International Journal of Basic and Applied Medical Sciences ISSN: 2277-2103
Online, International Journal, Available at http://www.cibtech.org/jms.htm
2025 Vol.15, pp. 19-43/Gurinder et al.
Research Article (Open Access)

BREAST CANCER IN MAJHA REGION, PUNJAB: AN EPIDEMIOLOGICAL ANALYSIS AND PUBLIC HEALTH IMPLICATIONS

Gurinder Kaur Walia, Diksha Chopra, Sonia, Lipsa Dehal*
Department of Zoology and Environmental Sciences, Punjabi University, Patiala
**Author for Correspondence: lipsa.dehal999@gmail.com*

ABSTRACT
Breast cancer continues to be the most common disease diagnosed in women and a major cause of cancer-related deaths globally, with regional variations caused by environmental, lifestyle, reproductive, and demographic variables. The prevalence is gradually increasing in India as well as in Punjab, especially the Majha region reporting alarming patterns. The purpose of this study was to investigate how lifestyle choices, environmental exposures, anthropometric measurements, and reproductive history interact to influence risk factors for breast cancer and survival rates in women living in Majha region of Punjab. Data on age, domicile, menopausal status, smoking, dietary habits, age at menarche, parity, and exposure history were collected from 117 women by cross-sectional questionnaire. Additionally, physical measures such as hip and waist circumferences were also recorded. Descriptive summary, chi-square tests, ANOVA, correlation analysis, and logistic regression were used for the statistical studies. The participants' average age was 54.5 years, indicating that post-menopausal women in the majority. The findings revealed no significant correlation between menopausal state and weight ($F=0.42$, $p=0.517$) or smoking ($\chi^2=1.97$, $p=0.578$). However, a moderately significant connection ($r=0.42$) between hip circumference and waist circumference was noted, underscoring the importance of abdominal obesity as a possible risk factor. Although robust regression analysis was limited by incomplete data, descriptive observations highlighted the potential influence of early menarche, age at first childbirth, increased parity, and environmental exposures on breast cancer vulnerability and survival outcomes. Overall, the results highlight the significance of thorough epidemiological studies and offer a regional perspective on Punjab's breast cancer risk trends. Despite its shortcomings, the study emphasizes the necessity of more extensive, thorough studies to confirm correlations and lead early detection, prevention, and region-specific approaches for the control of breast cancer in northern India.

Keywords: *Breast Cancer, Epidemiology, Punjab, Lifestyle, Survival Outcome*

INTRODUCTION
Cancer is a type of tumour that demonstrates abnormal cell growth with the potential to invade or spread from its originating organ (site) to other parts of the body (Arnold *et al.*, 2022; Desai *et al.*, 2024). Genetic and epigenetic changes that interfere with important cellular processes including DNA damage and repair, proliferation control, and immune destruction avoidance give tumors the ability to infiltrate or spread by developing various malignant phenotypes (Desai *et al*., 2024).
The breasts are a pair of glands situated above the pectoralis major muscle that vary in size and density. Parashar *et al*. (2023) explained that breast cancer usually starts in ductal carcinoma, but it can also develop in the breast lobules, a condition known as lobular carcinoma. This kind of cancer is tough to cure since it can spread to different body areas such the brain, liver, lungs, and bones (Fernandes *et al*., 2023; Mihai *et al*., 2024). The second most common cause of cancer-related fatalities among women worldwide is the breast cancer, which is also the most commonly diagnosed cancer in women (Hortobagyi *et al*., 2005; Arnold *et al.*, 2022). The most recent global figures show that 2.3 million new

instances of breast cancer occur every year in both genders and account for around 12.5% of all cancer diagnoses globally (Desai *et al*., 2024; Liao, 2024). Globally, the incidence has been increasing, and estimates suggest that it will continue until 2050 (Deng *et al*., 2025). Breast cancer killed about 685,000 women in 2020, making it the disease responsible for 16%, or 1 in 6 cancer-related deaths among women (Desai *et al*., 2024). Nonetheless, early detection raises survival rates and greatly increases the likelihood of a successful course of therapy (Gamal *et al*., 2024; Mihai *et al*., 2024; Rahman *et al*., 2024; Kumar *et al*., 2024). Early breast cancer stages I and II have five-year survival rates between 80 to 99%, however stage IV, which is an advanced stage with metastases, may have a survival rate of less than 30% (Fernandes *et al*., 2023).

From fourth on the list in the 1990s to first now, breast cancer has become the most frequent cancer among women in India. With more than 70,000 fatalities from breast cancer each year, India bears over one-third of the global burden (Gaitonde and Gajbhiye, 2022). Geographically and throughout various socioeconomic strata, the disease is growing more common (Mehrotra and Yadav, 2022; Kulothungan *et al*., 2024). With a mean diagnostic age of 53.66 years and notable differences in risk variables and clinical presentations, studies from several Indian states have shown clear regional trends (Gaitonde and Gajbhiye, 2022).

Punjab, an agricultural province in northwest India, the rise of cancer has been showing very alarming patterns. Majha, Malwa, and Doaba are the three main, separate regions that make up the state. Four districts (Amritsar, Tarn Taran, Gurdaspur, and Pathankot) make up the Majha region, which is a distinct geographic and cultural location with particular demographic traits and possible environmental factors (Kaur *et al*., 2021). The national average of 80 cancer cases per lakh is surpassed by the highest incidence of 136 cancer cases per 100,000 persons in the Malwa region (Nanda *et al*., 2016; Rana *et al*., 2021). There are a lot of incidences of breast cancer each year in Amritsar, Punjab's second-largest town. In southern Punjab, women made up 60.9% of the 1,328 cancer cases, with breast cancer accounting for the majority of these instances in women (35.7%). Moreover, in the different North Indian tertiary care survey, breast cancer accounted for 24.7% of all female cancer cases (Rana *et al*., 2021). Studies from Punjab have also identified specific risk patterns. Amritsar revealed that obesity and parity were significant risk factors, with different patterns in premenopausal and postmenopausal women as premenopausal women with 3 or fewer children had a higher risk (OR=5.54) compared to those with more than 3 children (Kour *et al*., 2019).

Numerous variables have been identified as contributing to the high prevalence of cancer in some areas of Punjab. The Malwa region or, cotton belt has seen a sharp increase in cancer cases, perhaps as a result of excessive pesticide use. The high rate of cancer in these regions is believed to be caused by environmental pollutants in the macro environment. The increasing burden may also be influenced by food habits, lifestyle choices, and socioeconomic shifts (Kour *et al*., 2019). Furthermore, understanding the epidemiological patterns in different regions can contribute to the broader knowledge of breast cancer etiology and help identify modifiable risk factors that are particularly relevant to specific populations. In light of this, it is clear that thorough regional epidemiological studies are required. With its distinct geographic position, farming methods, and demographic traits, Punjab's Majha region is a crucial place for in-depth epidemiological research. By bridging the gap between local health requirements and worldwide epidemiological trends, these research help improve breast cancer early detection, mitigation, and management options for this group.

MATERIALS AND METHODS

2.1 Study Design and Population

This study utilized a cross-sectional design to investigate the epidemiological patterns of breast cancer risk factors in the Majha region of Punjab, India (Figure 1). The study was designed to investigate demographic, lifestyle, reproductive and environmental risk factors associated with breast cancer

outcomes. Data was collected from 117 female participants who were selected using a convenience sampling method. The participants were recruited from local health centres, community outreach programs and other medical facilities within the region.

