.¹⁹ 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, lifestyle, socio-demographic factors, and cancer staging. These findings point toward regional variations in breast cancer trends from both Indian and Global patterns.

Survival Outcomes and Influencing Factors in Breast Cancer

Age at Diagnosis and survival:

This study revealed that the average age of breast cancer diagnosis is 44.5 years, significantly younger than in Western countries like the USA, where incidence typically peaks around 60 years.²⁰ These results are consistent with previous studies from NE India, which reported average ages at diagnosis of 44.6 and 44.5 years respectively, with peak onset in the 30-60 age range.²²⁻²³ Genetic, reproductive, dietary and lifestyle factors specific to Indian populations may contribute to this early onset.²⁴⁻²⁵ Notably, the younger average age in NE India highlights the importance of early screening initiatives.

Furthermore, our study demonstrated a strong correlation between age at diagnosis and survival rates. Younger patients (Below 45) showed a 95% 1-year and 89% 3-year survival rate, whereas middle or older patients (45 and above) had reduced survival rates (89% at 1 year and 45% at 3 years) (Table 2). The result also showed that median survival time of Younger patients Below 45 is 238 week, which is much higher than median survival time 154 week of middle or older patients(45 and above years) (Table 4). This indicated that the risk of death increases with the increasing age. And also the risk of hazard for middle or older patients is 2.03 times higher than younger age patients. (Table 5). This result was consistent with the study that increased age at diagnosis was associated with decrease survival.²⁶ This trend was attributed to younger individuals generally experiencing fewer co-morbidities and greater tolerance to aggressive treatments.²⁷ These findings suggest that age-appropriate screening and management strategies are crucial to optimize outcomes for middle or older age patients.

Socio-demographic patterns and survival:

The socio-demographic analysis of the study revealed 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 predominantly assume household roles, often engaging in informal labour, especially in rural and semi-urban areas.²⁸ Notably, literacy among women in this region was relatively high compared to other parts of India, potentially attributed to a strong emphasis on female education in states like Assam and Mizoram.²⁹

The data highlights the significance of education in breast cancer survival rates. Graduates demonstrated higher survival rates at all intervals: 99% at 1 year, 70% at 3 years, and 50% at 5 years. In contrast, illiterate sub group had lower survival rates: 81% at 1 year, 50% at 3 years, and 20% at 5 years (Table 2). Study also signifies that Illiterate patients have shorter survival time compared to literate patients. The median survival time is 147 week for illiterate patients compared to 221 weeks for Graduate patients, indicates illiteracy potentially being associated with shorter survival.(Table 4). Also the illiterate patients time to death breast cancer decreased by a factor 68.4% compared to literate patients, if we hold the other factor constant.(Table 5). The median survival time is 147 week for illiterate patients compared to 221 weeks for Graduate patients, indicates illiteracy potentially being associated with shorter survival.(Table5). Also the illiterate patients time to death breast cancer decreased by a factor 68.4% compared to literate patients, if we hold the other factor constant.(Table 6). This

highlights the vital role of education in health outcomes, likely due to improved health literacy, timely healthcare utilization and adherence to treatment protocols.^{14,30} Higher literacy levels can facilitate earlier medical attention for breast abnormalities, contributing to better health awareness. These findings emphasized the importance of educational initiatives targeting breast cancer awareness and self-breast examinations particularly in populations with lower literacy rates. By addressing education-related disparities; healthcare access and screening behaviors can be improved ultimately enhancing breast cancer outcomes in NE India.

Lifestyle Factors and Breast Cancer Risk: BMI, Diet, and Substance abuse:

A significant proportion (70.4%) of study participants had a normal BMI, averaging 22.21 kg/m² (Table 1). This contrasts with global findings where breast cancer incidence is often higher among individuals with overweight or obesity, particularly in post-menopausal women.³¹ 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 is a concern given high tobacco consumption rates in certain groups in NE India.³²⁻³³

BMI and Survival Outcomes

BMI also correlated with survival outcomes, with Normal BMI patients showing higher survival rates (90% at 1 year, 70% at 3 years, and 20% at 5 years) compared to obese individuals (Table 2). The present study shows that there is 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 suggest that obesity is associated with poorer overall breast cancer survival. On the other hand Cox model showed no evidence of significant correlation between the patients survival time with BMI. These results were contradict with the studies conducted by Singh *et. al.*, (2011) and Blair *et. al.*, (2019).³⁴⁻³⁵ 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.³⁶ 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.

