Review Journal of Social Psychology & Social Works
http://socialworksreview.com
ISSN-E: 3006-4724 Volume: 3
ISSN-P: 3006-4716 Issue: 1 (2025)

# A Case-Control Epidemiological Analysis of Breast Cancer Risk Factors Using Logistic Regression Modeling

**Ramla Shah[1], Sidra Younas[2], Dr. Maryam Sadiq[3]**

1. Department of Statistics, University of Azad Jammu & Kashmir
ramlashah193@gmail.com
2. Department of Statistics, University of Azad Jammu & Kashmir
sidra.younas@ajku.edu.pk
3. Department of Statistics, University of Azad Jammu & Kashmir
hussainulahmad@gmail.com

**Abstract**
Breast cancer continues to be a major health problem worldwide with a complex etiology involving genetic, environmental, and lifestyle causes. The aim of this research is to explore the relationship between the most significant risk factors and breast cancer as well as determine the most frequent symptoms and most affected breast side in women. The case-control study had a sample size of 459 women and used chi-squared tests of independence and binary logistic regression to compare possible risk factors. The findings indicate that age at diagnosis, residential location, level of education, age at menarche, family history of breast cancer, menopausal status, history of breastfeeding, use of oral contraceptives, radiation exposure, drug use, number of children, and night shift work were all significantly related to breast cancer risk. On the other hand, marital status, age at marriage, abortion history, removal of ovaries, and physical activity are not observed to be statistically significant with the development of breast cancer. The results emphasize the need for public health intervention and education to reduce the risk of breast cancer through increased awareness and preventive interventions.

**Keywords**: Risk Factors, Breast Cancer, Logistic Regression, Epidemiology, Chi-Square

## Introduction

Breast cancer is a leading cause of morbidity and mortality in women globally, including in Pakistan (Ferlay et al., 2010; Sohail & Alam, 2007). The increasing incidence and mortality rates warrant thorough knowledge of related risk factors to develop effective preventive strategies (Weir et al., 2007). Intensive studies highlight the multifactorial origin of breast cancer etiology, which includes genetic susceptibility, hormonal factors, lifestyle factors, and environmental factors (American Cancer Society, 2013). One of the best-known risk factors, age has a significant bearing, with rising incidence after the age of 50 years (Khokher et al., 2012). Family history also increases the risk, the likelihood depending upon the closeness of relationship to affected relatives. Other contributory causes are reproductive history, breast density, use of hormonal contraceptives, ionising radiation exposure, and life style factors like diet, smoking, alcohol intake, and night work (Weir et al., 2007; Shin et al., 2011; Zienolddiny et al., 2013). The major goal of this research is to examine and measure the relationship between these risk factors employing sound statistical techniques, hence making empirical findings available to strengthen breast cancer prevention efforts.

## Methods

### Study Design and Data Collection

The study population had 459 female participants, of which 153 were breast cancer cases and 306 were controls. Ethical approval was granted by the institutional ethics committee, and written consent was received from all respondents. Data were collected using structured questionnaires, encompassing demographic details, reproductive history, lifestyle factors, and medical history.

### Statistical Analysis

Descriptive statistical methods were employed to summarize the demographic cohort under study. The $\chi^2$ test of independence was applied to assess relationships between categorical variables and breast cancer occurrence. Binary logistic regression analysis was performed to compute odds ratios (ORs) and 95% confidence intervals (CIs) for statistically significant predictors. Statistical analysis was conducted using SPSS v16.0 and STATA v11.

### Statistical Techniques

The following statistical techniques were used.

#### 2.3.1 Chi-squared test of independence

Chi-squares test of independence is used to test the hypothesis about independence of binary dependent variables on each explanatory variable separately. The test-statistic is

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Where

$O_i$= Observed frequency

$E_i$ = Expected frequency

$n$ = the numberof cells in the table

### Logistic Regression

Logistic regression model with two explanatory variables is

$$logit\ (p) = \alpha_0 + \beta_1 X_1 + \beta_2 X_2$$

## Results

### Symptom Prevalence and Affected Breast Side

Among the 153 breast cancer patients, the most frequently reported initial symptom was a breast lump (92.8%), while other symptoms, including redness, pain, and swelling, were observed in 7.2% of cases. With respect to tumor localization, 51.6% of patients had left-sided breast cancer, 46.4% had right-sided involvement, and 2.0% exhibited bilateral breast cancer.