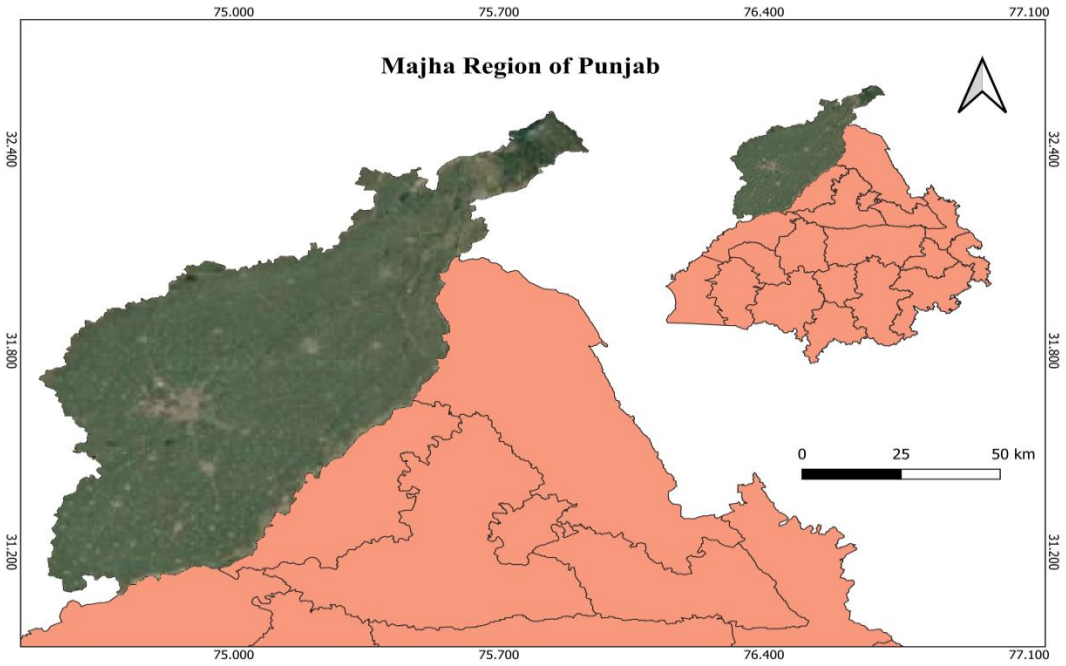

Figure 1: Map Showing Majha Region of Punjab

2.1.1 Inclusion criteria:

- Female participants aged 18 years and above.
- Residents of the Majha region of Punjab.
- No prior diagnosis of breast cancer at the time of data collection.

2.1.2 Exclusion criteria:

- Participants with incomplete or inconsistent data.
- Individuals with a history of breast cancer or other related comorbidities.

2.2 Ethical Considerations

The study was approved by the Institutional Review Board (IRB) of the participating health institutions. Informed consent was obtained from all participants before data collection. The purpose of the study was explained to participants, and confidentiality of the data was ensured by anonymizing responses. Participation was voluntary and participants could withdraw at any point during the study.

2.3 Data Collection

The data was collected over a period of three months, with interviews, surveys and questionnaire conducted in local clinics, health centres and community-based settings.

Questionnaire Structure

Section A: Demographic Details

1. Patient ID: _____
2. Age (in years): _____
3. Sex: ☐ Female ☐ Male
4. Residence: ☐ Urban ☐ Rural

Section B: Lifestyle and Personal Habits

5. Type of diet: ☐ Vegetarian ☐ non-vegetarian ☐ Mixed

6. Alcohol consumption: ☐ Never ☐ Occasionally ☐ Regularly
7. Smoking/tobacco use: ☐ Never ☐ Occasionally ☐ Regularly

Section C: Reproductive History (For female patients only)

8. Age at menarche: _____ years
9. Menopausal status: ☐ Premenopausal ☐ Postmenopausal
10. Number of full-term pregnancies: _____
11. Age at first childbirth: _____ years
12. Number of abortions (spontaneous or induced): _____

Section D: Anthropometric Measurements

13. Height: _____ cm
14. Weight: _____ kg
15. Waist circumference: _____ cm
16. Hip circumference: _____ cm

Note: Body Mass Index (BMI) and Waist-Hip Ratio (WHR) were calculated using the following formulas:

- BMI = Weight (kg) / [Height (m)]²
- WHR = Waist Circumference (cm) / Hip Circumference (cm)

Section E: Environmental and Occupational Exposure

17. History of exposure to pesticides, chemicals, radiation, or other toxins: ☐Yes ☐No
    If Yes, specify: _____
    Duration of exposure: _____ years

Section F: Clinical and Treatment Details

18. Side affected by cancer: ☐ Left ☐ Right ☐ Both

The data collection process involved the following key components:

1.3.1 Demographic and Lifestyle Information: Participants were asked to complete a structured questionnaire that gathered information on age, sex, residence, menopausal status, diet type, smoking status, alcohol consumption and exposure to environmental pollutants.

1.3.2 Health History: Data on reproductive history, including full-term pregnancies, age at first child, number of abortions and age at menarche.

1.3.3 Physical Measurements: Height was measured using a stadiometer, weight was measured using a calibrated digital scale and waist circumference and hip circumference was measured using a non-stretchable tape measure.

1.4 Variables

2.4.1 Independent Variables:

- Age (continuous): Participant's age in years.
- Menopausal status (categorical): Classified as PRE-menopausal and POST-menopausal.
- Exposure (categorical): Exposure to environmental factors such as dust, pesticides, radiation, etc.
- Diet type (categorical): Whether the participant followed a vegetarian or non-vegetarian diet.
- Smoking (binary): Whether the participant was a smoker (Yes/No).

2.4.2 Dependent Variable:

- Smoking status (binary): A categorical variable indicating whether the participant smokes (Yes/No).

2.5 Statistical Analysis

Data was analyzed using the Python programming language (with pandas for data manipulation and statsmodels for statistical analysis). The following statistical methods were employed:

2.5.1 Descriptive Statistics: Basic summary statistics (mean, median, standard deviation, range) were computed for continuous variables (e.g., Age, Weight, Waist circumference, Hip circumference).
2.5.2. Chi-Square Test of Independence: This test was used to assess the relationship between categorical variables (e.g., Menopausal status and Smoking). The null hypothesis assumed no association between the variables.
2.5.3. ANOVA (Analysis of Variance): ANOVA was used to compare the means of Weight across different Menopausal status groups (PRE and POST). A p-value of less than 0.05 was considered statistically significant.
2.5.4 Correlation Analysis: Pearson correlation coefficients were calculated to determine the strength and direction of the relationship between continuous variables, such as Weight and Waist circumference, Age and Waist circumference, and others.
2.5.5 Logistic Regression: A binary logistic regression model was developed to assess the influence of independent variables (e.g., Age, Waist circumference, Menopausal status, etc.) on the dependent variable (Smoking status). The logistic regression was run using statsmodels, with odds ratios (OR) calculated for each predictor.
2.5.6 Statistical Significance: All statistical tests were performed at the 95% confidence level ($\alpha = 0.05$). A p-value < 0.05 indicated statistical significance.