Clinical Characteristics and Cancer Staging:

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 national trend, where limited early screening and delayed healthcare access often resulted in late-stage diagnosis.³¹ 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.³⁷

Cancer Stage at Diagnosis and Survival:

Cancer stage at diagnosis remains critical 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 established evidence linking early-stage detection to improved survival rates in breast cancer (National Cancer Institute, 2020).³⁸ Our research also demonstrates that women's survival time after being diagnosed with breast cancer is significantly influence by her cancer stage. The average survival time for patients with Stage IV are 2.270 times shorter than those of reference group (Stage I), when all other parameter held constant(Table 5). This suggests that overall survival declines with increasing clinical stage. Similar finding indicated that late stage breast cancer patients are more likely to experience the event (death). On average

50% of BC cases in India presented at late stage (Stage III & stage IV). This condition is now the leading cause of breast cancer related deaths among Indian women.³⁹⁻⁴⁰ Stage wise comparison of estimated survival outcome with national and international studies is detailed in Table 6.1 & 6.2 respectively.⁴¹⁻⁴²

Breastfeeding History and Breast Cancer Risk:

A significant proportion (81.9%) of participants had breastfeed history (Table 1), which was associated with reduced breast cancer risk.⁴³ High breastfeeding rates particularly in NE India may contribute to potential protective effects that delay cancer onset.

Breastfeeding and Survival Outcomes:

Women with a breastfeeding history showed better survival outcomes (92% at 1 year and 58% at 3 years) than those who had not breastfeed (78% and 41%, a) (Table 2). Also, the patients with Breastfeeding history have more survival time (212 week) than those who had not Breastfeed (149 week). i.e., factor Breastfeeding is a significant predictive factor for survival time of Breast Cancer patients ($p=0.015$) (Table 4). On the other hand Cox model showed no evidence of significant correlation between the patients survival time with breastfeeding. This result contradicts with the findings of Pakseresht *et. al.*, (2009).⁴⁴ There is increasing risk of breastfeeding's protective effects which may be due to hormonal and cellular changes in breast tissue during lactation that lower tumor genesis likelihood.⁴³ This suggests breastfeeding may have post-diagnosis benefits, potentially influencing tumour biology or enhancing immune response.

Family History and Survival Outcomes:

Women without family history of cancer shows better survival outcomes (80% 1-year and 70% 3-year) than those with 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 who had family History (140 Weeks) (Table 4). The findings of the study also found that family history determines the survival of Breast Cancer patients. Patients who have breast cancer history in their family are more likely to the risk of death as compared to patients who have no breast cancer history in their family. Chang *et. al.*, (2008) conducted 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 all-cause 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.⁴⁵ Another study by Yusuf *et. al.*, (2024) found that age and family history had significant effects on the survival time ($p=0.003$) and risk of death was 3.244 (95% CI; 0.334,2.020). This suggests that a patient with family history of cancer is associated with survival outcome and there is significant relationship between family history of cancer and survival time.⁴⁶

Conclusion:

This study identified significant risk factors affecting the survival of breast cancer patients in Upper Assam, India. The key predictive factors for survival outcome were educational status, onset age of breast cancer diagnosis, BMI, stage at diagnosis, breastfeeding history and family history of cancer. The findings highlight the importance of early detection, education, better lifestyle and awareness in improving breast cancer survival rates. Future studies should focus on developing more robust models for estimating survival probability and assessing the level of awareness, treatment, and control of these risk factors.

References:

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et. al., Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca Cancer J. Clin.* 2024;74(3):229-263.