Table 1: Symptoms and effected side of cases

| Characteristic | Frequency | Percentage (%) |
|---|---|---|
| Symptoms of BC | | |
| Lump | 142 | 92.8 |
| Others | 11 | 7.2 |
| Effected side | | |
| Left | 79 | 51.6 |
| Right | 73 | 46.4 |
| Bilateral | 03 | 2.0 |

Table 2 presents a summary of the descriptive statistics for breast cancer cases and controls. The findings showed that the mean age of women with breast cancer was 39 years, whereas healthy women had a mean age of 45 ± 11 years. Notably, the median and modal ages for cases were 38 and 30 years, respectively. A significant proportion of female breast cancer patients (majority) fell within the 30-40 year age range. Furthermore, the distribution of marital status was comparable between cases and controls, with a non-significant difference. Nearly, 97% patients and 95% health respondents were married, while 3% of cases and 5% controls were unmarried. Majority of cases and controls (72% and 58%) significantly belonged to urban area. Breastfeeding was significant higher both in cases and controls (64.7% and 76.5%) while bottle feeding was lower. Family history of cancer was significant risk factor with 40% cases and 27% controls having positive family history. Positive history of abortions was non-significantly lower in cases and control (33% and 32 %). Most of cases (95%) and controls (92.5%) have no previous history of ovarian removal. Ovary removal was non-significant (p=0.357) factor of breast cancer. Use of OCP's was significantly lesser for cases (37%) and controls (24%). Eighty percent of the healthy women and seventy-one percent of female patients never received radiation therapy. Most of the control group (96%) and patients (91%) never consumed drugs. Most of the respondents 92% of cases and 91% of controls—never habitually performed physical activity. The proportion of groups who have never worked nights is higher (75 percent of cases and 83 percent of controls).

Sixty-three percent of the patients were literate. Most control women (52%), though, were not literate. The results indicated that most of the female patients (71%) started menstruating comparatively at a young age (less than 12). On the contrary, most of the control women (55%) started menstruating at a comparatively older age (more than 13). Based on the results, most patients (84 percent) and control women (65 percent) continue to experience menstruation, which means that they have not reached menopause. The findings indicated that most respondent got married after the 18 years of age. Nevertheless, a few of the responders got married prior to eighteen. Almost 54% patients have 1-3 kids. Additionally, more than three children are possessed by most control responders (55 %).