RESULTS

3.1 Descriptive Statistics

The dataset comprised 117 female participants and their details summarized in Table 1, with the following key demographics and health characteristics:

Table 1. Summary of Characteristics of the Study Population (n = 117)

Patient ID	Age	Sex	Residance	Type of diet	Alcohol Consumption	Age at menarche	Menopausal status	Full term Pregnancies	Age at 1st child	No. of Abortion	Height (survival)	Weight	Waist (circumference)	Hip (circumference)	Exposure	Location	Smoking
C13	32	F	URBAN	NON-VEG	NO	12	PRE	2	26	IN	5-5'4''	45	31	36	DUST	LEFT	NO
C14	38	F	URBAN	VEG	NO	13	POST	2	25	IN	5-5'4''	47	32	38	PESTICIDE	RIGHT	NO
C28	24	F	RURAL	VEG	NO	13	PRE	2	23	SP	>5'0"	45	33	40	PESTICIDE	LEFT	NO
C43	39	F	RUR	VEG	NO	13	POST	3	16	SP	>5'0"	48	34	42	DUST	RIGH	NO

			AL													T	
C51	37	F	RURAL	VEG	NO	13	PRE	3	17	SP	>5'0"	49	34	36	DUST	RIGHT	NO
C57	76	F	RURAL	VEG	NO	13	PRE	2	27	SP	>5'0"	55	33	38	DUST	RIGHT	NO
C64	55	F	RURAL	VEG	NO	13	POST	2	18	IN	>5'0"	57	32	40	DUST	LEFT	NO
C209	23	F	RURAL	VEG	NO	14	PRE	2	23	SP	>5'0"	56	31	42	PESTICIDE	RIGHT	NO
C211	27	F	URBAN	VEG	NO	14	POST	4	21	SP	>5'0"	59	40	44	DUST, PESTICIDE	RIGHT	NO
C214	24	F	URBAN	NON-VEG	NO	14	PRE	4	19	IN	>5'0"	58	42	46	PESTICIDE		NO
C242	24	F	URBAN	VEG	NO	14	POST	4	23	SP	5'-5'4''	60	44	48	DUST, PESTICIDE	RIGHT	NO
C273	29	F	URBAN	VEG	NO	14	PRE	3	17	IN	5'-5'4''	56	42	42	DUST, PESTICIDE	RIGHT	NO
C275	28	F	URB	VEG	NO	13	PRE	2	21	SP	5'-5'4''	46	40	40	DUST	LEFT	NO

			AN														
C279	27	F	RURAL	NON-VEG	NO	14	POST	2	24	IN	5'-5'4''	45	41	40	FERTILIZER	RIGHT	NO
C355	56	F	URBAN	VEG	NO	12	PRE	3	20	SP	>5'0"	48	42	40	DUST	RIGHT	NO
C139	43	F	URBAN	NON-VEG	NO	12	POST	4	19	IN	>5'0"	49	35	42	FERTILIZER	LEFT	NO
C78	60	F	RURAL	VEG	NO	12	PRE	2	22	SP	>5'0"	50	38	44	DUST	LEFT	NO
C213	25	F	URBAN	VEG	NO	12	POST	5	25	IN	>5'0"	51	38	46	FERTILIZER	RIGHT	NO
C138	50	F	SUB-URBAN	NON-VEG	NO	12	PRE	7	24	SP	5'-5'4''	55	38	50	DUST	LEFT	NO
UAT213	47	F	RURAL	VEG	NO	16	POST	4	22	IN	5'-5'4''	54	40	52	FERTILIZER	RIGHT	NO
C355	56	F	URBAN	VEG	NO	16	PRE	3	20	SP	5'-5'4''	53	44	45	DUST	RIGHT	NO
UAT12	56	F	RURA	VEG	NO	16	POST	2	25	IN	5'-5'4''	52	43	46	DUST,PESTICIDE		NO

5			L														
C 63	70	F	RURAL	VEG	NO	16	PRE	2	20	SP	5'-5'4''	52	52	48	DUST	RIGHT	NO
UA 20	65	F	RURAL	VEG	NO	16	POST	2	17	IN	5'-5'4''	53	44	44	FERTILIZER	RIGHT	NO
UAT 135	58	F	RURAL	VEG	NO	16	PRE	2	25	SP	5'-5'4''	54	36	42	DUST	LEFT	NO
C 68	45	F	RURAL	VEG	NO	16	POST	2	26	IN	5'-5'4''	55	38	46	MOBILE	LEFT	NO
UA 79	56	F	RURAL	VEG	NO	16	PRE	5	19	SP	5'-5'4''	51	28	42	DUST	RIGHT	NO
C 106	32	F	URBAN	VEG	NO	15	POST	4	20	IN	5'-5'4''	51	26	42	MOBILE	RIGHT	NO
C 330	52	F	RURAL	VEG	NO	15	PRE	6	19	SP	5'-5'4''	54	24	36	DUST	RIGHT	NO
C 174	60	F	RURAL	NON-VEG	NO	15	POST	3	18	IN	5'-5'4''	53	24	38	MOBILE	LEFT	NO
C 72	70	F	RURAL	VEG	NO	15	PRE	3	24	IN	5'-5'4''	50	24	40	DUST	LEFT	NO
C 12	53	F	UR	NO	NO	15	POST	2	26	SP	5'-5'4'	56	28	42	MOBIL	RIG	NO

2			BAN	N-VEG							'				E	HT	
C50	34	F	RURAL	NON-VEG	NO	14	PRE	2	18	SP	<5'4"	57	30	34	DUST	RIGHT	NO
UA14	60	F	RURAL	VEG	NO	14	POST	4	20	IN	<5'4"	58	30	36	MOBILE	RIGHT	NO
C128	45	F	URBAN	NON-VEG	NO	14	PRE	5	22	SP	<5'4"	51	30	40	MOBILE	RIGHT	NO
C91	46	F	RURAL	VEG	NO	12	POST	3	23	IN	<5'4"	52	26	24	MOBILE	LEFT	NO
UAT120	25	F	RURAL	NON-VEG	NO	13	PRE	6	24	SP	<5'4"	54	24	25	DUST, PESTICIDE	RIGHT	NO
UAT187	30	F	RURAL	VEG	NO	14	POST	6	20	IN	<5'4"	65	24	26	MOBILE	RIGHT	NO
C165	62	F	URBAN	VEG	NO	15	PRE	7	21	SP	<5'4"	66	28	27	DUST, PESTICIDE	RIGHT	CHULLAH
C132	50	F	RURAL	VEG	NO	16	PRE	3	21	IN	>5'0"	70	30	28	DUST, PESTICIDE	LEFT	NO
C19	65	F	RU	VEG	NO	12	POST	2	23	IN	>5'0"	45	30	29	DUST,	LEF	NO

9			RAL												PESTICIDE	T	
UA13	65	F	SUB-URBAN	VEG	NO	12	POST	2	22	SP	>5'0"	46	28	30	MOBILE	LEFT	NO
C341	50	F	RURAL	NON-VEG	NO	13	POST	2	21	SP	>5'0"	47	24	32	DUST, PESTICIDE	LEFT	NO
C79	45	F	RURAL	VEG	NO	14	POST	2	18	IN	>5'0"	48	24	34	MOBILE	RIGHT	NO
UAT148	45	F	RURAL	VEG	NO	14	POST	3	24	SP	<5'4"	49	26	24	DUST, PESTICIDE	RIGHT	NO
C53	35	F	RURAL	VEG	NO	14	POST	3	30	SP	<5'4"	61	28	28	RADIATION	LEFT	NO
C328	60	F	RURAL	VEG	NO	13	POST	3	30	SP	<5'4"	62	30	26	DUST, PESTICIDE	LEFT	NO
UA78	64	F	URBAN	VEG	NO	12	POST	4	21	SP	<5'4"	63	24	28	DUST, PESTICIDE	RIGHT	NO
C342	62	F	RUR	VEG	NO	16	POST	4	24	SP	<5'4"	64	26	25	RADIATIO	RIGH	NO