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et. al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca Cancer J. Clin.* 2021;71(3):209-249.
3. Breast Cancer statistics from Kolkata, Guwahati (https://www.breastcancerindia.net/statistics/stat_kolkata.html/ accessed on 06-11-2024)
4. Sathwara JA, Balasubramaniam G, Bobdey SC, Jain A, Saoba S. Sociodemographic factors and late-stage diagnosis of breast cancer in India: A hospital-based study. *IJMPO.* 2017;38(03):277-281.
5. Mehrotra R, Yadav K. Breast cancer in India: Present scenario and the challenges ahead. *World J. Clin. Oncol.* 2022;13(3):209-218.
6. Sheikh A, Hussain SA, Ghorri Q, Naeem N, Fazil A, Giri S, et. al., The spectrum of genetic mutations in breast cancer. *Asian Pac. J. Cancer Prev.* 2015;16(6):2177-2185.
7. Taplin SH, Ichikawa L, Yood MU, Manos MM, Geiger AM, Weinmann S, et. al., Reason for late-stage breast cancer: absence of screening or detection, or breakdown in follow-up?. *JNCI.* 2004;96(20):1518-1527.
8. Holmes MD, Hunter DJ, Colditz GA, Stampfer MJ, Hankinson SE, Speizer FE, et. al., Association of dietary intake of fat and fatty acids with risk of breast cancer. *Jama.* 1999;281(10):914-920.
9. DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013. *Ca Cancer J. Clin.* 2014;64(1):52-62.
10. Bhattacharjee S, Deka S. Analysis of breast cancer patients of NE India during 2016-2019. *JP J. Biostat.* 2022;19:1-3.
11. Dudley WN, Wickham R, Coombs N. An introduction to survival statistics: Kaplan-Meier analysis. *JADPRO.* 2016;7(1):91-100.
12. Ewertz M, Gillanders S, Meyer L, Zedeler K. Survival of breast cancer patients in relation to factors which affect the risk of developing breast cancer. *Int. J. Cancer.* 1991;49(4):526-530.
13. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J. Am. Stat. Asso.* 1958;53(282):457-481.
14. Siegel RL, Miller KD, Jemal A. Cancer statistics. *Ca Cancer J. Clin.* 2018;68(1):7-30.
15. Clark GM, McGuire WL. Steroid receptors and other prognostic factors in primary breast cancer. *Semin. Oncol.* 1988;15(2):20-25.
16. Cox DR. Regression models and life-tables. *J. R. Stat. Soc. Ser. B Stat. Method.* 1972;34(2):187-202.
17. Dikshit R, Gupta PC, Ramasundarahettige C, Gajalakshmi V, Aleksandrowicz L, Badwe R, et. al., Cancer mortality in India: a nationally representative survey. *Lancet.* 2012;379(9828):1807-1816.
18. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca Cancer J. Clin.* 2018;68(6):394-424.
19. ICMR-NCDIR, Profile of Cancer and Related Health Indicators in the North East Region of India – 2021, Bengaluru, India. (https://ncdirindia.org/All_Reports/NorthEast2021/resources/NE_Complete.pdf/ accessed on 06-11-2024)
20. Giaquinto AN, Sung H, Miller KD, Kramer JL, Newman LA, Minihan A, et. al., Breast cancer statistics, 2022. *Ca Cancer J. Clin.* 2022;72(6):524-541.
21. Agarwal G, Ramakant P. Breast cancer care in India: the current scenario and the challenges for the future. *Breast care.* 2008;3(1):21-27.
22. Gogoi G, Borgohain M, Saikia P, Fazal S. Profile of molecular subtypes of breast cancer with special reference to triple negative: A study from Northeast India. *Clin. Cancer Investig. J.* 2016;5(5-2016):374-383.