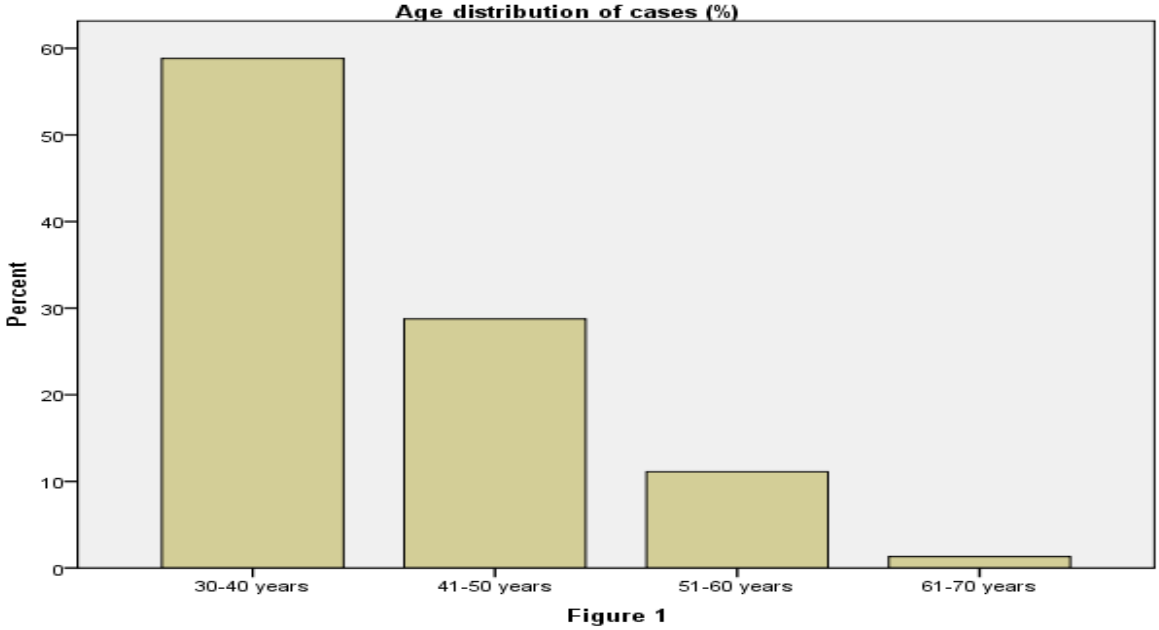

Figure 1

**Risk Factor Analysis**

The chi-squared test identified significant associations (p-value<0.05) between breast cancer and multiple variables, including urban residency, lower educational attainment, early menarche, menopausal status, fewer childbirths, lack of breastfeeding, family history of

breast cancer, oral contraceptive use, exposure to radiations, use of any drug, and night shift work.

Table 2: Basic Characteristics of cases and controls

| **Characteristics** | **Cases(n=153) n (%)** | **Controls(n=306) n (%)** | **p-value (Chi-squared test)** |
|---|---|---|---|
| Age at the time of diagnosis | | | 0.001 |
| 30-40 years | 90(59) | 104(34) | |
| 41-50 years | 44(29) | 105(34) | |
| 51-60 years | 17(11) | 70(23) | |
| 61-70 years | 02(1) | 27(9) | |
| Mean ± S.D | 39.7 ± 8.40 | 45.8 ± 10.5 | |
| Marital Status | | | 0.419 |
| Married | 148(97) | 291(95) | |
| Unmarried | 05(03) | 15(05) | |
| Type of area | | | 0.003 |
| Urban | 110 (72) | 177 (58) | |
| Rural | 43 (28) | 129(42) | |
| Breast feeding history | | | 0.008 |
| Positive | 99(64.7) | 234(76.5) | |
| Negative | 54(35.3) | 72(23.5) | |
| Family history | | | 0.006 |
| Positive | 61(40) | 83(27) | |
| Negative | 92(60) | 223(73) | |
| Abortion history | | | 0.888 |
| Positive | 50(33) | 98(32) | |
| Negative | 103(67) | 208(68) | |
| Ovary removal | | | 0.357 |
| Yes | 08(05) | 23(7.5) | |
| No | 145(95) | 283(92.5) | |
| Oral contraceptive pills | | | 0.004 |
| Used | 56(36.6) | 73(23.9) | |
| Never used | 97(63.4) | 233(76.1) | |
| Radiation treatment | | | 0.027 |
| Received | 44(29) | 60(19.6) | |
| Not received | 109(71) | 246(80.4) | |
| Drugs | | | 0.013 |
| Used | 14(9.2) | 11(3.6) | |
| Not used | 139(90.8) | 295(96.4) | |
| Physical activity | | | 0.907 |
| Yes | 13(8.5) | 27(8.8) | |
| No | 140(91.5) | 279(91.2) | |
| Night shift work | | | 0.039 |
| Yes | 39(25.5) | 53(17.3) | |
| No | 114(74.5) | 253(82.7) | |
| Education status | | | 0.001 |
| Illiterate | 31(20) | 158(51.6) | |
| Matric | 54(35) | 72(23.5) | |
| Graduation | 50(33) | 51(16.7) | |
| Masters | 18(12) | 25(8.2) | |
| Menstruation age | | | 0.001 |
| <12 years | 30(19.6) | 16(5.23) | |