			AL												N	T	
C353	50	F	RURAL	NON-VEG	NO	12	POST	4	22	SP	>5'0"	65	28	24	RADIATION	LEFT	NO
UAT215	60	F	RURAL	VEG	NO	13	POST	5	27	SP	>5'0"	63	30	40	RADIATION	LEFT	NO
UAT124	50	F	URBAN	VEG	NO	13	POST	5	19	SP	>5'0"	64	24	42	DUST, PESTICIDE	RIGHT	NO
UAT192	50	F	RURAL	VEG	NO	14	POST	3	20	SP	5'-5'4''	64	26	38	RADIATION	RIGHT	NO
C93	66	F	URBAN	VEG	NO	15	POST	3	19	SP	5'-5'4''	62	28	36	DUST, PESTICIDE	LEFT	NO
C77	63	F	RURAL	VEG	NO	15	POST	4	18	SP	5'-5'4''	61	30	40	RADIATION	LEFT	NO
C148	47	F	URBAN	VEG	NO	15	POST	4	24	SP	5'-5'4''	65	26	42	RADIATION	RIGHT	NO
UAT199	30	F	RURAL	NON-VEG	NO	13	POST	3	28	SP	<5'4"	63	34	28	RADIATION		NO
C147	44	F	URBA	NON-VE	NO	16	POST	2	20	SP	<5'4"	62	33	30	DUST	LEFT	NO

			N	G													
UA13	65	F	RURAL	VEG	NO	12	POST	2	22	SP	<5'4"	45	32	28	RADIATION	LEFT	NO
C124	54	F	URBAN	NON-VEG	NO	13	POST	2	24	SP	<5'4"	46	32	24	DUST		NO
C319	70	F	RURAL	NON-VEG	NO	13	POST	2	23	SP	<5'4"	47	31	26	DUST		NO
C326	63	F	RURAL	NON-VEG	NO	DON'TKNOW	POST	4	20	SP	5'-5'4''	48	34	24	DUST	LEFT	NO
C346	69	F	URBAN	VEG	NO	14	POST	3	21	SP	5'-5'4''	49	34	28	DUST		NO
C68	45	F	RURAL	VEG	NO	DON'TKNOW	POST	3	21	SP	5'-5'4''	60	34	24		LEFT	NO
UA30	54	F	URBAN	VEG	NO	DON'TKNOW	POST	3	22	SP	5'-5'4''	65	34	26			NO
PB101	43	F	URBAN	VEG	NO	12	POST	4	31	SP	<5'4"	63	35	26	DUST	LEFT	CHULAH
PB106	60	F	URBA	VEG	NO	13	PRE	4	20	IN	5'0''-5'4''	55	31	40	PESTICIDE	RIGHT	NO

			N														
PB105	25	F	RURAL	VEG	NO	13	POST	3	19	IN	5'0''-5'4''	56	34	42	PESTICIDE	RIGHT	NO
PB213	50	F	RURAL	NON-VEG	NO	13	PRE	5	18	SP	5'0''-5'4''	57	32	44	PESTICIDE	RIGHT	NO
PB418	48	F	RURAL	NON-VEG	NO	13	PRE	2	24	SP	5'0''-5'4''	58	33	46	PESTICIDE	RIGHT	NO
PB453	56	F	URBAN	VEG	NO	14	POST	2	26	SP	5'0''-5'4''	59	34	48	PESTICIDE	RIGHT	NO
PB458	56	F	URBAN	VEG	NO	14	POST	2	18	IN	5'0''-5'4''	60	32	48	PESTICIDE	LEFT	CHULLAH
PB432	70	F	RURAL	VEG	NO	12	POST	2	20	IN	5'0''-5'4''	60	32	50	PESTICIDE	RIGHT	NO
PB503	65	F	RURAL	VEG	NO	15	POST	3	22	SP	5'0''-5'4''	56	34	52	FERTILIZER	LEFT	NO
PB508	58	F	URBAN	NON-VEG	NO	16	PRE	4	26	SP	5'0''-5'4''	57	31	55	MOBILE	RIGHT	NO
PB523	45	F	RURAL	VEG	NO	13	POST	4	27	IN	5'0''-5'4''	58	34	54	FERTILIZER	RIGHT	NO
PB	58	F	RU	VEG	NO	12	PRE	4	20	SP	>5'0''	50	40	56	MOBIL	LEF	NO

520			RAL												E	T	
PB536	32	F	RURAL	VEG	NO	13	POST	3	21	SP	>5'0"	51	42	44	FERTILIZER	RIGHT	NO
PB537	52	F	URBAN	VEG	NO	16	PRE	4	21	IN	>5'0"	45	41	48	MOBILE	RIGHT	NO
PB549	60	F	RURAL	VEG	NO	16	POST	7	24	SP	>5'0"	45	40	50	DUST	RIGHT	NO
PB557	70	F	RURAL	NON-VEG	NO	15	PRE	3	22	SP	>5'0"	46	24	50	FERTILIZER	LEFT	NO
PB564	53	F	URBAN	VEG	NO	13	POST	7	18	IN	5'0'-5'4''	47	26	26	MOBILE	RIGHT	NO
PB563	34	F	SUB-URBAN	VEG	NO	12	PRE	2	24	IN	5'0'-5'4''	48	27	28	DUST	RIGHT	NO
PB566	58	F	URBAN	VEG	NO	14	POST	3	30	SP	5'0'-5'4''	49	28	30	FERTILIZER	LEFT	NO
PB567	45	F	URBAN	NON-VEG	NO	15	PRE	3	27	IN	5'0'-5'4''	53	30	32	MOBILE	RIGHT	NO
PB56	46	F	RUR	VEG	NO	15	POST	4	21	SP	5'0'-5'4'	53	30	30	FERTILIZ	LEFT	NO

9			AL								'				ER		
PB574	25	F	RURAL	VEG	NO	12	POST	4	24	SP	5'0''-5'4''	54	28	24	DUST	RIGHT	NO
PB478	30	F	RURAL	NON-VEG	NO	13	POST	3	22	SP	5'0''-5'4''	55	28	26	FERTILIZER	RIGHT	NO
PB215	62	F	RURAL	VEG	NO	14	PRE	2	27	IN	5'0''-5'4''	54	24	28	MOBILE	LEFT	NO
PB240	50	F	RURAL	NON-VEG	NO	15	POST	3	26	SP	5'0''-5'4''	53	26	30	FERTILIZER	RIGHT	NO
PB250	45	F	RURAL	NON-VEG	NO	12	PRE	4	25	SP	5'0''-5'4''	52	30	22	DUST	RIGHT	NO
PB259	65	F	RURAL	NON-VEG	NO	13	POST	2	23	SP	5'0''-5'4''	54	30	24	FERTILIZER	LEFT	NO
PB249	65	F	RURAL	VEG	NO	14	PRE	2	16	SP	5'0''-5'4''	55	24	26	MOBILE	LEFT	NO
PB240PB92	52	F	RURAL	NON-VEG	NO	15	POST	3	27	SP	>5'0"	51	26	28	MOBILE	RIGHT	PASSIVE-SMOKER
PB211	45	F	RURA	VEG	NO	16	PRE	3	18	IN	>5'0"	50	30	30	DUST	LEFT	NO

			L														
PB507	45	F	RURAL	VEG	NO	12	POST	2	22	SP	>5'0"	50	48	34	MOBILE	RIGHT	NO
PB510	35	F	RURAL	VEG	NO	13	PRE	3	21	SP	<5'4"	53	46	34	MOBILE	RIGHT	NO
PB518	58	F	RURAL	NON-VEG	NO	14	POST	6	19	IN	<5'4"	55	44	34	DUST	RIGHT	NO
PB552	65	F	URBAN	NON-VEG	NO	15	PRE	3	26	SP	<5'4"	45	42	50	MOBILE	RIGHT	NO
PB559	62	F	RURAL	VEG	NO	12	POST	5	21	SP	<5'4"	46	46	36	RADIATION	LEFT	NO
PB394	50	F	RURAL	VEG	NO	13	PRE	2	24	IN	<5'4"	47	48	38	RADIATION	RIGHT	NO
PB398	61	F	RURAL	VEG	NO	13	PRE	4	20	SP	<5'4"	48	42	40	RADIATION	RIGHT	NO
PB401	50	F	RURAL	VEG	NO	14	POST	3	19	SP	<5'4"	49	32	40	DUST	RIGHT	PASSIVE-SMOKER
PB405	50	F	URBAN	VEG	NO	15	PRE	2	22	SP	5'0''-5'4''	61	34	38	RADIATION	LEFT	NO

PB407	65	F	RURAL	NON-VEG	NO	15	POST	3	25	IN	5'0''-5'4''	65	32	36	RADIATION	RIGHT	NO
PB408	65	F	URBAN	VEG	NO	15	PRE	2	24	SP	5'0''-5'4''	63	31	40	DUST, PESTICIDE	RIGHT	NO
PB417	50	F	RURAL	VEG	NO	16	POST	3	22	SP	5'0''-5'4''	64	34	34	DUST	LEFT	NO
PB425	30	F	RURAL	VEG	NO	14	PRE	2	20	IN	>5'0"	64	33	36	DUST, PESTICIDE	RIGHT	NO
PB427	44	F	RURAL	NON-VEG	NO	14	POST	3	25	SP	>5'0"	63	31	38	DUST, PESTICIDE	RIGHT	NO
PB434	65	F	URBAN	VEG	NO	14	PRE	3	20	SP	>5'0"	64	32	40	DUST, PESTICIDE	LEFT	NO
PB448	50	F	RURAL	VEG	NO	14	POST	2	26	SP	<5'4"	65	34	38	DUST, PESTICIDE	RIGHT	NO
PB449	70	F	URBAN	VEG	NO	13	PRE	3	27	SP	<5'4"	65	36	36	DUST, PESTICIDE	RIGHT	NO
PB	62	F	UR	NO	NO	12	PRE	3	17	IN	<5'4"	67	38	38	DUST,	LEF	NO

92			BAN	N-VEG											PESTICIDE	T	
PB301	61	F	URBAN	NON-VEG	NO	12	POST	3	18	SP	>5'0"	49	40	40	DUST,PESTICIDE		NO
PB245	64	F	URBAN	VEG	NO	14	POST	3	19	SP	>5'0"	48	42	34	DUST,PESTICIDE		NO
PB608	66	F	URBAN	NON-VEG	NO	15	POST	3	23	SP	>5'0"	47	44	34	DUST, PESTICIDE	LEFT	NO
PB234	67	F	RURAL	VEG	NO	16	POST	4	24	IN	>5'0"	46	42	32	DUST	LEFT	PASSIVE SMOKER

3.1.1 Mean Age (Figure 2): The participants had a mean age of 54.5 years, with a standard deviation of 6.6 years. The age range spanned from 45 years to 70 years, representing a broad spectrum of adult women, which is important when considering age as a potential risk factor for breast cancer. Age is often associated with increased cancer risk, particularly post-menopause.

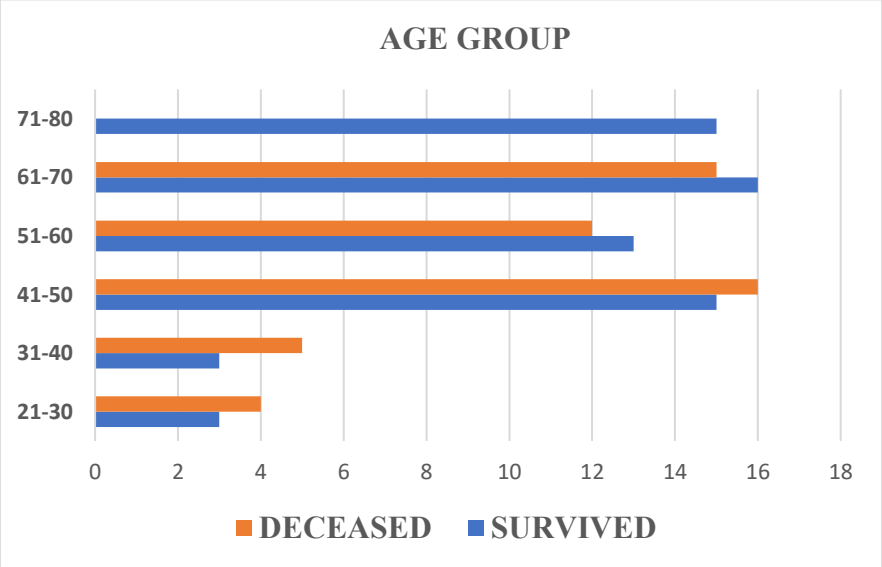

Figure 2: Age specific survival rate among breast cancer patients

3.1.2 Body Measurements:

3.1.2.1 Weight (Figure 3): The mean weight of participants was 54.55 kg (SD = 6.59 kg), with values ranging from 45 kg to 70 kg. Weight is an important factor in epidemiological studies, as being overweight or obese is a known risk factor for breast cancer.

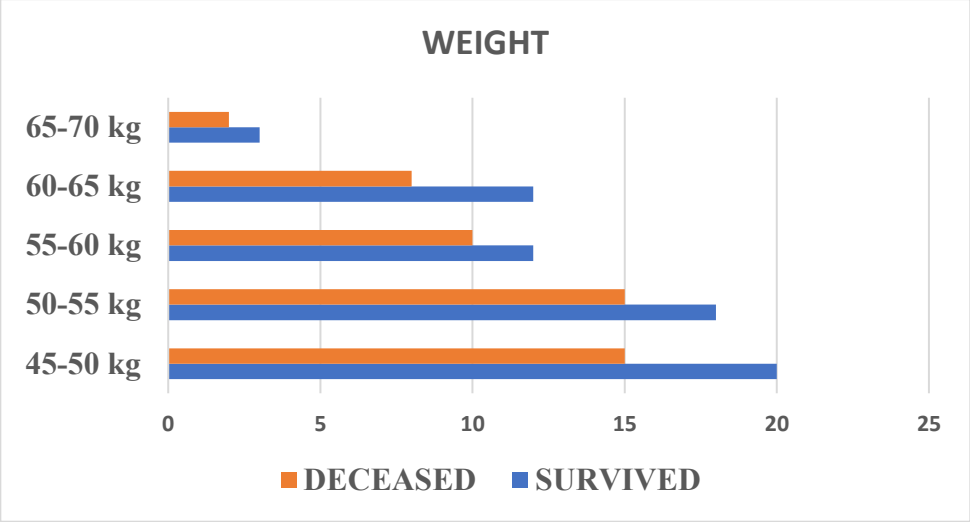

Figure 3: Weight-specific survival rate among breast cancer patients

3.1.2.2 Waist Circumference (Figure 4): The mean waist circumference was 33.25 cm, suggesting a relatively moderate level of abdominal fat in the population. The waist-to-hip ratio is often used to assess risk for diseases, including cancer, and this value may provide insight into fat distribution that could influence cancer risk.