23. Sharma M, Sharma JD, Sarma A, Ahmed S, Kataki AC, Saxena R, et. al., Triple negative breast cancer in people of North East India: Critical insights gained at a regional cancer centre. *Asian Pac. J. Cancer Prev.* 2014;15(11):4507-4511.
24. Admoun C, Mayrovitz HN. The etiology of breast cancer. In: Breast Cancer. Exon Publications, Brisbane (AU); 2022.
25. Vishwakarma G, Ndetan H, Das DN, Gupta G, Suryavanshi M, Mehta A, et. al., Reproductive factors and breast cancer risk: A meta-analysis of case-control studies in Indian women. *South Asian J. Cancer.* 2019;8(02):80-84.
26. Gajalakshmi CK, Shanta V, Swaminathan R, Sankaranarayanan R, Black RJ. A population-based survival study on female breast cancer in Madras, India. *Br. J. Cancer.* 1997;75(5):771-775.
27. American Cancer Society. Breast Cancer Facts & Figures 2022-2024. Atlanta: American Cancer Society, Inc. 2022.
(<https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/2022-2024-breast-cancer-fact-figures-acf.pdf>)
accessed on 09-11-2024)
28. Kumar A, Bhagabaty SM, Tripathy JP, Selvaraj K, Purkayastha J, Singh R. 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 Pac. J. Cancer Prev.* 2019;20(12):3711-3721.
29. International Institute for Population Sciences (IIPS) and ICF. (2021). National Family Health Survey (NFHS-5), 2019-21: India: Volume 1, Mumbai: IIPS.
(<https://dhsprogram.com/pubs/pdf/FR375/FR375.pdf>) accessed on 09-11-2024)
30. Gupta A, Shridhar K, Dhillon PK. A review of breast cancer awareness among women in India: Cancer literate or awareness deficit?. *Eur. J. Cancer.* 2015;51(14):2058-2066.
31. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. Continuous Update Project Expert Report 2018. (<https://www.wcrf.org/wp-content/uploads/2021/02/Summary-of-Third-Expert-Report-2018.pdf>) accessed on 10-11-2024)
32. Shanker N, Mathur P, Das P, Sathishkumar K, Shalini AM, Chaturvedi M. Cancer scenario in NE India & need for an appropriate research agenda. *Indian J. Med. Res.* 2021;154(1):27-35.
33. Singal K, Malik VS, Sachdeva M, Chauhan A, Singh M, Rana M, et al., Prevalence of tobacco consumption among the Northeast population of India: A systematic review and meta-analysis. *Int. J. Noncommun. Dis.* 2023;8(4):212-222.
34. Singh P, Kapil U, Shukla NK, Deo SV, Dwivedi SN. Association of overweight and obesity with breast cancer in India. *Indian J. Community Med.* 2011;36(4):259-262.
35. Blair CK, Wiggins CL, Nibbe AM, Storlie CB, Prossnitz ER, Royce M, et. al., Obesity and survival among a cohort of breast cancer patients is partially mediated by tumor characteristics. *NPJ Breast Cancer.* 2019;5(1):33-39.
36. Ligibel J. Obesity and breast cancer. *Oncology.* 2011;25(11):994-1000.
37. Dhillon PK, Mathur P, Nandakumar A, Fitzmaurice C, Kumar GA, Mehrotra R, et. al., The burden of cancers and their variations across the states of India: the Global Burden of Disease Study 1990–2016. *Lancet Oncol.* 2018;19(10):1289-1306.
38. 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).
39. Ata N, Özel G. Survival functions for the frailty models based on the discrete compound Poisson process. *J. Stat. Comput. Simul.* 2013;83(11):2105-2116.