| | | | |
|---|---|---|---|
| 12-13 years | 62(40.5) | 179(58.5) | |
| 14-15 years | 61(39.9) | 111(36.27) | |
| Menopausal status | | | 0.001 |
| Continued | 129(84) | 200(65) | |
| Up to 45 years | 09(06) | 27(09) | |
| Up to 50 years | 11(07) | 52(17) | |
| Up to 55 years | 04(03) | 27(09) | |
| Age at the time of wedding | | | 0.904 |
| Up to 18 years | 40(26.1) | 89(29.1) | |
| 19-25 years | 77(50.3) | 143(46.7) | |
| 26-30 years | 24(15.7) | 46(15) | |
| 31-40 years | 07(4.6) | 14(4.6) | |
| Never married | 05(3.3) | 14(4.6) | |
| Number of children | | | 0.008 |
| No children | 11(7.2) | 41(13.4) | |
| 1-3 children | 83(54.3) | 122(39.8) | |
| 4-6 children | 51(33.3) | 107(35) | |
| 7-9 children | 08(5.2) | 30(9.8) | |
| More than 9 children | 00 | 06(2) | |

**Logistic Regression Analysis**

The following significant predictors of breast cancer risk emerged from multivariate logistic regression analysis:

- **Urban residence** was linked to a nearly 2-fold elevated risk (OR=1.86, 95% CI: 1.23-2.83).
- **Non-breastfeeding mothers** exhibited an elevated risk (OR=1.77, 95% CI: 1.16-2.71).
- **Family history of breast cancer** increased the likelihood of developing breast cancer by 78% (OR=1.78, 95% CI: 1.18-2.69).
- **Oral contraceptive use** was linked to an 84% heightened risk (OR=1.84, 95% CI: 1.21-2.81).
- **Chest radiation exposure** was a significant risk factor (OR=1.66, 95% CI: 1.06-2.59).
- **Drug use** was associated with a nearly threefold increase in risk (OR=2.70, 95% CI: 1.20-6.10).
- **Night shift work** correlated with a 63% higher risk (OR=1.63, 95% CI: 1.02-2.61).

Table 3: Odd ratios and 95% C.I for significant risk factors

| Characteristics | O.R | CI (95%) |
|---|---|---|
| Age at the time of diagnosis | | |
| 30-40 years | 1.00 | |
| 41-50 years | 0.484 | 0.308-0.760 |
| 51-60 years | 0.281 | 0.154-0.512 |
| 61-70 years | 0.086 | 0.020-0.370 |
| Type of area | | |
| Rural | 1.00 | |
| Urban | 1.86 | 1.23-2.83 |
| Breast feeding history | | |
| Positive | 1.00 | |
| Negative | 1.77 | 1.16-2.71 |
| Family history | | |
| Negative | 1.00 | |

| | | |
|---|---|---|
| Positive | 1.781 | 1.182-2.685 |
| Oral contraceptive pills | | |
| Never used | 1.00 | |
| Used | 1.84 | 1.21-2.81 |
| Chest radiation treatment | | |
| Not received | 1.00 | |
| Received | 1.66 | 1.06-2.59 |
| Drugs | | |
| Not used | 1.00 | |
| Used | 2.70 | 1.20-6.10 |
| Night work | | |
| No | 1.00 | |
| Yes | 1.63 | 1.02-2.61 |
| Education status | | |
| Illiterate | 1.00 | |
| Up to Matric | 3.82 | 2.27-6.44 |
| Up to Graduation | 4.99 | 2.89-8.64 |
| Up to Masters | 3.67 | 1.79-7.52 |
| Menstruation age | | |
| < 12 years | 1.00 | |
| 12-13 years | 0.185 | 0.094-0.362 |
| 14-15 years | 0.293 | 0.148-0.580 |
| Menopausal status | | |
| Continued | 1.00 | |
| Up to 45 years | 0.517 | 0.235-1.134 |
| Up to 50 years | 0.328 | 0.165-0.652 |
| Up to 55 years | 0.230 | 0.079-0.672 |
| Number of children | | |
| No child | 1.00 | |
| 1-3 children | 2.54 | 1.23-5.22 |
| 4-6 children | 1.78 | 0.84-3.74 |
| 7-9 children | 1.00 | 0.36-2.77 |

**Discussion**

This study substantiates existing research linking reproductive and lifestyle factors with breast cancer risk. The observed association between early menarche and breast cancer aligns with literature indicating that prolonged estrogen exposure enhances susceptibility (Shin et al., 2011). Additionally, urban residency may be correlated with lifestyle factors that elevate risk, such as dietary patterns and reduced physical activity (Weir et al., 2007).