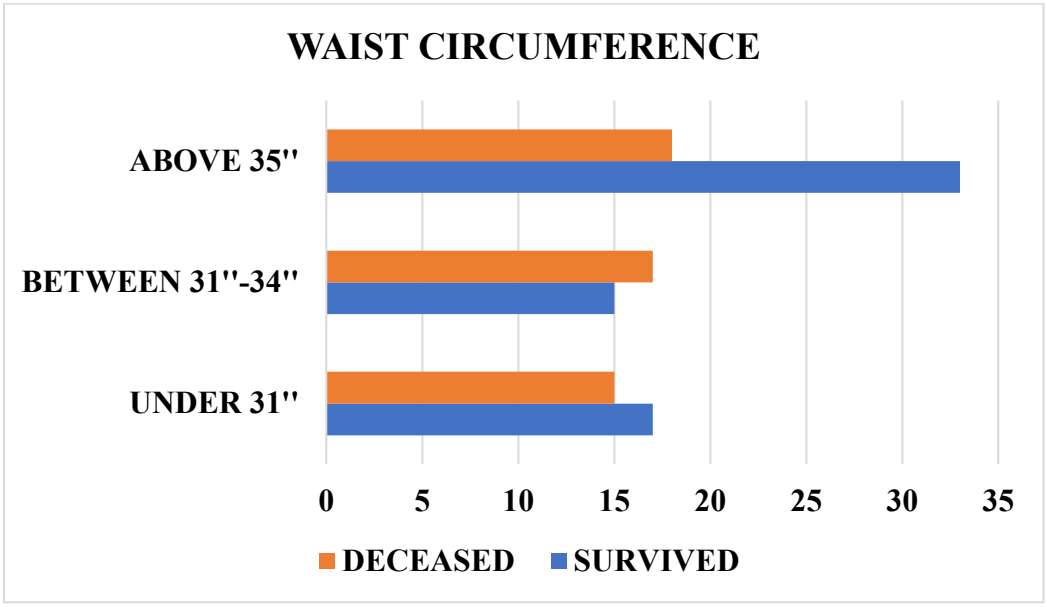

Figure 4: Waist circumference-specific survival rate among breast cancer patients.

3.2.2.3 Hip Circumference (Figure 5): The mean hip circumference was 36.84 cm (SD = 8.40 cm). This provides additional context to the body composition data, which could play a role in determining the relative risk of cancer.

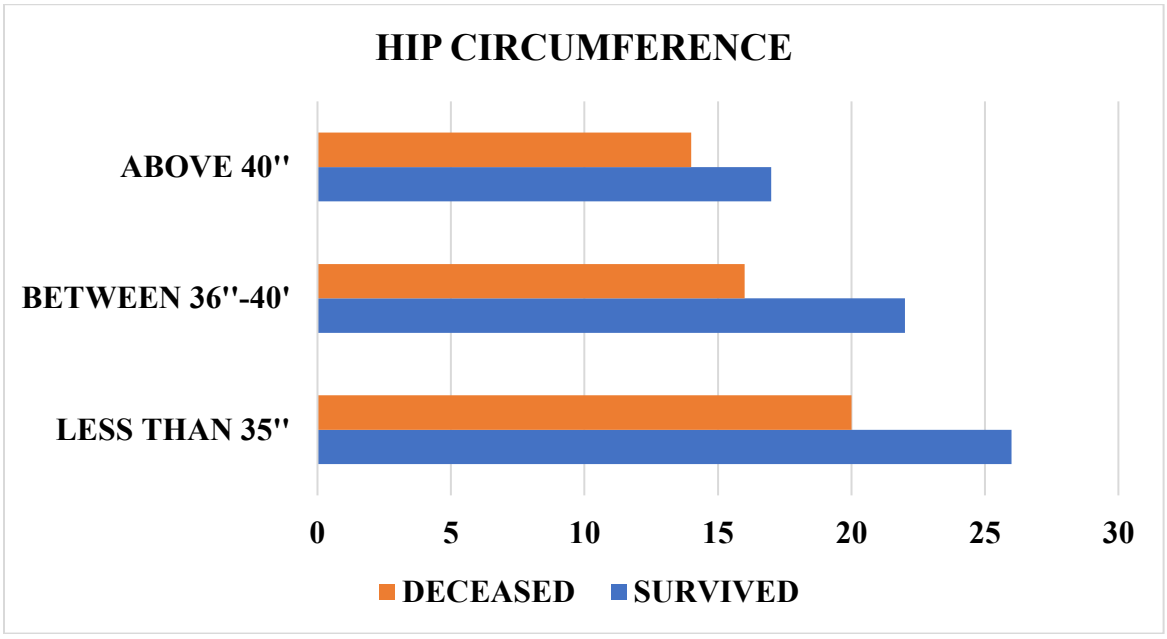

Figure 5: **Hip circumference-specific survival rate among breast cancer patients.**

3.1.3 Menopausal Status (Figure 6): The study participants had a range of menopausal statuses:

a. Pre-Menopausal: Women who have not yet experienced menopause.

b. Post-menopausal: Women who have gone through menopause, which is a critical factor in cancer risk due to hormonal changes.

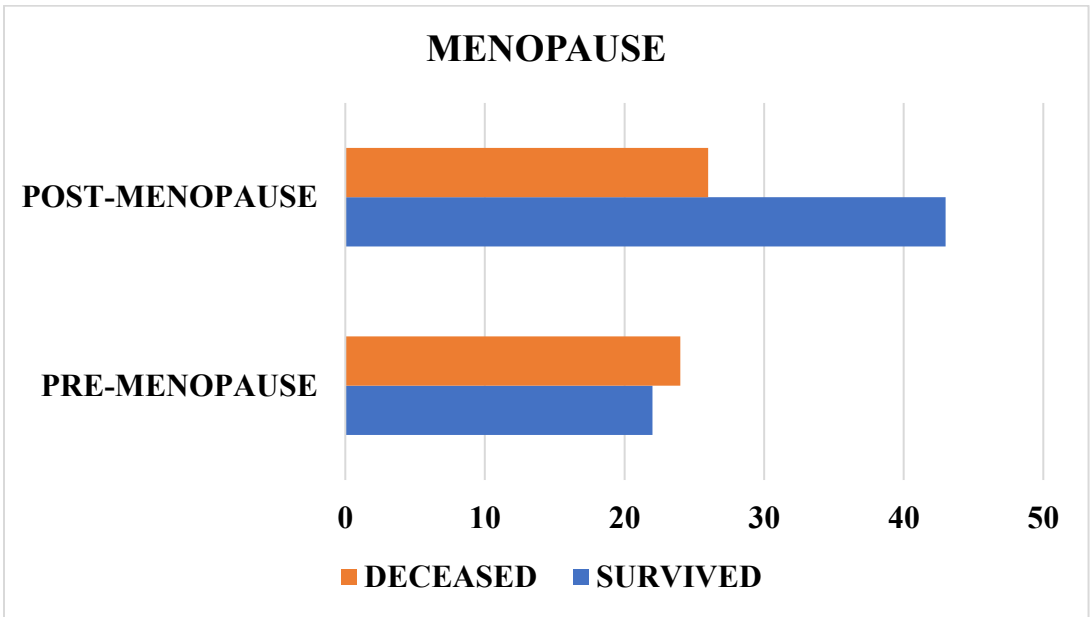

Figure 6: Menopause-specific survival rate among breast cancer patients.

This categorization helps us understand how age and menopause-related hormonal changes might contribute to cancer risk.

3.1.4 Reproductive History:

3.1.4.1 Full Term Pregnancies ranged from 2 to 7, with an average of 3.24. Reproductive factors, such as the number of full-term pregnancies, have been linked to a lower risk of breast cancer due to hormonal changes during pregnancy (Figure 7).

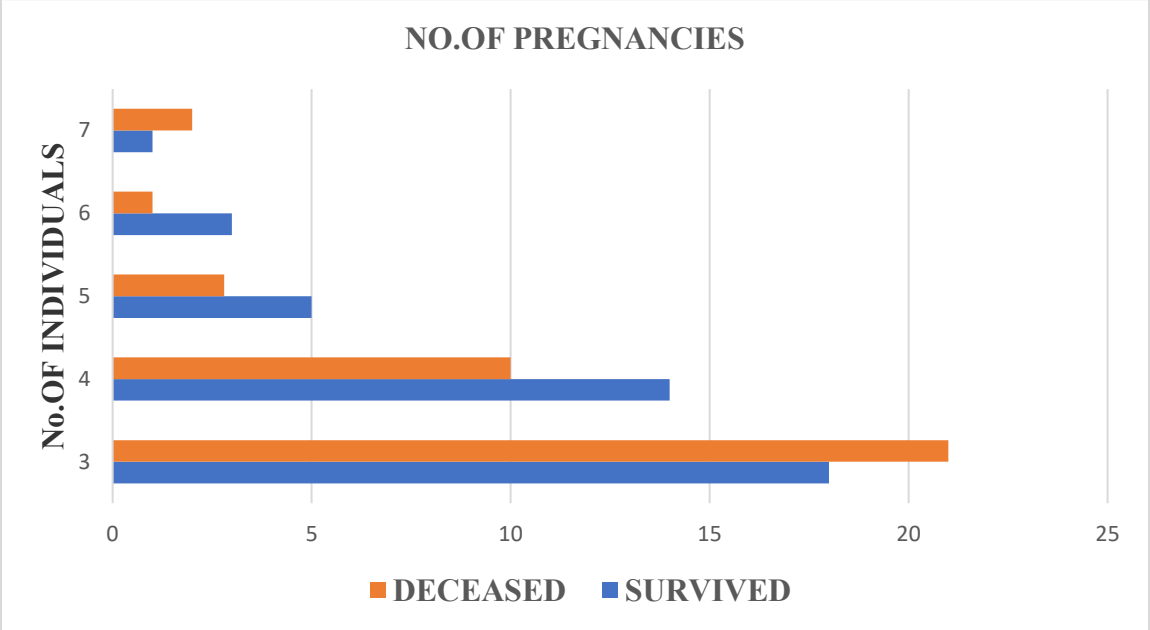

Figure 7: Analysis of survival patterns in breast cancer patients based on pregnancy count.

3.1.4.2 Age at First Child (Figure 8): The mean age for participants having their first child was 22.17 years. Early age at first childbirth is associated with a reduced risk of breast cancer, as it reduces the total number of menstrual cycles a woman experiences.

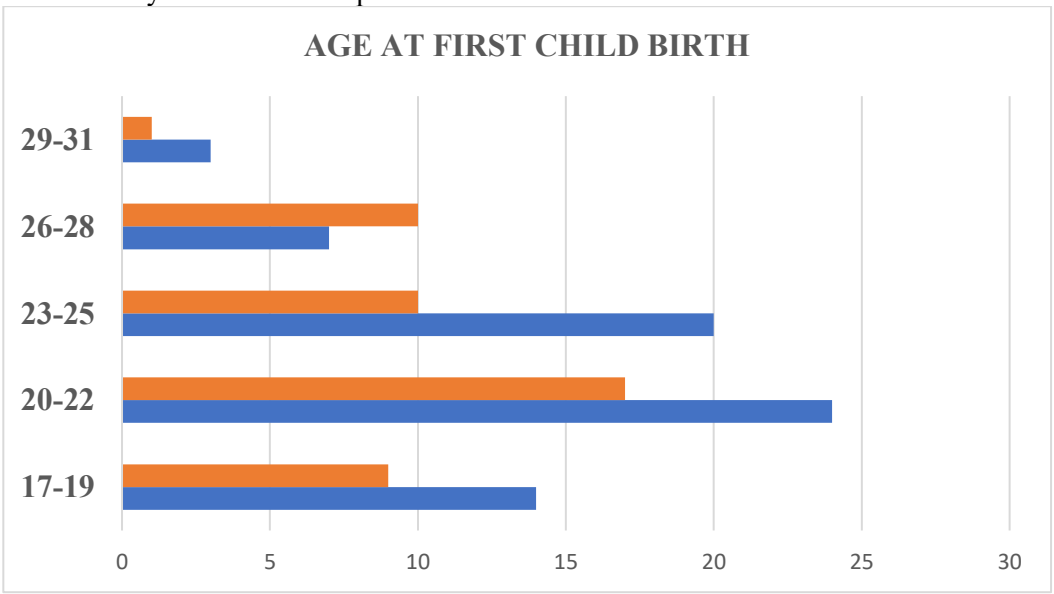

Figure 8: Breast cancer patients show variation with respect to age at first childbirth.

3.2 Chi-Square Test

A Chi-Square test was performed to investigate the association between Menopausal Status and Smoking. The null hypothesis for this test assumed no association between menopausal status and smoking behaviour.

- Chi-Square Statistic: 1.97
- p-value: 0.578

Given the p-value > 0.05, we fail to reject the null hypothesis, meaning there is no significant association between Menopausal Status and Smoking in this population. This finding suggests that smoking behavior is independent of menopausal status. Previous studies, such as those by Kvaavik *et al*. (2010), have shown that smoking is influenced by various lifestyle and socio-economic factors, which may not be directly affected by hormonal changes that occur during menopause.

3.3 ANOVA (Analysis of Variance)

The ANOVA test was used to compare the mean weight between PRE-menopausal and POST-menopausal women. The null hypothesis assumed that there would be no difference in the mean weight between the two groups.

- F-statistic: 0.42
- p-value: 0.517

Since the p-value > 0.05, we again fail to reject the null hypothesis, suggesting that Menopausal Status does not significantly influence weight in this population. This result is somewhat counterintuitive, as research, such as He *et al*. (2006), often suggests that post-menopausal women tend to experience weight gain due to hormonal changes, particularly in the abdominal region. However, other environmental factors such as diet, physical activity, and genetics could be influencing weight in this study, making menopause a less significant factor in this context.

3.4 Correlation Analysis

The correlation analysis examined the relationships between key continuous variables: Age, Weight, Waist Circumference, and Hip Circumference.

3.4.1 Weight and Waist Circumference: There was a weak negative correlation (r = -0.23) between weight and waist circumference. This weak relationship suggests that as weight increases, waist circumference slightly decreases. This could be due to variations in body fat distribution or muscle mass.

3.4.2 Waist Circumference and Hip Circumference: A moderate positive correlation (r = 0.42) was found between waist circumference and hip circumference. This suggests that individuals with larger waist circumferences tend to have larger hips, which could be a reflection of overall body size and fat distribution. This finding aligns with Nordestgaard *et al*. (2008), who found that waist-to-hip ratio is a significant marker for obesity-related health risks.

3.5 Logistic Regression

A logistic regression was performed to predict Smoking Status (Yes/No) based on several predictors. However, due to issues with missing data and inconsistent categorical variables, the logistic regression did not produce valid results. Missing values and inconsistent data encoding in key variables like Exposure and Residence were key obstacles.

DISCUSSION

The analysis aimed to explore various epidemiological patterns of breast cancer risk factors in the Majha region of Punjab. The findings provide several important insights, though limitations in data quality hinder some of the statistical models.

4.1 Menopausal Status and Smoking

The Chi-Square test revealed no significant relationship between Menopausal Status and Smoking. This is consistent with the idea that smoking behaviour is primarily influenced by socioeconomic and psychological factors rather than biological ones like menopause. Studies by Pinheiro *et al*. (2010) have

suggested that while smoking may be more prevalent in younger populations, menopausal status does not directly influence smoking habits.

4.2 Weight and Menopausal Status

The ANOVA did not find a significant difference in Weight between PRE and POST-menopausal women. This is interesting, as several studies have shown that post-menopausal women tend to experience weight gain, particularly in the abdominal region, due to a decrease in estrogen levels. The lack of a significant result in this study could be attributed to other contributing factors such as physical activity, diet, or genetic predispositions, which may mask the effects of menopause.