40. Richards MA, Westcombe AM, Love SB, Littlejohns P, Ramirez AJ. Influence of delay on survival in patients with breast cancer: a systematic review. *Lancet*. 1999;353(9159):1119-1126.
41. Nair MK, Sankaranarayanan R, Nair KS, Amma NS, Varghese C, Padmakumari G, et. al., Overall survival from breast cancer in Kerala, India, in relation to menstrual, reproductive, and clinical factors. *Cancer*. 1993;71(5):1791-1796.
42. Feleke B, Tesfaw LM, Mitku AA. Survival analysis of women breast cancer patients in Northwest Amhara, Ethiopia. *Front oncol*. 2022;12:1041245.
43. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50 302 women with breast cancer and 96 973 women without the disease. *Lancet*, 2002;360(9328):187-195.
44. Pakseresht S, Ingle GK, Bahadur AK, Ramteke VK, Singh MM, Garg S, et. al., Risk factors with breast cancer among women in Delhi. *Indian J. Cancer*. 2009;46(2):132-138.
45. Chang ET, Milne RL, Phillips KA, Figueiredo JC, Sangaramoorthy M, Keegan TH, et. al., Family history of breast cancer and all-cause mortality after breast cancer diagnosis in the Breast Cancer Family Registry. *Breast Cancer Res. Treat*. 2009;117:167-176.
46. Yusuf S, Okeke EN, Lasisi KE. Survival Analysis of Covariates Influencing Breast Cancer Treatment: A Case Study of North Eastern Nigeria. *Biom. Lett*. 2024;61(1):1-5.

Acknowledgement:

The authors are thankful to Dr. Dharmendra Dubey, Assistant Professor, Department of Biostatistics, AIIMS, New Delhi and Miss Asmita Tamuli, Research Scholar, Department of Statistics, Dibrugarh University for their support.

Conflict of interest:

NIL

List of Abbreviations:

Abbreviation	Definition
BC	Breast Cancer
BMI	Body Mass Index
NE	North-East India
CI	Confidence Interval
SD	Standard Deviation

List of Tables: **Table 1:** Distribution of socio-demographic, clinical characteristics with frequency of substance consumption among n=210 breast cancer participants.

Socio-demographic characteristics		
Distribution of Educational Status		
Literate	168	80.0 %
Illiterate	42	20.0%
Breakdown of literates:		
Primary	92	43.8%
Secondary	57	27.2 %
Graduate	19	9.05 %
Distribution of Age:		
Below 45	76	36.2%
45 & Above	134	63.8%
Distribution of Employment Status		
House wife	206	98.0 %
Tea garden worker	4	2.0 %
Caste Distribution		
Unreserved	22	10.5%
Reserved	188	89.5%
Marital Status		
Married	202	96.2 %
Unmarried	8	3.8 %
Distribution of Religion		
Hindu	198	94.2%
Muslim	12	5.7 %
Clinical characteristics		
Mode of delivery		
Vaginal	180	85.8 %
Caesarean	30	14.2 %
Breast feeding status		
Yes	172	81.9 %
No	38	18.1%
Cancer on Breast		
Left	112	53.4 %
Right	97	46.2 %
Breast cancer stage distribution		
First (I)	47	22.4 %
Second (II)	48	22.8 %
Third (III)	59	28.1%
Fourth (IV)	47	26.7 %
Family history of cancer		
Yes	19	9.1 %
No	191	90.9 %
BMI		
Underweight	14	6.7%
Normal	148	70.4%
Overweight	35	16.7%
Obese	13	6.2%
Status of substance consumption		

Alcohol Consumption		
Yes	11	5.3%
No	199	94.7 %
Use of tobacco		
Yes	20	9.5 %
No	190	90.5%

Table 2: Survival rates of n=210 BC patients according to significant socio-demographic and clinical factor.

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	95	70	39
	45 & Above	89	49	15
BMI	Underweight	90	71	-
	Normal	90	65	21
	Overweight	90	41	20
	Obese	70	41	-
Cancer Stage	Stage I	95	75	75
	Stage II	95	70	60
	Stage III	89	60	22
	Stage IV	85	31	5
Breastfeeding	Womenwith children (Yes)	92	58	21
	Women with no Children (No)	80	41	-
Family History	Yes	70	30	-
	No	90	55	21
Overall Survival		95	70	38

