JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Assessment of Sociodemographic, Lifestyle and Biochemical Risk Factors of Breast Cancer using Binary Logistic Regression: A Case Study of Breast Cancer Patients in Kalaburagi City, Karnataka State, India

SUJATHA INGINSHETTY [1*], SHWETA.S[1], M.N. MEGERI [2]

[1*]Professor, Department of Statistics, Gulbarga University, Kalaburagi-585106, Karnataka.
Corresponding Author E-mail address: sujathainginshetty22@gmail.com
[1] Research Scholar, Department of Statistics, Gulbarga University, Kalaburagi, Karnataka.

[2] Professor, Department of Statistics, Karnatak Science College, Dharwad-580001, Karnataka.

ABSTRACT:

Background: Breast cancer remains one of the most predominant cancers among women worldwide, significantly impacting public health. Various sociodemographic, lifestyle, and biochemical factors contribute to breast cancer risk. Identifying these factors through statistical modelling aids in early diagnosis and targeted interventions.

Aim & Objective: To examine the association between sociodemographic, lifestyle, and biochemical factors with risk of breast cancer among women.

Methods: A total of 290 women cancer patients, including 140 with breast cancer and 150 with other cancers were randomly selected. Data were collected using a structured questionnaire covering lifestyle habits, biochemical risk factors, cancer stage, and diagnosis. Statistical analyses were conducted using SPSS software, estimating crude odds ratios and 95% confidence intervals. A binary logistic regression model was utilized to identify significant risk factors.

Results: A total of 290 women patients with cancer, 51.7% had different types of cancer and 48.3% had breast cancer. Age, literacy level, Tobacco consumption, physical inactivity, minimal intake of fruit and vegetable, BMI, cancer stage and elevated CRP/cholesterol levels were strongly associated with the risk of breast cancer. Sleep duration, stress, sugary beverage intake, and red/processed meat consumption

showed no significant association. Advanced-stage cases (58.5%) were more common than early-stage cases (29.4%), reinforcing the need for early detection.

Conclusion: This study highlights the significant role of age, literacy, lifestyle habits, and biochemical markers in breast cancer risk. Early detection, awareness programs, and preventive strategies are crucial for reducing the burden of breast cancer. The binary logistic regression model effectively determined key risk factors, making it a valuable tool for risk estimation. Further longitudinal research is recommended to evaluate long-term impacts.

Keywords **: Breast cancer, Sociodemographic characteristics, Risk factors, Odds ratio, Binary logistic regression**

I. INTRODUCTION

Breast cancer stands as one of the most predominant cancer affecting women across the world, posing a significant public health challenge (Sung et al.,2021). The complex etiology of breast cancer involves a multitude of interacting factors, encompassing sociodemographic characteristics, lifestyle choices, and biochemical markers (Anand et al., 2020). Understanding the risk factors is crucial for early detection, targeted interventions, and ultimately, improved patient outcomes. Epidemiological studies play a vital role in identifying these associations and quantifying their impact on breast cancer risk (Liao et al., 2022). Several sociodemographic factors, such as age, marital status, and socioeconomic status, have been implicated in breast cancer risk (Key et al., 2017). For instance, increasing age is a well-established risk factor for breast cancer, with incidence rates rising significantly after menopause (DeSantis et al., 2019).

Lifestyle factors, including tobacco use, physical inactivity, dietary/nutrition habits, and BMI, also contribute to develop breast cancer (Aune et al., 2017). Tobacco consumption has been found to be linked with a higher incidence of multiple malignancies, including breast cancer mainly in females in the reproductive phase (Gandini et al., 2017). Conversely, regular physical activity has been shown to acquire a protective implication against breast cancer, possibly through its influence on hormone levels and immune function (Schmid & Leitzmann, 2014).

Food consumption patterns, particularly large amounts of red and processed meats in the diet and low consumption of fruits and vegetables, have been linked to increased breast cancer risk (Bradlow et al., 2022). Furthermore, obesity and overweight, as measured by BMI, are recognized key influences on postmenopausal breast cancer occurrence (Dixon et al., 2021).

Biochemical markers like C-reactive protein (CRP), homocysteine and vitamin D have also emerged as potential indicators of breast cancer risk (Goodman et al., 2017). CRP, an inflammatory marker, has been connected with increased risk of developing breast cancer, suggesting a role for chronic inflammation in breast cancer development (Pierce et al., 2016).