4.3 Waist and Hip Circumference

The moderate positive correlation between Waist Circumference and Hip Circumference aligns with existing literature, which suggests that abdominal obesity (higher waist circumference) is often correlated with overall body fat, including in the hips. These measurements are useful in assessing overall body composition, and a higher waist-to-hip ratio is known to be associated with increased cancer risk.

4.4 Challenges with Logistic Regression

While the logistic regression model could not be executed successfully due to issues with missing data and inconsistent coding, the intent was to predict smoking status based on various lifestyle and demographic factors. Smoking is a well-known risk factor for breast cancer, particularly in post-menopausal women. The failure to run this model underlines the importance of data quality. Missing values and inconsistent categorical encoding can lead to incorrect conclusions and weak model predictions.

4.5 Limitations of the study:

4.5.1 Data Quality: A significant amount of missing and inconsistent data affected the statistical analyses. Many variables, such as Exposure and Smoking Status, had missing values that limited the scope of our analysis, particularly for logistic regression.

4.5.2 Sample Size: The relatively small sample size (117 participants) could limit the power of the statistical tests, making it difficult to detect subtle relationships between variables.

4.5.3 Measurement Error: Waist and Hip Circumference were self-reported in some cases, which may introduce measurement bias. Accurate and consistent measurement protocols are critical for epidemiological studies.

CONCLUSION AND FUTURE DIRECTIONS

Future research should aim to improve data collection methods, particularly focusing on accurately capturing lifestyle behaviours (e.g., smoking) and environmental exposures (e.g., dust, pesticides). Additionally, a larger and more diverse sample would increase the power of the statistical analyses and improve the generalizability of the findings.

In conclusion, while this study provides valuable insights into the epidemiological patterns of breast cancer risk factors in the Majha region of Punjab, the results should be interpreted with caution due to data quality issues. The study highlights the importance of addressing data gaps and ensuring robust data collection in future research efforts.

REFERENCES

Arnold M, Morgan E, Rumgay H, Mafra A, Singh D, Laversanne M, Vignat J, Gralow JR, Cardoso F, Siesling S, Soerjomataram I (2022). Current and future burden of breast cancer: Global statistics for 2020 and 2040. *Breast* **66, 15–23.**

Deng T, Zi H, Guo XP, Luo LS, Yang YL, Hou JX, Zhou R, Yuan QQ, Liu Q, Huang Q, Wu GS (2025). Global, regional, and national burden of breast cancer, 1990–2021, and projections to 2050: A systematic analysis of the Global Burden of Disease Study 2021. *Thoracic Cancer* **16(9), 1–14.**

Desai K, Baralo B, Kulkarni A, Keshava VE, Iqbal S, Ali H, Prabhakaran SY, Thirumaran R (2024). Cancer statistics: The United States vs worldwide. *Journal of Clinical Oncology*, e23276**.**
Fernandes JO, Machado BF, Cardoso-Filho C, Nativio J, Cabello C, Vale DB (2023). Breast cancer survival after mammography dissemination in Brazil: A population-based analysis of 2,715 cases. *BMC Women's Health* **23(1), 1–7.**
Gaitonde RU, Gajbhiye N (2022). Epidemiology of breast cancer in India. *Asian Journal of Microbiology Biotechnology and Environmental Sciences* **23(1), 204–207.**
Gamal O, Eldin MH, Refaat AA, Hassan RY (2024). Advances in nanocomposites-based electrochemical biosensors for the early diagnosis of breast cancer. *Frontiers in Sensors* **5, 1–14.**
He J, Klag MJ, Whelton PK, Chen JY, Qian MC, He GQ (1994). Body mass and blood pressure in a lean population in southwestern China. *American Journal of Epidemiology* **139(4), 380–389.**
Hortobagyi GN, de la Garza Salazar J, Pritchard K, Amadori D, Haidinger R, Hudis CA, Khaled H, Liu MC, Martin M, Namer M, O'Shaughnessy JA (2005). The global breast cancer burden: Variations in epidemiology and survival. *Clinical Breast Cancer* **6(5), 391–401.**
Kaur H, Singh R, Singh K, Kaur S, Jairath M, Sidhu SK (2021). Pervasiveness and epidemiological profile of SARS-CoV-2 infection among the population of Majha region of Punjab, India. *Journal of Clinical and Diagnostic Research* **15(6), 1–5.**
Kour A, Sharma S, Sambyal V, Guleria K, Singh NR, Uppal MS, Manjari M, Sudan M, Kukreja S (2019). Risk factor analysis for breast cancer in premenopausal and postmenopausal women of Punjab, India. *Asian Pacific Journal of Cancer Prevention* **20(11), 3299–3304.**
Kulothungan V, Ramamoorthy T, Sathishkumar K, Mohan R, Tomy N, Miller GJ, Mathur P (2024). Burden of female breast cancer in India: Estimates of YLDs, YLLs, and DALYs at national and subnational levels based on the national cancer registry programme. *Breast Cancer Research and Treatment* **205(2), 323–332.**
Kumar P, Bolshette NB, Jamdade VS, Mundhe NA, Thakur KK, Saikia KK, Lahkar M (2013). Breast cancer status in India: An overview. *Biomedical Preventive Nutrition* **3(2), 177–183.**
Kvaavik E, Batty GD, Ursin G, Huxley R, Gale CR (2010). Influence of individual and combined health behaviors on total and cause-specific mortality in men and women: The United Kingdom health and lifestyle survey. *Archives of Internal Medicine* **170(8), 711–718.**
Liao L (2024). Inequality in breast cancer: Global statistics from 2022 to 2050. *Breast* **7, 103851.**
Mehrotra R, Yadav K (2022). Breast cancer in India: Present scenario and the challenges ahead. *World Journal of Clinical Oncology* **13(3), 209–218.**
Mihai AM, Ianculescu L, Cretoiu D, Suciu N (2024). Breast cancer screening in Romania: Challenges and opportunities for early detection. *Acta Endocrinologica (Bucharest)* **20(1), 45–50.**
Nanda S, Kumar M, Kumar A, Behal S, Nanda S (2016). Malwa region, the focal point of cancer cases in Punjab: A review study. *International Journal of Current Research in Multidisciplinary* **1(3), 41–45.**
Nordestgaard BG, Palmer TM, Benn M, Zacho J, Tybjaerg-Hansen A, Davey Smith G, Timpson NJ (2012). The effect of elevated body mass index on ischemic heart disease risk: Causal estimates from a Mendelian randomisation approach. *PLoS Medicine* **9(5), e1001212.**
Parashar A, Ansari H, Sharma AK, Khan D (2023). Epidemiological trends of breast cancer: A five-year analysis in a tertiary care centre. *International Journal of Multidisciplinary Research (IJFMR)* **5(6), 1–12.**
Pinheiro SP, Holmes MD, Pollak MN, Barbieri RL, Hankinson SE (2005). Racial differences in premenopausal endogenous hormones. *Cancer Epidemiology Biomarkers and Prevention* **14(9), 2147–2153.**
Rahman MA, Hamada M, Sharmin S, Rimi TA, Talukder AS, Imran N, Kobra K, Ahmed MR, Rabbi M, Matin MMH, Ali MA (2024). Enhancing early breast cancer detection through advanced data analysis. *IEEE Access* **12, 161941–161953.**

Rana MK, Barwal TS, Sharma U, Bansal R, Singh K, Rana APS, Jain A, Khera U, Singh Sr K (2021). Current trends of carcinoma: Experience of a tertiary care cancer center in North India. *Cureus* **13(6), 1–8.**