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

Ris k Fac tors	Status	No. of cases	Perce nt	Mean Survival Time(MST) in Weeks	SD@ of MST	Media n Survi val Time in Weeks	Test Stati stic	p-value
1.Educa tion status	Primary	92	43.8	109.8	66.7	111.0	Kruskalwalli s Test U=2.221	0.528
	Seconda ry	57	27.2	101.5	63.0	100.0		
	Graduat e	19	9.05	121.1	63.1	127.0		
	Illiterate	42	20.0	115.2	67.2	114.0		

2.Caste	Unreserved	22	10.5	82.7	55.2	75.0	Mann-Whitney Test U=1534.0	0.048*
	Reserved	188	89.5	111.6	65.8	113.5		
3. Age	Below 45	76	36.2	106.7	66.6	97.5	Mann-Whitney Test U=4948.5	0.733
	45 & above	134	63.8	109.7	64.7	111.0		
4.4.Religion	Hindu	198	94.2	108.6	65.4	105.0	Mann-Whitney Test U=1688.0	0.365
	Muslim	12	5.7	108.9	66.2	126.0		
5.Marital status	Married	202	96.5	108.4	65.9	104.0	Mann-Whitney U=2277.0	0.174
	Unmarried	8	3.8	114.3	49.2	129.0		
6.Employment Status	Housewife	206	98.0	110.5	65.8	108.0	Mann-Whitney U=10441.5	0.028*
	Professional	4	2.0	120.0	23.07	114.0		
7.Cancer Stage	Stage I	47	22.4	105.4	63.3	96.0	Kruskalwallis Test= 4.903	0.179
	Stage II	48	22.8	105.0	60.2	100.0		
	Stage III	59	28.1	107.2	65.8	109.0		
	Stage IV	56	26.7	116.0	71.3	114.0		
8.Breastfeeding Status	Yes	172	81.9	107.9	64.8	104.0	Mann-Whitney U=1673	0.024*
	No	38	18.1	143.9	62.0	142.0		
9.Delivery mode	Vaginal	180	85.8	106.3	65.4	104.0	Mann-Whitney U=2886.0	0.492
	Cesarean	30	14.2	123.9	65.0	125.0		
10.cancer Breast	Left	112	53.4	108.9	60.0	111.0	Mann-Whitney U=7994.0	0.632
	Right	97	46.2	107.3	70.9	100.0		
11.Family History of cancer	Yes	19	9.1	95.3	66.4	84.0	Mann-Whitney U=1963.5	0.851
	No	191	90.9	109.9	65.2	108.0		
12.BMI	Underweight	14	6.7	76.1	60.0	54.2	Kruskalwallis Test= 0.931	0.818
	Normal	148	70.4	111.0	65.6	111.0		
	Overweight	35	16.7	109.4	61.9	104.0		
	Obesity	13	6.2	113.6	72.5	126.0		
13.Consumed alcohol	Yes	11	5.3	102.0	72.4	83.0	Mann-Whitney U=1873.5	0.788
	No	199	94.7	108.9	65.0	109.0		
14.Used tobacco	Yes	20	9.5	105.5	66.5	84.0	Mann-Whitney U=0.3341	0.982
	No	190	90.5	109.9	65.2	108.0		

@SD=Standard Deviation

*=Significant at 5% level of Significance

Table 4: Estimated Mean and Median Survival Times with the individual factors and Log Rank Test value and p-value used in survival analysis to compare the survival distributions of two or more group.