The findings reinforce the protective role of breastfeeding, in accordance with previous studies demonstrating that prolonged breastfeeding reduces breast cancer risk via hormonal modulation. The association between oral contraceptive use and radiation exposure underscores the importance of vigilant monitoring of medical treatments in high-risk populations.

**Conclusion and Proposed Interventions**

This study highlights multiple preventable and modifiable determinants of Breast Cancer, emphasizing the need for targeted public health interventions. Healthcare professionals should educate women about risk reduction strategies, including breastfeeding promotion, lifestyle modifications, and careful monitoring of hormonal contraceptive use and radiation

exposure. Future research should investigate gene-environment interactions to refine risk assessment models further.

**References**


Breast cancer and hormonal contraceptives: further results. Collaborative Group on Hormonal Factors in Breast Cancer. (1996). *Contraception*, *54*(3 Suppl), 1S–106S. https://doi.org/10.1016/s0010-7824(15)30002-0

Center, C. American Cancer Society Inc.(2013). *Cancer facts and figures 2013*.

Collaborative Group on Hormonal Factors in Breast Cancer (2001). Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet (London, England)*, 358(9291), 1389–1399. https://doi.org/10.1016/S0140-6736(01)06524-2

Collaborative Group on Hormonal Factors in Breast Cancer (2002). Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet (London, England)*, *360*(9328), 187–195. https://doi.org/10.1016/S0140-6736(02)09454-0

Collaborative Group on Hormonal Factors in Breast Cancer. 2002. Alcohol, tobacco andbreast cancer – collaborative reanalysis of individual data from 53epidemiological studies, including 58 515 women with breast cancer and 95 067 women without the disease. Br J Cancer, 87(11): 1234–1245.

Fatima, N., Zaman, M. U., Maqbool, A., Khan, S. H., & Riaz, N. (2013). Lower incidence but more aggressive behavior of right sided breast cancer in Pakistani women: does right deserve more respect?. *Asian Pacific Journal of Cancer Prevention*, *14*(1), 43-45.

Ferlay, J. (2010). GLOBOCAN 2008 v2. 0, Cancer incidence and mortality worldwide: IARC CancerBase No. 10. *http://globocan. iarc. fr*.

Khokher, S., Qureshi, M. U., Riaz, M., Akhtar, N., & Saleem, A. (2012). Clinicopathologic profile of breast cancer patients in Pakistan: ten years data of a local cancer hospital. *Asian Pacific Journal of Cancer Prevention*, *13*(2), 693-698.

Kotsopoulos, J., Lubinski, J., Salmena, L., Lynch, H. T., Kim-Sing, C., Foulkes, W. D., ... & Hereditary Breast Cancer Clinical Study Group. (2012). Breastfeeding and the risk of breast cancer in BRCA1 and BRCA2 mutation carriers. *Breast cancer research*, *14*, 1-7.

Mailk, A. A., Wani, K. A., & Ahmad, S. R. (2012). Breast conservative therapy. *JMS SKIMS*, *15*(1), 7-14.

Shin, A., Song, Y. M., Yoo, K. Y., & Sung, J. (2011). Menstrual factors and cancer risk among Korean women. *International journal of epidemiology*, *40*(5), 1261-1268.

Sohail, S., & Alam, S. N. (2007). Breast cancer in Pakistan-awareness and early detection.

Weir, R., Day, P., & Ali, W. (2007). Risk factors for breast cancer in women. *New Zealand Health Technology Assessment (NZHTA) Report*, *10*(2), 1-328.

Zienolddiny, S., Haugen, A., Lie, J. A. S., Kjuus, H., Anmarkrud, K. H., & Kjærheim, K. (2013). Analysis of polymorphisms in the circadian-related genes and breast cancer risk in Norwegian nurses working night shifts. *Breast Cancer Research*, *15*, 1-16.