The present study aimed to explore the association between a range of sociodemographic, lifestyle, biochemical factors & risk of breast cancer in women from Kalaburagi city. We employed a binary logistic regression analysis for examining influential risk factors and estimate their impact on breast cancer risk. This statistical approach is well-suited for analyzing dichotomous outcomes, such as the presence or absence of breast cancer, and allows for the simultaneous assessment of multiple predictor variables. By identifying key risk factors, this study contributes to the growing body of knowledge on breast cancer epidemiology and informs the development of targeted prevention and intervention strategies.

II. METHODOLOGY

A cross-sectional study was performed using a pre-tested questionnaire to ensure validity and reliability. The aim of the study was thoroughly described to the participants, & informed consent was secured before data collection commenced. Overall, 290 women (140 diagnosed with breast cancer & 150 with other types of cancers) were randomly selected and interviewed from two prominent cancer care hospitals in Kalaburagi city over a three-months period. The study was limited to these two hospitals, and data collection focused exclusively on interviewing cancer patients to assess their sociodemographic, lifestyle, and biochemical risk factors.

The breast cancer status of women is considered as the dependent variable. For analysis purposes, the breast cancer status is grouped as 1 if diagnosed with breast cancer and 0 if diagnosed with other cancers, making it a dichotomous response variable. Apart from the response variable, data on independent variables were obtained from the structured questionnaire.

In the logistic model, we focus on cases where the response variable is a categorical random variable with only two possible outcomes, binary or dichotomous.

It is standard practice to let Y (Breast cancer status) be a two-category binary/dichotomous response variable, which is defined as:

$$\textbf{Y (Breast cancer status)} = \begin{cases} 1, & \text{if diagnosed with breast cancer,} \\ 0, & \text{if diagnosed with other cancers} \end{cases}$$

To explore a relationship between a set of covariates i.e. $X^{1} = (X_1, X_2, ..., X_p)$ and the presence or absence of breast cancer in the study, binary logistic regression analysis was applied. The crude odds ratio and 95% confidence interval were computed for each predictor variable to assess the strength of the connection with breast cancer risk.

III. RESULTS & DISCUSSION

The study sample included 290 female cancer patients, of whom 140 (48.3%) were diagnosed with breast cancer, while the remaining 150 (51.7%) had other types of cancers.

Table 1: Association of various Sociodemographic Factors with Risk of Breast Cancer

Parameters	Breast Cancer		Total (N = 290)	Crude odds ratio (OR)	95 % C. I		p-value
	No (N=150)	Yes (N=140)			Lower	Upper	
Age group							
20-49	49 (36.6)	85 (63.4)	134 (46.2)	0.3139	0.1940	0.5079	**0.0001***
50 and above	101(64.7)	55 (35.3)	156 (53.8)				
Literacy Level							
Illiterate	115 (63.2)	67 (36.8)	182 (62.8)	3.5800	2.1646	5.9208	**0.0001***
literate	35 (32.4)	73 (67.6)	108 (37.2)				
Employment Status							
Unemployed	134 (53.4)	117 (46.6)	251 (86.6)	1.6464	0.8303	3.2646	0.1534
Employed	16 (41.0)	23 (59.0)	39 (13.4)				
Annual Income							
Less than 3 lakhs	124 (51.5)	117 (48.5)	241 (83.1)	0.9375	0.5067	1.7347	0.8372
3 lakhs and above	26 (53.1)	23 (46.9)	49 (16.9)				

Note: The figures within the parenthesis are the percentage values

* Indicates that the result is significant at 5 % level of significance.

The above table results suggested that the risk of breast cancer significantly decreases with age. Women aged 20–49 years had a higher prevalence of breast cancer (63.4%) compared to those aged 50 years and above (35.3%). The estimated odds ratio of age is (OR = 0.31) , indicating that older women have significantly lower odds of developing breast cancer compared to younger women.

The proportion of illiterate women is higher overall (62.8%), but among breast cancer patients, literate women (67.6%) have a higher proportion compared to illiterate women (36.8%). The estimated odds ratio of literacy level is (OR = 3.58), indicating that literate women have significantly higher odds of having breast cancer compared to illiterate women. However, since the overall number of illiterate women is higher, illiteracy remains an important factor in breast cancer risk. This suggests that lack of knowledge and awareness about breast cancer among illiterate women may contribute to late detection and diagnosis, emphasizing the need for targeted awareness programs. The result is statistically significant, indicating a strong association between literacy level and breast cancer risk.

Employment status did not show a significant association with breast cancer, although employed women had a slightly higher prevalence (59.0%) compared to unemployed women (46.6%). The estimated odds ratio of employment status is (OR=1.65), suggesting that employment alone may not be a major determinant of breast cancer risk.

Similarly, annual income did not show a strong association with breast cancer risk. Women earning 3 lakh rupees and above had a slightly lower prevalence of breast cancer (46.9%) compared to those earning less than 3 lakh rupees (48.5%). The estimated odds ratio of income level is (OR = 0.937), indicating no statistically significant difference in breast cancer risk based on income.

These results highlight the significant association of age and literacy level with breast cancer risk, while employment and income do not show a statistically significant relationship. The findings emphasize the importance of awareness initiatives and early screening, particularly for older and less educated women, to improve early detection and breast cancer outcomes.

Table 2: Association of Lifestyle, Biochemical, and Clinical Factors with Risk of Breast Cancer

Parameters	Breast cancer		Total (N=290)	Crude odds ratio (OR)	95 % C. I		p-value
	No (N=150)	Yes (N=140)			Lower	Upper	
Consumption of tobacco products							
No	50 (32.9)	102 (67.1)	152 (52.4)	0.1863	0.1125	0.3083	0.0001*
Yes	100 (72.5)	38 (27.5)	138 (47.6)				
Physical activity							
No	135 (54.4)	113 (45.6)	248 (85.5)	2.1504	1.0907	4.2398	0.0271*
Yes	15 (35.7)	27 (64.3)	42 (14.5)				
Sleep per night							
Less than 8 hours	78 (56.5)	60 (43.5)	138 (47.6)	1.4444	0.9088	2.2957	0.1198
More than 8 hours	72 (47.4)	80 (52.6)	152 (52.4)				
Stress							
No	21 (46.7)	24 (53.3)	45 (15.5)	0.7868	0.4161	1.4879	0.4608
Yes	129 (52.7)	116 (47.3)	245 (84.5)				
Consumption of fruits & vegetables							
Low	127 (49.2)	131 (50.8)	258 (89.0)	0.3794	0.1690	0.8514	0.0188*
High	23 (71.9)	9 (28.1)	32 (11.0)				

Consumption of Sugary beverages							
Low	38 (50.0)	38 (50.0)	76 (26.2)	0.917	0.5395	1.5337	0.7263
High	112 (52.3)	102 (47.7)	214 (73.8)				
Eating Red or Processed meats							
Low	71 (50.4)	70 (49.6)	141 (48.6)	0.8987	0.5668	1.4250	0.6499
High	79 (53.0)	70 (47.0)	149 (51.4)				
BMI							
Low BMI	41 (66.1)	21 (33.9)	62 (21.4)	2.1315	1.1855	3.8323	0.0115*
High BMI	109 (47.8)	119 (52.2)	228 (78.6)				
Biochemical risk factor 1							
High CRP/ Cholesterol	86 (46.0)	101 (54.0)	187 (64.5)	0.5189	0.3175	0.8479	0.0088*
Low Hb/High vitamin D	64 (62.1)	39 (37.9)	103 (35.5)				
Biochemical risk factor 2							
High antioxidants /Omega-3 fatty acids	58 (58.6)	41 (41.4)	99 (34.1)	1.5233	0.9322	2.4858	0.0931
High homocysteine /Calcium	92 (48.2)	99 (51.8)	191 (65.9)				
Stages							
Early stage	72 (70.6)	30 (29.4)	102 (35.2)	0.3846	2.0216	5.6666	0.0001*
Advanced stage	78 (41.5)	110 (58.5)	188 (64.8)				
Diagnosis							
Less than 1 year	109 (50.7)	106 (49.3)	215 (74.1)	0.8527	0.5032	1.4450	0.5538
1-3 years	41 (54.7)	34 (45.3)	75 (25.9)				

Note: The figures within the parenthesis are the percentage values.

***** Indicates that the result is significant at 5 % level of significance.

Tobacco consumption was found to be significantly associated with breast cancer risk, with an odds ratio of 0.1863 (p = 0.0001), indicating that individuals who do not consume tobacco have lower odds of developing breast cancer.

Physical inactivity was found to be significantly associated with breast cancer risk, with an odds ratio of 2.1504. Among breast cancer patients, 80.7% were physically inactive, whereas 19.3% were physically active. This indicates that individuals with lower physical activity levels are more likely to develop breast cancer.

As for sleep duration, there was no strong association with breast cancer risk (OR = 1.4444). Stress showed no significant relationship with breast cancer risk (OR = 0.7868). Among the total study population, 84.5% reported experiencing stress, while 15.5% did not experience stress. Among breast cancer patients, 47.3% experienced stress, meanwhile, 53.3% did not. This finding suggests that stress levels were high among participants, but the study results suggest that stress is not significantly associated with breast cancer risk.

Consumption of fruits and vegetables showed a significant correlation with a higher risk of breast cancer (OR = 0.3794). Among breast cancer patients, 50.8% had low fruits and vegetables intake, compared to 28.1% who had a high intake. This finding suggests that inadequate fruits and vegetables intake may be a contributing factor to breast cancer risk. High consumption of sugary beverages was not found to have a significant association with breast cancer risk, with an odds ratio of 0.917. Among breast cancer patients, 50.0% consumed low amounts of sugary beverages, whereas 47.7% consumed high amounts. This indicates that sugary beverage intake does not appear to have a strong influence on breast cancer risk in this study. A high intake of red or processed meats did not show a significant association with breast cancer risk, with an odds ratio of 0.8987.

BMI was found to be a significant factor for breast cancer risk (OR = 2.1315). Among breast cancer patients, 33.9% had a low BMI, whereas 52.2% had a high BMI. This finding implies that individuals with a higher BMI may be more susceptible to developing breast cancer. However, some individuals who were overweight before diagnosis may have shifted to the normal BMI category due to weight loss caused by treatment and disease progression. This post-diagnosis weight reduction should be considered when interpreting BMI-related risk factors.

Among biochemical risk factors, women with elevated CRP levels showed a significantly raised prevalence of breast cancer (54.0%) compared to those with low haemoglobin or elevated vitamin D (37.9%). The estimated odds ratio of CRP levels is (OR=0.52), suggests a significant relationship between high CRP levels & breast cancer risk.

The stage of cancer at diagnosis was an important aspect. Patients whose cancer was detected in its early stage had lower odds of developing advanced breast cancer (OR = 0.3846). Among breast cancer patients, 29.4% were diagnosed at an early stage, while 58.5% were diagnosed at an advanced stage. This study indicates that detecting and diagnosing breast cancer at an early stage may help lower the risk of progression to advanced stages**.** The time elapsed since diagnosis showed no strong association with breast cancer risk (OR = 0.8527).

The findings suggest that consumption of tobacco products, physical inactivity, inadequate fruit and vegetable intake, and BMI are significantly correlated with an increased risk of developing breast cancer, while sleep duration, stress, sugary beverage consumption, red/processed meat intake, and time since diagnosis do not show strong associations**.** Early-stage diagnosis is linked to a reduced risk of progressing to advanced breast cancer**,** emphasizing the need for early diagnosis and timely intervention.

IV. CONCLUSION

In this study, we identified that various sociodemographic, lifestyle, and biochemical factors were significantly associated with the risk of developing breast cancer. Women aged 50 and above had significantly lower odds of developing breast cancer compared to younger women. However, literacy level was significantly associated, with illiterate women having substantially higher odds of developing breast cancer compared to literate women. Employment status and income level did not show a significant association with breast cancer risk. Individuals who do not consume tobacco have lower odds of developing breast cancer. Physically inactive women had significantly higher odds of developing breast cancer, emphasizing the role of an active lifestyle in cancer prevention. Additionally, inadequate fruit and vegetable intake was associated with a higher risk of breast cancer.

BMI was a significant factor, with some overweight individuals shifting to normal BMI due to post-diagnosis weight loss. High CRP/cholesterol levels were significantly related with breast cancer risk, whereas other biochemical markers showed no strong relationship.

Advanced-stage breast cancer cases were more common than early-stage diagnoses, highlighting the importance of early detection and timely medical intervention. The binary logistic regression model effectively identified key predictors of breast cancer, reinforcing the need for self-awareness and healthcare-seeking behaviour in women, strengthening awareness programs, promoting early screening, and implementing targeted preventive strategies to reduce the burden of breast cancer. Further research with a longitudinal perspective is recommended to better understand the long-term impact of these factors on breast cancer development and progression.

V. ACKNOWLEDGMENT

We thank the cancer care hospitals and study participants for their co-operation. We also acknowledge the support of all those who contributed to this research.

VI. REFERENCES

[1] Anand, Preetha, Kunnumakkara, Ajaikumar B., Sundaram, Chitra, Harikumar, Kuzhuvelil B., Tharakan, Sheeja T., Lai, Oiki S., Sung, Bokyung, & Aggarwal, Bharat B. (2020). Cancer is a preventable disease that requires major lifestyle changes. *Pharmaceutical Research, 25*(9), 2097–2116.

[2] Aune, Dagfinn, Giovannucci, Edward, Boffetta, Paolo, Fadnes, Lars T., Keum, NaNa, Norat, Teresa, Greenwood, Darren C., Riboli, Elio, Vatten, Lars J., & Tonstad, Serena. (2017). Fruit and vegetable intake and the risk of cardiovascular disease, total cancer, and all-cause mortality—a systematic review and dose-response meta-analysis of prospective studies. *International Journal of Epidemiology, 46*(3), 1029–1056.

[3] Bradlow, H. Leon, Michnovicz, Jerzy J., Telang, Nitish T., & Osborne, Marian P. (2022). Effects of dietary indole-3-carbinol on estradiol metabolism and spontaneous mammary tumors in mice. *Carcinogenesis, 12*(9), 1571–1574.

[4] DeSantis, Carol E., Ma, Jiemin, Gaudet, Mia M., Newman, Lori A., Miller, Kimberly D., Sauer, Agnes G., Jemal, Ahmedin, & Siegel, Rebecca L. (2019). Breast cancer statistics, 2019. *CA: A Cancer Journal for Clinicians, 69*(6), 438–451

[5] Dixon, Suzanne C., Nagle, Christina M., Wark, Jessica D., Courneya, Kerry S., DeFazio, Anna, Friedenreich, Christine M., & Webb, Penelope M. (2021). Patterns of weight change and their association with lifestyle factors in Australian ovarian cancer survivors. *Cancer Causes & Control, 32*(7), 757–770.

[6] Gandini, Sara, Botteri, Edoardo, Iodice, Stefania, Boniol, Mathieu, Lowenfels, Albert B., Maisonneuve, Patrick, & Boyle, Peter. (2017). Tobacco smoking and cancer: A meta-analysis. *International Journal of Cancer, 122*(1), 155–164.

[7] Goodman, Marc, Bostick, Roberd M., Kucuk, Omer, & Jones, Derek P. (2017). Clinical trials of antioxidants as cancer prevention agents: Past, present, and future. *Free Radical Biology and Medicine, 51*(5), 1068–1084.

[8] Key, Timothy J., Appleby, Paul N., Reeves, Gillian K., Travis, Ruth C., Alberg, Anthony J., Barricarte, Aurelio, Berrino, Franco, Krogh, Vittorio, Sieri, Sabina, & Bueno-de-Mesquita, H. Bas. (2017). Carotenoids, retinol, tocopherols, and risk of breast cancer: A pooled analysis of eight cohort studies. *American Journal of Clinical Nutrition, 83*(3), 357–373.

[9] Liao, Min-Hsiang, Hsieh, Ming-Ju, Shih, Yi-Wen, Hsin, In-Yu, Wu, Jun-Yi, Chang, Chia-Wei, Cheng, Chia-Fu, Chou, Yu-Ting, Chen, Yu-Jen, & Lin, Han-Chung. (2022). Impact of social determinants on breast cancer screening adherence in Asian countries. *Cancer Epidemiology, 77*, 102116.

[10] Pierce, Brandon L., Ballard-Barbash, Rachel, Bernstein, Leslie, Baumgartner, Kathy B., Neuhouser, Marian L., Wener, Mark H., Baumgartner, Richard N., Gilliland, Frank D., Sorensen, Barry, & McTiernan, Anne. (2016). Elevated biomarkers of inflammation are associated with reduced survival among breast cancer patients. *Journal of Clinical Oncology, 27*(21), 3437–3444.

[11] Schmid, Daniela, & Leitzmann, Michael F. (2014). Association between physical activity and mortality among breast cancer and colorectal cancer survivors: A systematic review and meta-analysis. *Annals of Oncology, 25*(7), 1293–1311.

[12] Sung, Hyuna, Ferlay, Jacques, Siegel, Rebecca L., Laversanne, Mathieu, Soerjomataram, Isabelle, Jemal, Ahmedin, & Bray, Freddie. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 71*(3), 209–249.