Factor	Status	Cases	Died (%)	Mean Survival Time (in weeks)	Median Survival Time (in weeks)	SD @ of MST	CI for MST	Log rank value	p-value
Education	Illiterate	42	33	127.6	147.0	31.5	[85.2, 208.7]	19.7	0.0002*
	Primary	92	24	197.8	265.0	0.0	[-, -]		
	Secondary	57	17	170.8	195.0	31.5	[133.1, 256.8]		
	Graduate	19	4	178.4	212.0	45.2	[132.3, 309.6]		
Age	Below 45	76	13	205.4	238.0	31.5	[176.2, 299.9]	12.2	0.0004*
	45& Above	134	65	155.8	154.0	5.12	[143.9, 164.0]		
Caste	Unreserved	22	4	163.4	221.0	77.7	[68.6, 373.3]	0.190	0.633
	Reserved	188	74	170.4	167.0	9.6	[148.2, 185.3]		
Religion	Hindu	198	73	172.9	177.0	16.9	[143.8, 210.2]	0.614	0.433
	Muslim	12	5	159.8	162.0	31.1	[101.1, 222.8]		
Marital Status	Married	202	76	170.7	171.0	14.8	[142.0, 200.3]	0.132	0.717
	Unmarried	8	2	160.0	167.0	0.0	[-, -]		
Employment status	Housewife	206	75	172.4	171.0	15.7	[140.4, 201.7]	0.910	0.340
	Professional	4	3	134.3	164.0	19.0	[76.7, 151.3]		
Stages	Stage I	47	8	204.4	212.0	13.6	[-, -]	16.3	0.001*
	Stage II	48	9	203.8	-	10.3	[-, -]		
	Stage III	59	21	170.4	171.0	13.9	[143.6, 198.3]		
	Stage IV	56	40	139.4	140.0	20.8	[99.2, 180.8]		
Breastfeeding status	Yes	172	51	182.7	195.0	21.8	[152.3, 237.7]	9.93	0.002*
	No	38	27	132.3	140.0	26.8	[88.7, 191.3]		
Delivery Mode	Vaginal	180	70	166.7	162.0	9.2	[144.1, 179.9]	0.574	0.449
	Cesarean	30	8	174.9	221.0	45.8	[131.2, 310.8]		

Cancer Breast	Left	112	45	156.8	157.0	4.8	[147.7, 166.2]	1.69 2	0.193
	Right	97	33	179.3	204.0	22.7	[159.6, 248.4]		
Family History	Yes	19	13	129.3	140.0	38.8	[63.9, 216.0]	7.5	0.006*
	No	191	65	192.6	206.0	9.64	[187.1, 224.9]		
BMI	Underweight	14	2	156.7	-	-	[-, -]	8.8	0.032*
	Normal	148	47	180.3	195.0	19.6	[156.6, 233.4]		
	Overweight	35	19	146.9	143.0	22.7	[98.4, 187.6]		
	Obese	13	10	126.8	139.0	35.4	[69.5, 208.5]		
Consumption of Alcohol	Yes	11	2	198.8	-	-	[-, -]	0.16 4	0.685
	No	199	76	169	205.0	13.6	[182.4, 216.7]		
Consumption of tobacco	Yes	20	6	164.3	162.0	14.5	[141.0, 236.9]	0.21 4	0.644
	No	190	72	170.2	167.0	10.4	[148.1, 185.9]		

@ = Standard Deviation; *= Significance at 95% CI

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

Factors	B	$SE^{\dagger} of \beta$	e^{β}	$CI^{\dagger} of e^{\beta}$	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)	0.837	0.323	2.309	[1.226, 4.384]	2.592	0.009*
BMI(Underweight)	<i>Reference</i>					
BMI(Normal)	0.295	0.442	1.343	[0.564, 3.201]	0.668	0.504
BMI(Overweight)	0.598	0.470	1.819	[0.723, 4.575]	1.217	0.203
BMI(Obese)	0.890	0.526	2.437	[0.867, 6.846]	1.691	0.090
Cancer stage(I)	<i>Reference</i>					
Cancer stage(II)	0.345	0.500	1.412	[0.529, 3.765]	0.689	0.491
Cancer stage(III)	0.498	0.439	1.646	[0.529, 3.894]	1.136	0.256
Cancer stage(IV)	0.819	0.411	2.270	[1.014, 5.092]	1.994	0.036*
Breastfeeding(No)	<i>Reference</i>					
Breastfeeding (Yes)	0.548	0.275	1.731	[1.009, 2.969]	1.513	0.046*
Family History(No)	<i>Reference</i>					

Family History(Yes)	-0.594	0.326	0.552	[0.2910, 1.046]	-1.823	0.063
Likelihood ratio test= 59.78 on 15 df, p=3e-07 Wald test = 62.19 on 15 df, p=1e-07 Score (logrank) tes = 71.52 on 15 df, p=2e-09						

*= Significant at 95% CI

Table 6.1: Stage wise comparison of estimated Survival results with National Figure.^[41]

Factor	5- year survival(OS) rate		Median Survival Time		p-value	
	Estimated from present study (%)	National Reference (%)	Estimated from present study (Weeks)	National Reference (weeks)	Estimated from present study	National Reference
Stage I	75	98	212	-	0.000	<0.01
Stage II	60	65	-	-		
Stage III	22	33	171	130.36		
Stage IV	5	6	140	78.27		

Table 6.2: Stage wise comparison of estimated survival results with International Figure.^[42]

Factor	Mean survival time		Median survival time		Hazard ratio		p-value	
	Estimated from present study (Weeks)	International Reference (Weeks)	Estimated from present study (Weeks)	International Reference (Months)	Estimated from present study	International Reference	Estimated from present study	International Reference
Stage I	220.3	197.4	212	-	-	-	-	-
Stage II	199.6	220.3	-	-	1.412	3.973	0.491	0.195
Stage III	107.7	152	165	156.5	1.646	14.997	0.256	0.008
Stage IV	158.6	146.1	159	147.8	2.27	11.936	0.036	0.016

Figure Legends:

Figure 1: Kaplan-Meier curve for overall survival function of breast cancer patients.

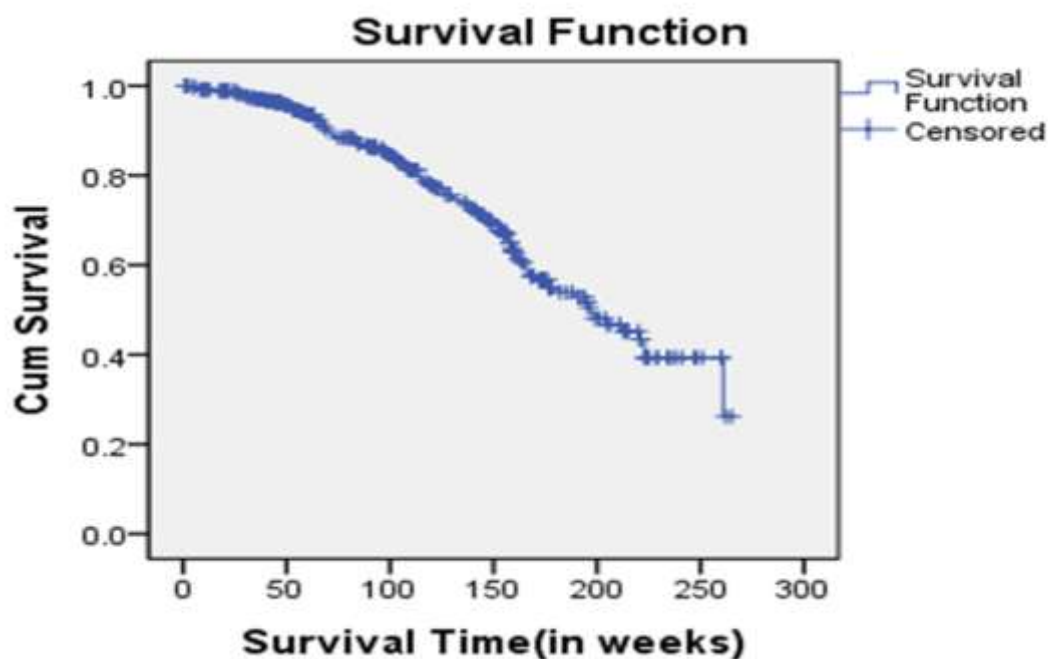


Figure 2: Kaplan-Meier curve for survival function of breast cancer patients in association with socio-demographic and other studies factors: (a) education status (b) Age (c) BMI (d) breastfeeding (e) stage of breast cancer at diagnosis (f) family history of cancer:

