Available online at https://asric.africa/health-sciences
ASRIC Journal of Health Sciences Vol.5(1)(2025) 121-153

Histological and Molecular Subtyping of Breast Cancer with Epstein–Barr Virus LMP1 Expression among Northern Nigerian Women: A Combined Survey

Maryam Ibrahim Rimi[1], Aminu Zakari Mohammed[2], Isah Abubakar Aliyu[3]

[1]Department of Medical Laboratory Science, Faculty of Allied Health Science, Kaduna State University, Kaduna, Nigeria, bindawa.maryam@kasu.edu.ng
[2]Pathology Department, Aminu Kano Teaching Hospital, Kano, Nigeria, aminuzm@yahoo.com
[3]Department of Medical Laboratory Science, Faculty of Allied Health Science, College of Health Sciences, Bayero University Kano, Nigeria, iaaliyu.mls@buk.edu.ng

Corresponding Author: bindawa.maryam@kasu.edu.ng

Received 16 July 2025; revised 21 August 2025; accepted 29 September 2025

Abstract

Epstein–Barr virus (EBV) has been linked to several cancers, but its involvement in breast cancer is not well understood in sub-Saharan Africa. This study examined the sociodemographic, lifestyle, and reproductive health characteristics of women in Katsina and Kaduna States, Nigeria, and investigated histological and molecular subtypes of breast cancer with attention to EBV latent membrane protein 1 (LMP1) expression. **Methods:** Three hundred and twenty-five confirmed breast cancer cases were reviewed. Questionnaire data were used to capture sociodemographic, lifestyle, and reproductive factors. Histological classification was carried out on all cases, while immunohistochemistry was performed on 91 and 63 samples to determine molecular subtypes and EBV LMP1 expression resoectively. Odds ratio analysis of EBV exposure and breast cancer risk in Africa was drawn from pooled data. **Results:** Most participants were married housewives with limited education. Differences between states were observed in education, physical activity, contraceptive use, and cultural habits such as grooming kit sharing and communal eating. Almost all women avoided smoking and alcohol, ate vegetables regularly, breastfed, and reported no current pregnancy. Invasive ductal carcinoma was the most frequent histological subtype in both Katsina (81.2%) and Kaduna (84.0%), with no significant differences between them. Of the 91 cases analyzed by immunohistochemistry, the triple-negative molecular subtype was most common (47.2%), followed by HER2-positive (35.8%). EBV LMP1 expression was not detected in any of the samples. Odds ratio analysis, however, showed a strong association between EBV exposure and breast cancer risk in Africa (OR = 24.17, 95% CI: 12.46–46.87, $p<0.0001$). The absence of LMP1 expression prevented biomarker correlation. **Conclusion:** Breast cancer in Northern Nigeria is mainly invasive ductal carcinoma and most frequently triple-negative in molecular profile. While EBV exposure appears strongly associated with breast cancer risk in Africa, the lack of detectable LMP1 expression suggests latent infection or protein suppression, limiting its value as a biomarker. Further molecular studies are needed to clarify the role of EBV in breast cancer development.

Keywords: Epstein–Barr virus (EBV), Breast cancer, Immunohistochemistry, Histological types, Moleculer subtyping, Scarff-Bloom Richardson Grading.

Introduction

Breast cancer continues to be the most prevalent cancer among women globally and a leading cause of cancer-related mortality. In 2022, approximately 2.3 million women were diagnosed with breast cancer worldwide, resulting in about 670,000 deaths (WHO, 2024). In Nigeria, breast cancer is the leading cancer among women, with increasing incidence and mortality rates, largely due to late presentation, limited access to healthcare, and sociocultural barriers (Jedy-Agba *et al.,* 2020). Across sub-Saharan Africa, breast cancer poses a growing public health challenge, often diagnosed at advanced stages, with survival rates significantly lower than in high-income countries (Fungisai *et al.,* 2021). The burden is compounded by a lack of organized screening programs, inadequate infrastructure for early detection, and insufficient public awareness (Sengayi-Muchengeti *et al.,* 2022). Similarly, In Egypt, breast cancer remains the most common malignancy among females, accounting for a significant proportion of cancer cases (International Agency for Research on Cancer [IARC], 2025)

The role of oncogenic viruses in breast cancer has been a subject of research for decades. Recent studies have highlighted the potential involvement of viruses such as Epstein-Barr Virus (EBV), high-risk Human Papillomavirus (HPV), Mouse Mammary Tumor Virus (MMTV), and Cytomegalovirus (CMV) in breast carcinogenesis (Glenn *et al.,* 2012; Richardson *et al.,* 2020). Co-infection with multiple oncogenic viruses, particularly EBV and HPV, has been observed in breast cancer tissues, suggesting a synergistic effect in tumor development (Dai *et al.,* 2024).

The presence of EBV and HPV co-infection in breast cancer has been associated with more aggressive tumor characteristics and poorer prognosis (Gupta *et al.,* 2020). EBV contributes to oncogenesis by expressing viral proteins that interfere with cellular pathways, promoting immune evasion and inflammation (Lehle *et al.,* 2024). Similarly, high-risk HPV types disrupt tumor suppressor genes, facilitating uncontrolled cell proliferation (Dai *et al.,* 2024).

Recent research has also focused on the role of Apolipoprotein B mRNA editing enzyme catalytic subunit-like protein 3B (APOBEC3B), a DNA cytosine deaminase, in breast cancer. Aberrant expression of APOBEC3B leads to increased mutagenesis, contributing to tumor heterogeneity and therapy resistance (Zhang *et al.,* 2023). The upregulation of APOBEC3B has been linked to viral infections, including HPV, suggesting a possible connection between viral oncogenesis and APOBEC3B-mediated mutagenesis (Lehle *et al.,* 2024).

Molecular techniques such as Polymerase Chain Reaction (PCR) and Immunohistochemistry (IHC) remain the most sensitive methods for detecting viral presence in malignant tissues (Glenn *et al.,* 2012). The Aim, Objectives and Hypothesis of this study are as follows:

Aim

To investigate the relationship between Epstein–Barr virus (EBV) infection and breast cancer among women in Northern Nigeria by examining sociodemographic, lifestyle, reproductive health factors, and the distribution of histological and molecular subtypes, with emphasis on EBV latent membrane protein 1 (LMP1) expression.

Objectives

1. To describe the sociodemographic, lifestyle, and reproductive health profiles of women with breast cancer in Katsina and Kaduna States.
2. To determine the distribution of histological subtypes and SBR grading of breast cancer in the study participants.
3. To classify the molecular subtypes of breast cancer using immunohistochemistry (IHC).
4. To detect EBV latent membrane protein 1 (LMP1) expression in breast cancer tissues using IHC.
5. To assess the association between EBV exposure and breast cancer risk in the study setting.

Hypotheses

1. **H_0 (null):** The distributions of sociodemographic, lifestyle, and reproductive health characteristics do not differ between women with breast cancer in Katsina and those in Kaduna.
2. **H_1 (alternative):** The distributions of sociodemographic, lifestyle, and reproductive health characteristics differ between women with breast cancer in Katsina and those in Kaduna.
3. **H_0:** The proportionate distribution of histological subtypes (e.g., IDC, ILC, others) is uniform across the study population (no subtype predominates or no difference between locations).
4. **H_1:** The distribution of histological subtypes is not uniform across the study population (one or more subtypes predominate and/or proportions differ by location).
5. **H_0:** The distribution of molecular subtypes determined by IHC (e.g., triple-negative, HER2-enriched, luminal A/B) does not differ between Katsina and Kaduna and shows no clear dominant subtype.
6. **H_1:** The distribution of molecular subtypes determined by IHC differs between Katsina and Kaduna and/or a specific molecular subtype is dominant in the study population.
7. **H_0:** EBV LMP1 expression is not detectable in breast cancer tissue samples by IHC (no cases show LMP1 positivity).
8. **H_1:** EBV LMP1 expression is detectable in at least some breast cancer tissue samples by IHC (one or more cases show LMP1 positivity).
9. **H_0:** There is no association between EBV exposure and breast cancer risk in the study population (EBV exposure does not affect the odds of breast cancer).
10. **H_1:** There is an association between EBV exposure and breast cancer risk in the study population (EBV exposure is associated with increased odds of breast cancer).

Statement of the Problem

Breast cancer remains a significant public health concern in Nigeria, characterized by high incidence rates and often poor prognoses. Recent studies indicate that a substantial proportion of breast cancer patients in Nigeria are young, premenopausal women who frequently present with advanced-stage disease, leading to challenges in treatment compliance and outcomes (Agolli *et al.,* 2023). This scenario underscores the need for comprehensive research into potential etiological factors, including viral associations.

Epstein-Barr Virus (EBV), a ubiquitous gamma-herpesvirus, has been implicated in various malignancies and is hypothesized to play a role in breast carcinogenesis. Evidence suggests that EBV may contribute to breast cancer development through mechanisms such as DNA methylation alterations and immune modulation (Abdallah *et al.,* 2018). However, the exact relationship between EBV and breast cancer, particularly in the Nigerian context, remains inadequately explored.

Geographical variations in EBV prevalence and its association with breast cancer have been observed. For instance, studies have reported varying EBV detection rates in breast cancer tissues across different regions, with higher prevalence noted in parts of Africa and Asia compared to Western countries (Agolli *et al.,* 2023). These disparities highlight the necessity for region-specific investigations to elucidate the potential role of EBV in breast cancer within Nigeria.

Justification

The association between EBV and breast cancer remains a subject of ongoing debate. While some studies have identified EBV DNA in breast cancer tissues, others have failed to establish a significant correlation (Agolli *et al.,* 2023). This inconsistency may be attributed to differences in detection methods, sample types, and population genetics (Agolli *et al.,* 2023).

In Nigeria, there is a paucity of comprehensive studies examining the prevalence and impact of EBV in breast cancer patients. Identifying the presence of EBV in breast cancer tissues and understanding its potential role in tumorigenesis could provide valuable insights into disease mechanisms and inform targeted therapeutic strategies (Abdallah *et al.,* 2018).

Furthermore, exploring the epigenetic modifications associated with EBV infection, such as DNA methylation patterns, may reveal critical pathways involved in breast cancer development. Such investigations are essential for developing novel diagnostic and prognostic biomarkers, ultimately improving patient outcomes in Nigeria and similar settings (Abdallah *et al.,* 2018). These methods have been instrumental in identifying the association between viral infections and breast cancer, providing insights into potential diagnostic and therapeutic targets.

Materials and Methods

Description of the Study Area

This study was a combined survey carried out both in Katsina and Kaduna State; Nigeria at Federal Medical Teaching Hospital and Barau Dikko Specialist Hospital respectively. Both States are in the Northwest of Nigeria's six geopolitical zones, a geographical and political representation of the country; other states include Jigawa, Kano, Kebbi, Sokoto, and Zamfara (Quadri, 2021). They together cover a total area of 42,752 sq mi and an area rank of 4th of 36. They have a total population of 14,833,784 people going by 2022 census leaving them both the 3rd of 36 in rank and a density of 200.4/km^2 (340/sq mi) population density with 2.5% annual population change (2006-2022). They are populated by about 59 to 63 different ethnic groups if not more (Rimi *et al*., 2024); where Hausa and Fulani are the dominant ethnic groups. Its water supply is sourced through damping of rivers and digging of wells and boreholes. Kaduna State consists of twenty-three (23) Local Government Areas while Katsina on the other hand has 34 known villages (Rimi *et al*., 2024).

Ethics Statement

The ethical clearance certificates to conduct this study were obtained from the human research committee of the Ministry of Health Kaduna State and from Ethical committee, Federal Teaching Hospital, Katsina, Nigeria with the following NREC numbers respectively

- NHREC/17/03/2018 (Appendix I)
- FTHKTNHREC.REG.24/06/22C/83 (Appendix II)

This study adhered to the ethical principle of Helsinki for collection and handling of samples from human subjects.

Consent Form Design

Written and verbal consents were obtained from all study participants after explaining the objectives of the study with right to withdraw at any point in the research before recruiting them into the study. Confidentiality was also upheld (Appendix III).

Questionnaire Design

Demographic data, location, contacts, life style, contraceptives and other drugs use, disease history, anthropometric among others were included to fit the study design. (Appendix IV).

Study Design and Study Population

This is cross-sectional with the following description; a molecular pathology study with a correlational/association and comparative analysis component performed between 4th July; 2023 to 20th September; 2025.

Inclusion Criteria

The inclusion criteria were defined as female patients with confirmed histopathological evidence of breast cancer and accessibility of FFPE blocks. No limitations in age, type of cancer, tumour size, and stage was considered for the inclusion of participants in this study.

Exclusion Criteria

Inadequate or poor-quality tissue blocks (insufficient for histology or IHC). Patients with incomplete clinical records. Duplicate or extensively depleted archival samples. Breast tumours that were of metastatic nature whose origin is not breast cancer.

Steps Taken to Obtain the Tissue Blocks

Sampling: Purposive sampling technique was employed in obtaining samples

Sample Size Determination

The Single proportion formula was used to calculate number of study participants and to estimate the sample size for determining the prevalence of EBV targeting the EBNA1-positive samples. An adjustment factor of 10–15% was applied to account for potential non-response or sample loss:

$$n_0 = \frac{Z^2 \cdot p(1-p)}{d^2}$$

Where:

- n_0 = required **s**ample size
- Z = Z–score corresponding to desired confidence level (e.g., 1.96 for 95%)
- p = estimated prevalence (Dimkpa *et al*., 2024)
- 1−p = complement
- d^2 = margin of error (precision)

$$d = Z \cdot \sqrt{\frac{p(1-p)}{n}}$$

$$d = 1.96 \cdot \sqrt{\frac{0.10 \times 0.90}{325}} = 1.96 \cdot \sqrt{0.000277}$$

= 0.0311

With 311 samples, the prevalence of EBV was initially estimated with a margin of error of ±3.25%, offering greater precision than the conventional ±5%. To account for potential non-response or poor-quality samples, the sample size was rounded up to 325. This sample size was statistically justified given the expected low prevalence of EBV in breast cancer, the goal of achieving high precision (±3%), and the planned stratified analyses by age and tumour grade.

Formula for sample size determination of LMP1 and Molecular Subtyping

The finite population correction was used to calculate the sample size for LMP1 expression and molecular subtyping from 243 and 325 initial sets as follows for adjustment**;** For the IHC and Molecular Subtyping Subset.

$$n_{\text{adj}} = \frac{n_0}{1 + \left(\frac{n_0 - 1}{N}\right)}$$

This is not random, so no formal probabilistic sampling formula applies. However, to justify it based on comparison, correlation, and statistical power, the following was incorporated: estimating a proportion of prevalence of LMP1-positive sample expression in a subgroup. Out of 325 positive breast cancer samples, finite population correction (FPC) was used, drawing a subset from a known population of 243 and 325 respectively as follows.

$n_{adj}=n_0 1+(n_0-1N)n\_$

Where: n_0= sample size from the basic formula (as above) and N = total population (243)

If n_0=68 and N=243

$$n_{\text{adj}} = \frac{68}{1 + \left(\frac{68-1}{243}\right)} = \frac{68}{1 + \left(\frac{67}{243}\right)} = \frac{68}{1.275} = 53.3$$

Fifty four (54) samples were calculated, but this was rounded up to 63 for adequacy. As such Sixty three (63) FFPE blocks were utilized for LMP1 expression detection and 91 samples for molecular subtyping. Table Summary as follows:

Finite Population Correction (FPC) Results

Case	Population Size (N)	Initial Sample Size (n_0)	Formula	Adjusted Sample Size (nadjn_(text(adj)))
PCR Analysis	243	68	$\frac{68}{1 + \left(\frac{68-1}{243}\right)} \quad \frac{68}{1.275}$	53.3 approximately 53 (Liu *et al*., 2023) Rounded to 63 samples for adequacy
Tissue Blocks	325	109	$\frac{109}{1 + \left(\frac{109-1}{325}\right)} \quad \frac{109}{1.332}$	91.8 approximately 91 (Belachew *et al*., 2023)

This way, the two adjusted sample sizes are clearly justified with the same formula.

- The n_0=68 used was derived from that basic formula for the PCR subpopulation (N = 243).
- The n_0=109 was similarly calculated for the larger population (N = 325 tissue blocks).

Sample Requirement:

According to the inclusion and exclusion criteria, 325 breast cancer patients were included as the participants, calculation followed in the subsequent chapter. 63 tissue blocks were collected from the patients in the period stated. All tissue samples were formalin fixed paraffin embedded. Furthermore, the stage of the tumours according to the cTNM and pTNM system were provided by consulting an expert cancer team consisting of an oncologist, a radiologist, and a cancer surgeon. All breast tissue biopsies were accompanied by duly completed forms and properly fixed from theatre and all samples were registered properly in the laboratory and unit register after ensuring the above.

Retrospective Procedure

Archived formalin-fixed paraffin-embedded (FFPE) breast tissue blocks were obtained from both study centers following presentation of ethical approval certificates. Histologically normal slides were selected as controls for comparison with breast cancer photomicrographs. Tumor staging was determined according to the cTNM and pTNM systems for clinical and pathological respectively by a specialist cancer team comprising an oncologist and a general surgeon. All archived biopsies were properly documented with completed laboratory request forms and processed in accordance with standard histopathological procedures.

Prospective Procedure

In addition to archived material, prospectively collected samples were obtained from newly diagnosed breast cancer patients who consented to participate in the study. After informed consent and questionnaire administration, tissue biopsies were retrieved directly from the theatre, registered, fixed, grossed, processed, embedded, sectioned, stained, and mounted for histopathological analysis following the standard operating procedures postulated below. Photomicrographs were subsequently prepared for further evaluation.

Sample Processing and Analyses (Appendix V)

This involve histopathological analyses of tissues according to WHO standards and SOPs involving grossing, processing, embedding, microtomy, staining and photomicrography as follows:

Grossing Procedure:

The specimen and orientation marker were identified. Dimensions (length, breadth, and width) were measured. The external surface was examined in detail. Breast tissues were sectioned at 1 cm intervals, and cut surfaces were described with any pathological processes identified. Suspected lesions were measured and documented, noting colour, consistency, and shape. Surgical margins were assessed. Smaller biopsies were identified, numbered, aggregated, and measured. Fragment shape, colour, and consistency were recorded. Care was taken to avoid transection of inked margins. All small biopsies were secured in cassettes to prevent tissue loss during processing. All fragments were submitted, dipped in eosin for visibility, and wrapped in filter paper. Macroscopic findings were documented on the accompanying form. Well-grossed and labelled tissues were placed in cassettes for loading into the automatic tissue processor. Labelling was maintained throughout until microscopic examination.

Tissue Processing using the Automatic Tissue Processor
Procedure:

Tissue samples obtained through excision biopsy (at least 10 mm thick) were placed in labelled cassettes and loaded into a carrier basket. The basket was inserted into an Automatic Tissue Processor, which contained a series of beakers with 10% formalin, graded alcohols (70%, 90%, and absolute alcohol in three changes), xylene (three changes), and two changes of molten paraffin wax. Processing was carried out using

program 2, which runs for approximately 23–24 hours. After processing, the carrier was removed and the tissues were transferred to the embedding chamber. At this stage, the tissues were well-dehydrated, infiltrated with paraffin, and ready for embedding.

Embedding

Using pre-warmed forceps (to prevent paraffin from adhering), the cassette containing the impregnated tissue was removed from the paraffin reservoir. A mould was first placed on a warm plate (30–80 °C) and then transferred to the cold plate (–5 to –15 °C), allowing a thin solid layer of paraffin to form at the base. The tissue surface was carefully positioned against this solid layer to ensure proper orientation for sectioning. The cassette was then placed firmly on top of the mould, which was subsequently filled with molten paraffin. The combined mould and cassette were cooled on the cold plate of the embedding machine for about 15 minutes until solidification occurred. Finally, the mould was separated from the cassette, leaving the tissue embedded in a solid paraffin block, ready for microtome sectioning.

Microtomy (Formalin Fixed Paraffin-Embedded Block of Tissue)

Sample Requirement:

All paraffin-embedded tissue blocks were checked to ensure that each was correctly labelled and accompanied by the corresponding request form. The blocks were then placed on an ice tray for 20 minutes before sectioning. A Rotary Cut 5062 microtome was set up according to SLEE Medical operating instructions, along with the tissue floating bath and stretching table. The tissue surface was exposed by trimming at 20 µm, after which the blocks were again placed on ice for 20 minutes. Blocks were labelled on the frosted end of the slides with their identification numbers. Thin sections of 3 µm were cut to produce ribbons, which were then stretched to remove folds using 20% alcohol and the floating bath maintained at 40–50 °C (5–10 °C below paraffin melting point). Sections were transferred onto slides, drained on a slide dryer at 50 °C for 10 minutes, and arranged in racks for staining. Finally, the slides were incubated in an oven for 10 minutes to achieve dewaxing.

Staining Procedure:

Labelled slides were arranged in staining racks and sequentially processed as follows: immersion in xylene I for 5 minutes, followed by xylene II for another 5 minutes. Slides were then passed through absolute alcohol (5 minutes), 90% alcohol (2 minutes), 70% alcohol (2 minutes), and rinsed in water for 2 minutes. Staining was performed with Cole's haematoxylin for 5 minutes, followed by rinsing and bluing in tap water for 5 minutes. Counterstaining was done with eosin yellow for 1 minute 30 seconds, after which the slides were rinsed in water and dehydrated through ascending grades of alcohol. Finally, slides were cleared in xylene and mounted with DPX mountant. Photomicrographs were captured using Optika imaging software.

Immunohistochemistry Analyses for LMP Detection

This was carried out at National Hospital. Paraffin-embedded tissue blocks were sectioned at 4 µm thickness. Sections were placed on a hot plate for at least one hour to melt the paraffin and incubated alongside a positive control. The sections were then rehydrated in water, and antigen retrieval was performed by heating in citrate buffer at 100 °C for 30 minutes. Endogenous peroxidase activity was blocked by incubating the sections with a peroxidase-blocking solution for 20 minutes, followed by thorough washing with phosphate-buffered saline (PBS). A liquid repellent marker (A-PAP pen) was used to demarcate the sections and retain reagents during staining.

Sections were incubated with the primary antibody against LMP1, ER, PR and Her2neu prepared by diluting 10 µL of antigens of interests' stock solution in 1 mL PBS, at room temperature for one hour to allow antigen-antibody binding. After washing with PBS, sections were incubated with the secondary antibody for 20 minutes, followed by incubation with horseradish peroxidase (HRP) conjugate for 40 minutes to enable secondary antibody binding. Excess HRP was removed by washing the sections five

times with PBS, after which 3,3'-diaminobenzidine (DAB) chromogen was applied for 5 minutes to visualize LMP1, ER, PR and Her2neu expression.

The sections were then rinsed in water, counterstained with hematoxylin for 15 seconds to provide contrast, dehydrated in graded alcohols, and dried in a hot air oven. Finally, the sections were mounted with DPX mountant, and photomicrographs were captured using a light microscope.

Statistical Analyses

The data was edited and summarised in Microsoft Excel Spreadsheet. Missing values and outliers were identified and recorded appropriately. Normality test and tests of significance were conducted using SPSS version 27. Categorical variables from the participant's clinical attributes were summarized using frequency and percentages. Continuous variables were summarised as mean, standard deviation, and degrees of freedom. Chi-square tests were conducted to assess p-values set at $P<0.05$ to determine the level of significance. Incidence Risk (IR) and Risk Ratios (RR) were computed using IBM Statistics version 27. The odds ratio was used to quantify the strength of association between Epstein-Barr virus (EBV) presence and breast cancer among the study participants. Additionally, the Phi coefficient (ϕ) was calculated to evaluate the correlation between LMP1 expression in immunohistochemically analysed EBV-positive breast cancer samples and EBV-negative samples.

Results

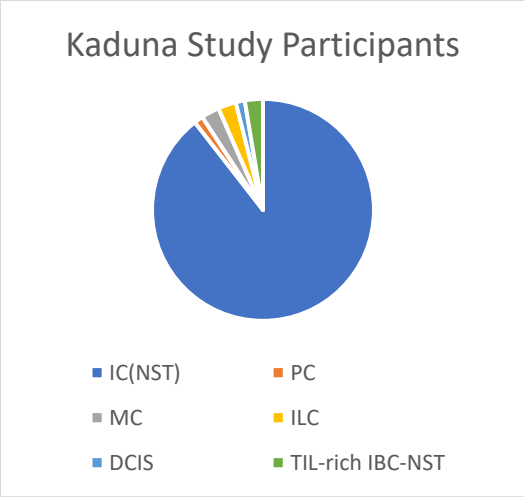
1: Histological Types of Breast Cancer of Study Participants

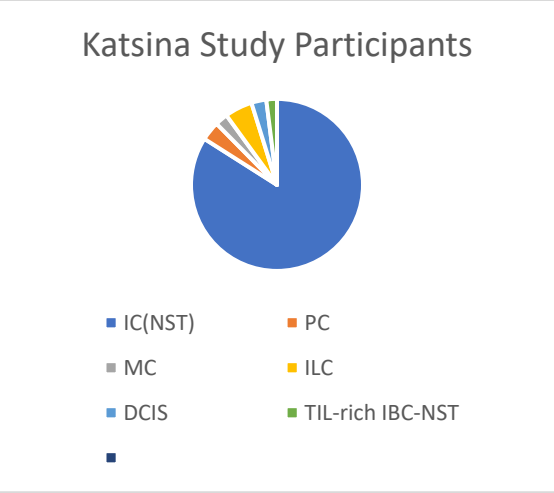
The predominant histological type observed was invasive carcinoma of no special type (IC-NST), accounting for 85.5% of all cases, followed by invasive lobular carcinoma (4.7%). Less frequent types included papillary carcinoma (3.1%), mucinous carcinoma (2.5%), ductal carcinoma in situ (2.5%), and tumour-infiltrating lymphocyte–rich IC-NST with a medullary pattern (2.2%). The distribution of histological subtypes was similar across both study locations, with no statistically significant difference ($\chi^2(5) = 2.68$, $p = 0.75$) (Table 1).

Table 1: Histological Types of Breast Cancer of Study Participants

Histological Type	Katsina	(%)	Kaduna	(%)	Total	(%)	χ2, p-value
IC(NST)	210	84%	68	90.7%	278	85.5%	
PC	9	3.6%	1	1.3%	10	3.1%	
MC	6	2.4%	2	2.7%	8	2.5%	
ILC	13	5.2%	2	2.7%	15	4.7%	$\chi^2(5) = 2.68$, $p>0.05$
DCIS	7	2.8%	1	1.3%	8	2.5%	
MeC	5	2.0%	2	2.7%	7	2.2%	
Total	**250**	**100.0%**	**75**	**100.0%**	**325**	**100.0%**	
							P-value = 0.75

Key: IC(NST)-Invasive Carcinoma of No Special Type, PC-Papillary Carcinoma, MC-Mucinous Carcinoma, ILC-Invasive Lobular Carcinoma, DCIS-Ductal Carcinoma InSitu, MeC- Medullary Carcinoma



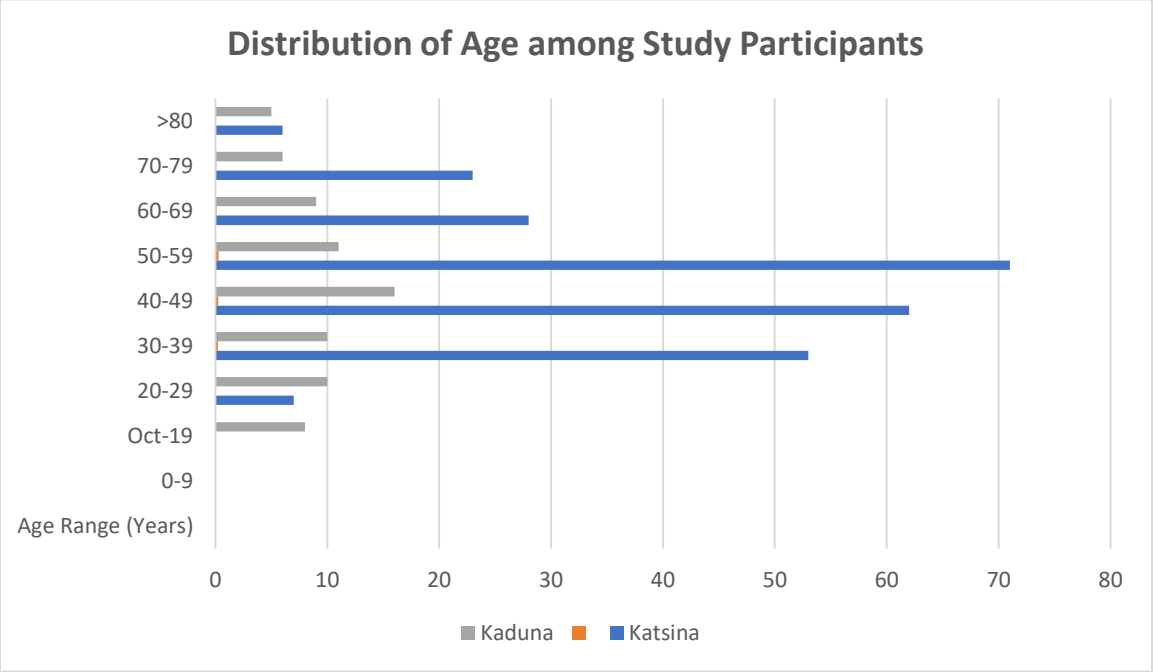


χ2, p-value: $\chi^2(5) = 2.68$, $p>0.05$, P-value = 0.75

Key: IC(NST)-Invasive Carcinoma of No Special Type, PC-Papillary Carcinoma, MC-Mucinous Carcinoma, ILC-Invasive Lobular Carcinoma, DCIS-Ductal Carcinoma InSitu, MeC - medullary Carcinoma

2: Age Distribution of Study Participants

The mean age of participants was 49.63 ± 17.56 years. The largest proportion of cases occurred in the 40–49 years (24.0%) and 50–59 years (25.2%) age groups, consistent with the middle-aged peak incidence of breast cancer. Fewer participants were observed at the extremes of age (<20 years and >80 years). A significant difference in age distribution was observed between the two study locations ($\chi^2(4) = 40.19$, $p < 0.001$), with a relatively higher proportion of younger participants (10–29 years) in Kaduna compared to Katsina (Figure 2).



χ2, p-values $\chi2(4) = 40.19$, $p<0.001$ p = 0.0000000223, Mean ±SD= 49.63± 17.56

Figure 2: Distribution of Age among Study Participants

The majority of participants were within the 40–59 years age range, with peaks in the 40–49 and 50–59 year groups, reflecting the typical age of breast cancer presentation in this cohort. Fewer cases were recorded at the extremes of age (<20 years and >80 years).

3: SBR Grading of Study Participant

The distribution of study participants according to the Scarff-Bloom-Richardson (SBR) grading system. Among participants from Katsina, the majority were in grade 9 (27.7%), followed by grades 6 (17.4%) and 7 (16.3%). In Kaduna, the highest proportion was observed in grade 5 (27.0%), with grades 6 (18.9%) and 7 (17.6%) following closely. Overall, combining both locations, grade 9 constituted the largest proportion of cases (23.3%), followed by grades 6 (17.8%) and 7 (16.7%).

The differences in SBR grade distribution between Katsina and Kaduna were statistically significant ($\chi^2(2) = 10.29$, $p < 0.01$; overall $p = 0.0058$), indicating variation in tumor differentiation among the study populations. Lower-grade tumors (grades 0–4) were relatively uncommon, accounting for 16.0% of cases overall, whereas higher-grade tumors (grades 5–9) represented the majority (84.0%), suggesting a predominance of moderately to poorly differentiated breast cancer in the study cohort (Figure 4.6).

SBR histological grading showed significant associations with patients' level of education ($\chi^2=12.37$, $p<0.05$) and physical activity ($\chi^2=10.37$, $p<0.01$), but not with Duration of Symptoms ($\chi^2=0.19$, $p>0.05$). Specifically, lower educational attainment and sedentary lifestyle were associated with higher histological grades (Grade III), suggesting socio-behavioral influences on tumor aggressiveness (Table 2).

Table 2: SBR Grading of Study Participants

SBR Grading	Katsina	(%)	Kaduna	(%)	Total	(%)	χ2, p-value
0	18	9.8%	7	9.5%	25	9.7%	
1	1	0.5%	1	1.4%	2	0.8%	
2	1	0.5%	1	1.4%	2	0.8%	
3	1	0.5%	1	1.4%	2	0.8%	
4	10	5.4%	2	2.7%	12	4.7%	$\chi2(2) = 10.29$, $p<0.01$
5	17	9.2%	20	27.0%	37	14.3%	
6	32	17.4%	14	18.9%	46	17.8%	
7	30	16.3%	13	17.6%	43	16.7%	
8	24	13.0%	9	12.2%	33	12.8%	
9	51	27.7%	9	12.2%	60	23.3%	
Total	**184**	**100.0%**	**74**	**100.0%**	**258**	**100.0%**	
							P-value=0.0058

4: Results of Molecular Subtyping among study participants

Molecular subtyping of study participants was based on immunohistochemical (IHC) markers, estrogen receptor (ER), progesterone receptor (PR), and Her2neu (ERBB2). In Katsina, most cases were triple-negative (ER–, PR–, Her2neu–), accounting for 52.6%, followed by ER–, PR–, Her2neu+ at 28.9%. In Kaduna, 33.3% were triple-negative and 53.3% were ER–, PR–, Her2neu+. Other subtypes, including ER+/PR+/Her2neu+ (1.9%) and ER+/PR+/Her2neu– (7.5%), were much less common. Overall, triple-negative tumors represented 47.2% of cases, ER–/PR–/Her2neu+ tumors 35.8%, and all other subtypes together 17.0%. Statistical analysis showed no significant difference in IHC distribution between Katsina

and Kaduna ($\chi^2(1) = 1.61$, $p = 0.204$), suggesting similar receptor profiles in both populations, with a predominance of aggressive subtypes.

When analyzed by parity, triple-negative tumors were more frequent among women with higher parity (≥6 children), accounting for nearly half of cases in this group. In contrast, luminal subtypes appeared less often overall but were relatively more common among women with fewer children (≤2). Although the relationship between parity and molecular subtype did not reach statistical significance ($\chi^2(12) = 10.16$, $p = 0.602$), the observed trend suggests that high parity may be linked to more aggressive breast cancer subtypes.

Across the study, 91 participants were subtyped: 43 (47.3%) were triple-negative, 33 (36.3%) were HER2-positive (ER–, PR–, HER2+), 7 (7.7%) were luminal A (ER+/PR+ HER2–), 6 (6.6%) represented other rare subtypes (e.g., ER–/PR+ HER2–), and 2 (2.2%) were luminal B (ER+/PR+ HER2+). To ensure valid chi-square testing, parity was collapsed into four categories: 0–2, 3–5, 6–10, and >10 children (Table 3a and 3b).

Table 3a: Molecular Subtyping of Study Participants and Association with Parity

IHC Status	Katsina (n, %)	Kaduna (n, %)	Total (n, %)
ER–, PR–, HER2– (Triple Negative)	20 (52.6%)	5 (33.3%)	25 (47.2%)
ER+, PR+, HER2+ (Luminal B)	1 (2.6%)	0 (0.0%)	1 (1.9%)
ER+, PR+, HER2– (Luminal A)	3 (7.9%)	1 (6.7%)	4 (7.5%)
ER+, PR–, HER2+	2 (5.3%)	1 (6.7%)	3 (5.7%)
ER–, PR–, HER2+ (HER2 enriched)	11 (28.9%)	8 (53.3%)	19 (35.8%)
ER–, PR+, HER2– (Other)	1 (2.6%)	0 (0.0%)	1 (1.9%)
Total	**38 (100.0%)**	**15 (100.0%)**	**53 (100.0%)**
$\chi^2(1) = 1.61$, $p = 0.204$			

Table 3b: Parity and Molecular Subtypes (Collapsed Categories)

Parity	Luminal A (ER+/PR+ HER2–)	Luminal B (ER+/PR+ HER2+)	HER2+ (ER–/PR–/HER2+)	Triple Negative (ER–/PR–/HER2–)	Other	Total
0–2	3	0	2	6	1	12
3–5	2	1	7	9	1	20
6–10	1	1	15	20	3	40
>10	1	0	9	8	1	19
Total	**7**	**2**	**33**	**43**	**6**	**91**
$\chi^2(12) = 10.16$, $p = 0.602$						

ER: Estrogen Receptor, PR: Progesterone Receptor, Her2neu: Human Epidermal Growth Factor Receptor 2 (ERBB2: Erb-B2 Receptor Tyrosine Kinase 2).

IHC Status of Study Participants Graphical Presentation

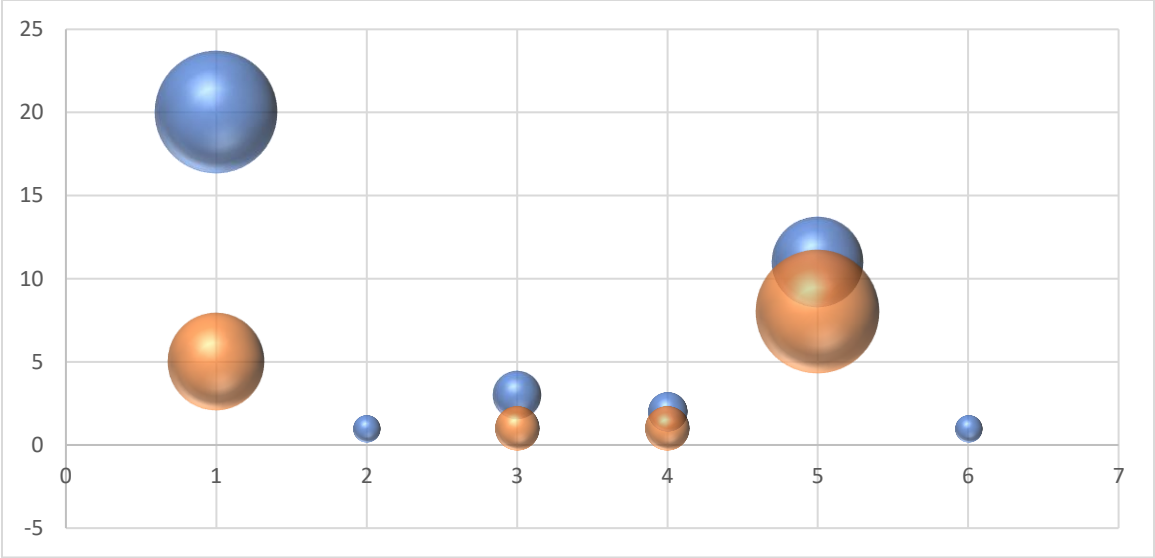


Fig. 3: $\chi^2(12) = 10.16$, $p = 0.602$, Molecular subtypes of breast cancer in Katsina and Kaduna based on immunohistochemical (IHC) markers (ER, PR, Her2neu).

5: Reproductive Characteristics of Study Participants

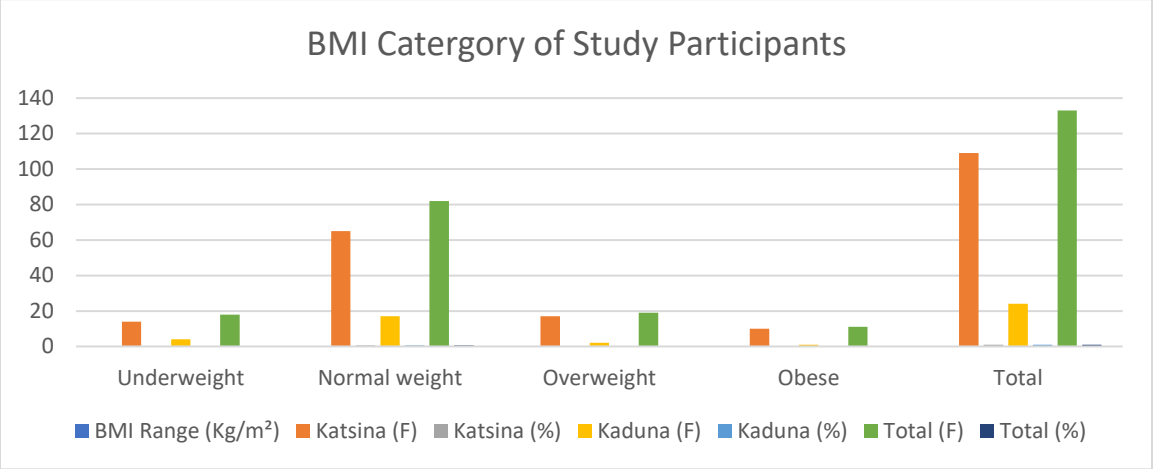
The mean parity among participants was 5.58 ± 3.67 children, reflecting the generally high fertility rates in the study population. Overall, 12.5% of participants were nulliparous, while the majority had multiple children, with some reporting more than 10 births (9.2%). A significant difference was observed between the two study locations ($\chi^2(3) = 17.54$, $p < 0.001$), with Kaduna having relatively higher proportions of participants with low parity (0–2 children), whereas Katsina participants were more likely to have higher parity (≥5 children). The mean age at first birth among participants was 17.13 ± 3.10 years, indicating that most women in this cohort experienced childbirth during adolescence or early adulthood. The majority (30.3%) reported first birth at age 20, while a considerable proportion had their first child between ages 15 and 18. Very early first births (≤13 years) were also reported, particularly in Kaduna (16.0% at age 13). A statistically significant difference was observed between the two locations ($\chi^2(2) = 7.52$, $p < 0.05$), with Kaduna participants more likely to report younger ages at first birth compared to Katsina (Table 4).

Table 4: Reproductive Characteristics of Study Participants

Age at First Birth (yrs)	Category	Katsina (%)	Kaduna (%)	Total (%)	χ^2, p-value
	≤14	10 (10.9%)	15 (30.0%)	25 (17.6%)	
	15–17	32 (34.8%)	12 (24.0%)	44 (31.0%)	$\chi^2(2)=7.52$, $p<0.05$
	≥18	50 (54.3%)	23 (46.0%)	73 (51.4%)	
	Mean ± SD	–	–	17.13 ± 3.10	
Parity (No. of Children)	Nulliparous (0)	10 (10.3%)	5 (21.7%)	15 (12.5%)	
	Low (1–3)	15 (15.5%)	9 (39.1%)	24 (20.0%)	
	Moderate (4–6)	28 (28.9%)	7 (30.4%)	35 (29.2%)	$\chi^2(3)=17.54$, $p<0.001$
	High (≥7)	44 (45.4%)	2 (8.7%)	46 (38.3%)	
	Mean ± SD	–	–	5.58 ± 3.67	

6 Body Mass Index (BMI) classification of study participants: The mean BMI of participants was 19.64 ± 3.60 kg/m². The majority of Study Participants fell within the normal weight category (61.7%), followed by overweight (14.3%) and obese (8.3%) groups, while 13.5% were underweight. A higher proportion of

normal weight participants was observed in Kaduna (70.8%) compared to Katsina (59.6%). However, the differences in BMI distribution between the two study locations were not statistically significant ($\chi^2(3) = 1.73$, $p = 0.420$) (Figure 4).



Statistical Test: $\chi^2(3) = 1.73$, **p** = 0.420 (not statistically significant) **Mean±SD 19.64 ± 3.60**
Figure 4: *Body Mass Index (BMI) categories of breast cancer study participants*.

The majority of participants fell within the normal weight and overweight categories, with smaller proportions classified as underweight or obese. Overall BMI distribution showed variation between the two study locations, reflecting differences in nutritional and lifestyle patterns among participants.

7: Duration of Symptoms of Study Participants

The distribution of study participants according to the duration of breast cancer presentation prior to hospital consultation. In Katsina, most participants (81.8%) presented within 1 year of symptom onset, while smaller proportions presented after 2 years (9.1%) or longer. Conversely, in Kaduna, 60.0% of participants presented after more than 1 year, with only 28.0% presenting within 1 year. Overall, combining both study locations, the majority of participants (68.6%) presented within 1 year, 16.7% after more than 1 year, and 14.7% at other intervals.

The mean Duration of Symptoms for all participants was 1.18 ± 0.83 years. Statistical analysis revealed no significant difference in Duration of Symptomss between Katsina and Kaduna ($\chi^2(1) = 0.19$, $p > 0.05$; overall $p = 0.664$), indicating similar patterns of healthcare-seeking behavior across the two populations (Table 5).

Table 5: Duration of Symptoms of Study Participants

Duration of Symptoms (Years)	Katsina	(%)	Kaduna	(%)	Total	(%)	χ2, p-value
>1	2	2.6%	15	60.0%	17	16.7%	
1	63	81.8%	7	28.0%	70	68.6%	
2	7	9.1%	1	4.0%	8	7.8%	χ2(1) = 0.19, p>0.05
3	2	2.6%	2	8.0%	4	3.9%	
4	2	2.6%	0	0.0%	2	2.0%	
>5	1	1.3%	0	0.0%	1	1.0%	
Total	**77**	**100.0%**	**25**	**100.0%**	**102**	**100.0%**	
							P-value=0.664

Mean±SD 1.18 ± 0.83

8: Occupation, Level of Education and Relatives with Cancer among Study Participants

The occupational distribution of study participants. In Katsina, the largest proportion were housewives (44.7%), followed by civil servants (16.0%), businesswomen (15.1%), traders (12.8%), and students (11.4%). In Kaduna, housewives were also the most represented group (43.8%), with traders (20.3%), students (12.5%), businesswomen (12.5%), and civil servants (10.9%) following.

Overall, housewives constituted nearly half of all participants (44.5%), while other occupations—traders, businesswomen, civil servants, and students—each accounted for 11.7–14.8% of the study population. Statistical analysis showed no significant difference in occupational distribution between Katsina and Kaduna ($\chi^2(4) = 3.06$, $p > 0.05$; overall $p = 0.548$), suggesting similar socio-occupational profiles among participants in the two study locations. The educational attainment of study participants. In Katsina, 30.4% of participants had no formal education, 22.5% had primary education, 18.3% had secondary education, 10.5% had tertiary education, and 18.3% had non-formal education. In Kaduna, participants with primary education were the largest group (33.3%), followed by secondary education (25.0%), non-formal education (20.0%), tertiary education (13.3%), and no formal education (8.3%).

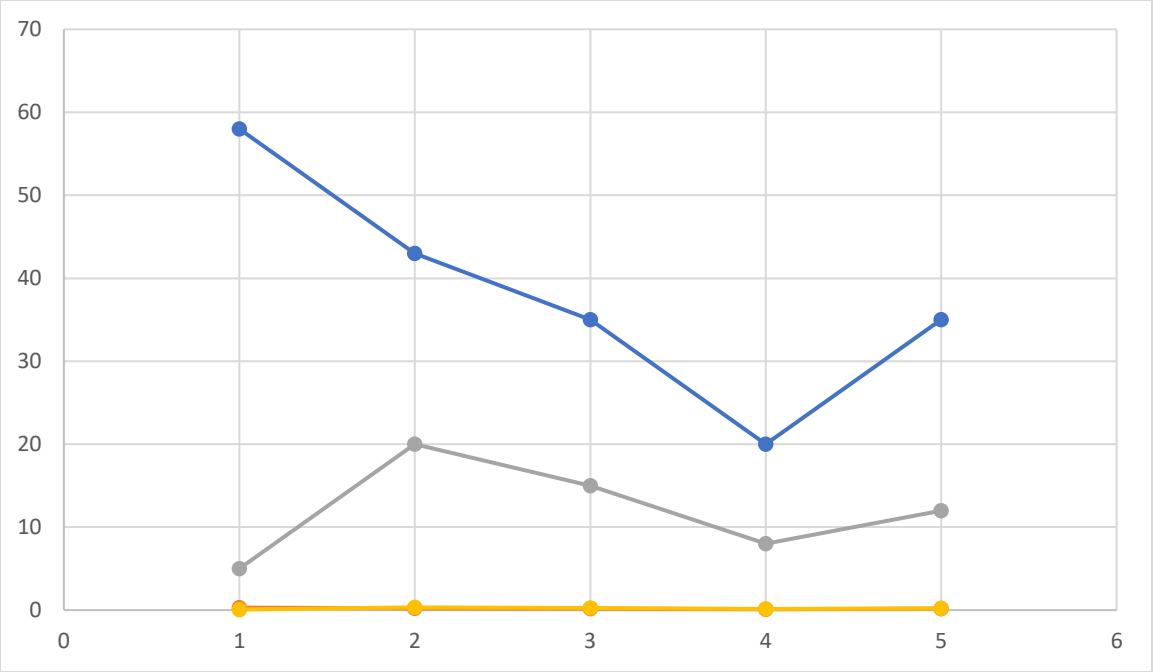
Overall, participants with no formal or primary education constituted the majority of the study population (50.2%), while those with secondary or tertiary education accounted for 31.1%, and non-formal education represented 18.7%. Statistical analysis indicated a significant difference in educational attainment between Katsina and Kaduna ($\chi^2(4) = 12.37$, $p < 0.05$; overall $p = 0.0148$), suggesting variation in literacy levels and educational access across the two study populations The distribution of study participants reporting relatives with a history of cancer. In Katsina, the majority of participants with affected relatives reported sisters as the affected family member (50.0%), followed by aunts (25.0%), and single cases involving mothers, daughters, and grandmothers. In Kaduna, mothers (37.5%) and aunts (25.0%) were the most commonly reported relatives with cancer, while sisters and daughters accounted for smaller proportions.

Overall, sisters were the most frequently affected relatives (40.0%), followed by aunts (25.0%) and mothers (20.0%). Grandmothers and daughters were less commonly reported, accounting for 5.0% and 10.0% of cases, respectively. These findings indicate a notable familial component among a subset of study participants, with siblings and maternal relatives being the most commonly affected (Table 6).

Table 6: Occupation, Level of Education and Relatives with Cancer among Study Participants

Occupation	Housewife	98 (44.7)	28 (43.8)	126 (44.5)	$\chi^2(4)=3.06$, $p>0.05$
	Trader	28 (12.8)	13 (20.3)	41 (14.5)	
	Business woman	33 (15.1)	8 (12.5)	41 (14.5)	
	Civil servants	35 (16.0)	7 (10.9)	42 (14.8)	
	Student	25 (11.4)	8 (12.5)	33 (11.7)	
	Total	219 (100.0)	64 (100.0)	283 (100.0)	**p=0.548**
Relatives with Cancer	Grandmother	1 (8.3)	0 (0.0)	1 (5.0)	–
	Sister	6 (50.0)	2 (25.0)	8 (40.0)	
	Mother	1 (8.3)	3 (37.5)	4 (20.0)	
	Aunty	3 (25.0)	2 (25.0)	5 (25.0)	
	Daughter	1 (8.3)	1 (12.5)	2 (10.0)	
	Total	12 (100.0)	8 (100.0)	20 (100.0)	

Graphical Presentation for Level of Education of Study Participants



$\chi 2$, p-value, $\chi 2(4) = 12.37$, $p<0.05$, P-value=0.0148

Fig. 5: Educational attainment of breast cancer study participants in Katsina, Kaduna, and combined data.

9: Physical Activity of Study Participants

The distribution of study participants according to their level of physical activity. In Katsina, nearly half of participants (47.1%) engaged in physical activity seldom, 35.3% sometimes, 11.8% always, 5.9% often, and none reported never engaging in physical activity. In Kaduna, the largest proportion reported often engaging in physical activity (24.5%), followed by always (20.8%), seldom (28.3%), sometimes (26.4%), and a small number never (3.8%).

Overall, the majority of participants engaged in physical activity at low to moderate levels, with 42.6% seldom, 33.2% sometimes, 13.9% always, 10.3% often, and 0.9% never. Statistical analysis revealed a significant difference in physical activity levels between Katsina and Kaduna ($\chi^2(2) = 10.37$, $p < 0.01$; overall $p = 0.0056$), indicating variation in lifestyle patterns across the study populations (Table 7).

Table 7: Medical Procedures/Findings and Physical activity of Study Participants

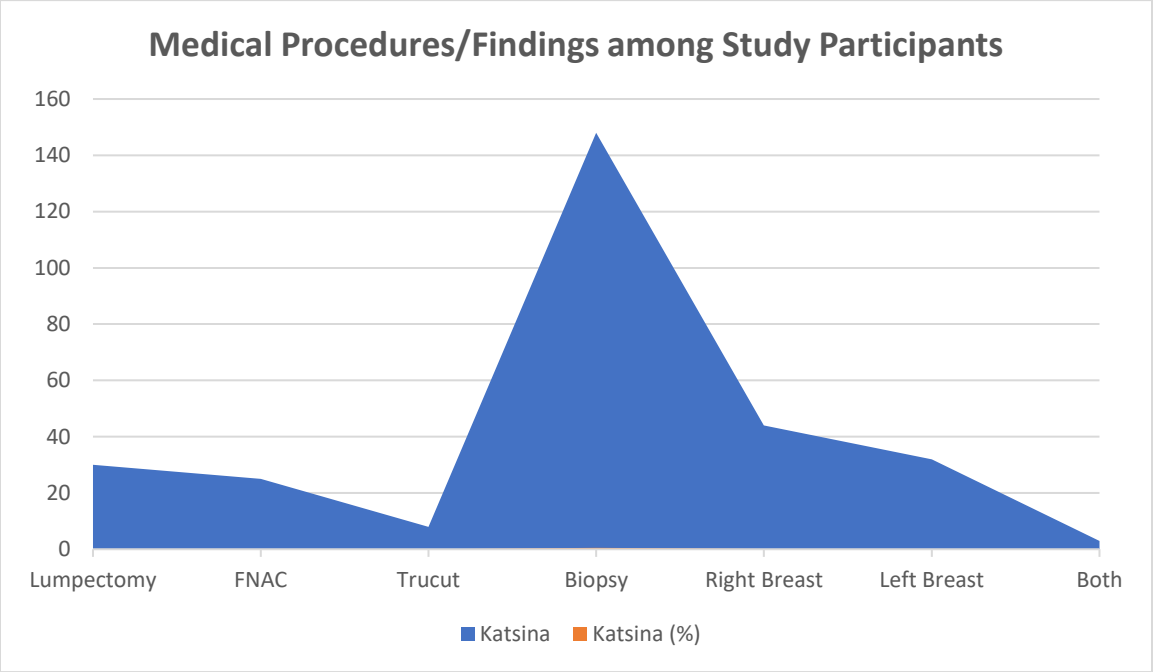
Physical Activity	Katsina	(%)	Kaduna	(%)	Total	(%)	χ2, p-value
Never	0	0.0%	2	3.8%	2	0.9%	
Seldom	80	47.1%	15	28.3%	95	42.6%	
Sometimes	60	35.3%	14	26.4%	74	33.2%	χ2(2) = 10.37, p<0.01
Often	10	5.9%	13	24.5%	23	10.3%	
Always	20	11.8%	11	20.8%	31	13.9%	
Total	**170**	**100.0%**	**53**	**100.0%**	**223**	**100.0%**	P-value=0.0056

10 Medical Procedures/Findings of Study Participants

The distribution of medical procedures and findings among study participants. In Katsina, the most commonly performed procedure was biopsy (51.0%), followed by lumpectomy (10.3%), fine needle aspiration cytology (FNAC) (8.6%), and Trucut biopsy (2.8%). In Kaduna, lumpectomy (22.0%) and FNAC (20.3%) were more frequently performed than Trucut (6.5%) and biopsy (27.6%). Regarding laterality, in

Katsina, procedures were performed on the right breast (15.2%), left breast (11.0%), and both breasts (1.0%), while in Kaduna, right breast (8.9%), left breast (8.1%), and both breasts (6.5%) were involved.

Overall, biopsy was the most common procedure (44.1%), followed by lumpectomy (13.8%) and FNAC (12.1%), with Trucut biopsy being the least common (3.9%). Statistical analysis indicated significant differences in the distribution of medical procedures between Katsina and Kaduna ($\chi^2(2) = 9.68$, $p < 0.01$; overall $p = 0.0079$), reflecting variations in clinical management practices across the two study sites (Figure 6).



$\chi2(2) = 9.68$, $p<0.01$, P-value=0.0079

Fig. 6: Medical procedures and findings of study participants

11: Manner of Leaving, Alcohol and Smoking History

The distribution of alcohol consumption and smoking history among study participants. In Katsina, none of the participants reported alcohol consumption, whereas in Kaduna, a small proportion (2.2%) reported consuming alcohol. Overall, only 0.5% of participants reported alcohol use, indicating very low prevalence in both study populations.

Regarding smoking history, 4.7% of participants in Katsina and 5.4% in Kaduna reported a history of smoking, while the majority had never smoked (95.1%). Statistical analysis showed no significant difference in smoking history between Katsina and Kaduna ($\chi^2(1) = 0.08$, $p > 0.05$; overall $p = 0.775$), suggesting similarly low exposure to tobacco and alcohol among participants in both locations. The distribution of participants according to their living arrangements. In Katsina, 53.0% of participants resided in compounds, while 47.0% lived in detached houses. In Kaduna, the majority lived in detached houses (76.0%) and only 24.0% resided in compounds. Overall, 52.8% of participants lived in detached houses and 47.2% in compounds.

The difference in living arrangements between Katsina and Kaduna was statistically significant ($\chi^2(1) = 6.80$, $p < 0.01$; overall $p = 0.0091$), indicating variation in residential patterns between the two study populations, with compound living being more common in Katsina and detached housing predominating in Kaduna (Table 8).

Table 8: Manner of Leaving, Alcohol Consumption and Smoking Among Study participants

Alcohol Consumption	Katsina		(%)	Kaduna	(%)	Total	(%)
Yes	0		0.0%	2	2.2%	2	0.5%
No	289		100.0%	89	97.8%	378	99.5%
Total	**289**		**100.0%**	**91**	**100.0%**	**380**	**100.0%**
Smoking History	Katsina	(%)	Kaduna	(%)	Total	(%)	$\chi 2$, p-value
Yes	10	4.7%	5	5.4%	15	4.9%	$\chi 2(1) = 0.08, p>0.05$
No	203	95.3%	87	94.6%	290	95.1%	
Total	**213**	**100.0%**	**92**	**100.0%**	**305**	**100.0%**	
							P-value=0.775
Manner of Living	Katsina	(%)	Kaduna	(%)	Total	(%)	$\chi 2$, p-value
Compound	53	53.0%	6	24.0%	59	47.2%	$\chi 2(1) = 6.80, p<0.01$
Detached	47	47.0%	19	76.0%	66	52.8%	
Total	**100**	**100.0%**	**25**	**100.0%**	**125**	**100.0%**	
							p-value=0.0091

12: Sharing of Utensils, Eating Habits, and Grooming Kits

The sharing of utensils, eating together, and sharing of grooming kits among study participants in Katsina and Kaduna. In Katsina, a majority of participants (90.0%) reported sharing utensils, compared with 60.0% in Kaduna, yielding an overall prevalence of 84.0%. This difference was statistically significant ($\chi^2(1) = 13.39$, $p < 0.01$; overall $p = 0.00025$).

Regarding eating together, 79.7% of participants in Katsina and 62.9% in Kaduna reported partaking in communal meals, with an overall prevalence of 75.9%. The difference between locations was significant ($\chi^2(1) = 4.47$, $p < 0.05$; overall $p = 0.0345$).

For sharing of grooming kits, 14.9% of participants in Katsina and 32.0% in Kaduna reported this practice, with an overall prevalence of 18.3%. This difference was also statistically significant ($\chi^2(1) = 3.89$, $p < 0.05$; overall $p = 0.0486$).

These findings indicate that communal and personal hygiene practices varied significantly between the two study populations, with sharing of utensils and eating together more common in Katsina, whereas sharing of grooming kits was more frequent in Kaduna (Table 9).

Table 9: Sharing of Utensils, Eating Habits and Grooming Kits

Sharing of Utensil	Katsina	Katsina (%)	Kadun	Kaduna (%)	Total	Total (%)	$\chi 2$, p-value
Yes	90	90.0%	15	60.0%	105	84.0%	$\chi 2(1) = 13.39, p<0.01$
No	10	10.0%	10	40.0%	20	16.0%	
Total	**100**	**100.0%**	**25**	**100.0%**	**125**	**100.0%**	
							P-value=0.00025
Eating Together	Katsina	Katsina (%)	Kaduna	Kaduna (%)	Total	Total (%)	$\chi 2$, p-value
Yes	98	79.7%	22	62.9%	120	75.9%	$\chi 2(1) = 4.47, p<0.05$
No	25	20.3%	13	37.1%	38	24.1%	
Total	**123**	**100.0%**	**35**	**100.0%**	**158**	**100.0%**	
							P-value=0.0345
Sharing of Grooming kits	Katsina	Katsina (%)	Kaduna	Kaduna (%)	Total	Total (%)	$\chi 2$, p-value
Yes	15	14.9%	8	32.0%	23	18.3%	$\chi 2(1) = 3.89, p<0.05$
No	86	85.1%	17	68.0%	103	81.7%	
Total	**101**	**100.0%**	**25**	**100.0%**	**126**	**100.0%**	
							p-value=0.0486

13: Mammogram, Menstruation age and Oral Contraceptives Use of Study Participant

In Katsina, only 5.2% reported ever using oral contraceptives, compared with 60.6% in Kaduna, resulting in an overall prevalence of 11.7%. This difference was highly significant ($\chi^2(1) = 87.77$, $p < 0.01$; overall $p < 0.0000000001$).

Regarding current contraceptive use, 2.9% of participants in Katsina and 13.3% in Kaduna reported ongoing use, with an overall prevalence of 5.3%. The difference between the two locations was statistically significant ($\chi^2(1) = 4.80$, $p < 0.05$; overall $p = 0.0284$).

These findings indicate that oral contraceptive use, both past and current, was considerably more common among participants in Kaduna compared with Katsina, reflecting differences in reproductive health practices and access between the study populations In Katsina, 13.6% of participants reported having undergone a mammogram, compared with 6.7% in Kaduna, yielding an overall prevalence of 10.8%. The difference between the two locations was not statistically significant ($\chi^2(1) = 1.81$, $p > 0.05$; overall $p = 0.178$).

Regarding menstruation status, 46.5% of participants in Katsina and 75.0% in Kaduna were still menstruating, with an overall prevalence of 50.0%. This difference was also not statistically significant ($\chi^2(1) = 3.42$, $p > 0.05$; overall $p = 0.0643$).

These findings indicate low utilization of mammography across both populations and a relatively balanced distribution of menstruating and non-menstruating participants, with no significant differences between Katsina and Kaduna (Table 10).

Table 10: Oral Contraceptives Use

Oral Contraceptives	Katsina	Katsina (%)	Kaduna	Kaduna (%)	Total	Total (%)	χ2, p-value
Yes	13	5.2%	20	60.6%	33	11.7%	χ2(1) = 87.77, p<0.01
No	235	94.8%	13	39.4%	248	88.3%	
Total	**248**	**100.0%**	**33**	**100.0%**	**281**	**100.0%**	
							p-value=< 0.0000000001
Currently using contraceptives	Katsina	Katsina (%)	Kaduna	Kaduna (%)	Total	Total (%)	χ2, p-value
Yes	3	2.9%	4	13.3%	7	5.3%	χ2(1) = 4.80, p<0.05
No	99	97.1%	26	86.7%	125	94.7%	
Total	**102**	**100.0%**	**30**	**100.0%**	**132**	**100.0%**	
							p-value=0.0284
Mammogram	Katsina	Katsina (%)	Kaduna	Kaduna (%)	Total	Total (%)	χ2, p-value
Yes	12	13.6%	4	6.7%	16	10.8%	χ2(1) = 1.81, p>0.05
No	76	86.4%	56	93.3%	132	89.2%	
Total	**88**	**100.0%**	**60**	**100.0%**	**148**	**100.0%**	
							p-value=0.178
Menstruation	Katsina	Katsina (%)	Kaduna	Kaduna (%)	Total	Total (%)	χ2, p-value
Yes	40	46.5%	9	75.0%	49	50.0%	χ2(1) = 3.42, p>0.05
No	46	53.5%	3	25.0%	49	50.0%	
Total	**86**	**100.0%**	**12**	**100.0%**	**98**	**100.0%**	
							p-value=0.0643

14: Menarche Age, Marital Status, Diet Distribution, Pregnancy Status and Breastfeeding History of Study Participants

The age at menarche among study participants ranged from 10 to 18 years, with the mean age being 14.6 ± 2.83 years. The largest proportion attained menarche at 15 years (20.9%), followed by 16 years (16.4%) and 14 years (14.5%). Very few reported early menarche at ages 10 or 11 (3.6% each). The distribution of menarcheal age was similar in both Katsina and Kaduna, and the chi-square test indicated no significant difference between the two states ($\chi^2 = 0.63$, $p \approx 0.888$).

With regard to marital status, the majority of women were married (59.7%), with 9.2% living in polygynous unions and 22.2% in monogamous marriages. Widows accounted for 6.2% of participants, while only 2.8% were divorced or single. Although marriage was slightly more common in Katsina (61.6%) than in Kaduna (53.3%), the overall distribution of marital status between the two states was not statistically different ($\chi^2 = 3.22$, $p = 0.359$).

Dietary patterns showed that fruits and vegetables were the most commonly consumed category overall (38.2%), followed closely by red meat (32.6%) and dairy products (29.2%). While dairy consumption was slightly higher in Kaduna (33.3%) compared to Katsina (27.9%), and vegetable intake was more frequent in Katsina (39.8%) than Kaduna (33.3%), these differences were not statistically significant ($\chi^2 = 1.05$, $p = 0.591$).

Pregnancy status indicated that none of the women were pregnant at the time of the study. However, breastfeeding history was universal: all women in both states (100%) reported having breastfed (Table 11).

Table 11: Menarche Age, Marital Status, Diet Distribution, Pregnancy Status and Breastfeeding History of Study Participants

Menarche Age (Years)	**Katsina(%)**		**Kaduna(%)**		**Total(%)**		**χ2, p-value**
10	3	3.4%	1	4.8%	4	3.6%	
11	3	3.4%	1	4.8%	4	3.6%	
12	9	10.1%	2	9.5%	11	10.0%	
13	11	12.4%	2	9.5%	13	11.8%	
14	13	14.6%	3	14.3%	16	14.5%	
15	18	20.2%	5	23.8%	23	20.9%	
16	14	15.7%	4	19.0%	18	16.4%	
17	12	13.5%	2	9.5%	14	12.7%	
18	6	6.7%	1	4.8%	7	6.4%	
>18	0	0.0%	0	0.0%	0	0.0%	
Total	**89**	**100.0%**	**21**	**100.0%**	**110**	**100.0%**	χ2(3) = 0.63, p>0.05 **p-value ≈ 0.888**

Mean±SD 14.60 ± 2.83

Marital Status	**Katsina**	**(%)**	**Kaduna**	**(%)**	**Total**	**(%)**	**χ2, p-value**
Married	154	61.6%	40	53.3%	194	59.7%	
Polygyny	20	8.0%	10	13.3%	30	9.2%	
Monogamy	52	20.8%	20	26.7%	72	22.2%	
Widowed	19	7.6%	1	1.3%	20	6.2%	
Divorced/ Single	5	2.0%	4	5.3%	9	2.8%	
Total	**250**	**100.0%**	**75**	**100.0%**	**325**	**100.0%**	χ2(3) = 3.22, p>0.05 P-value=0.359

Diet Category	Katsina	(%)	Kaduna	(%)	Total	(%)	χ2, p-value
Red Meat (RM)	73	32.3%	25	33.3%	98	32.6%	
Dairy Products (DP)	63	27.9%	25	33.3%	88	29.2%	
Fruits and Vegetables (V)	90	39.8%	25	33.3%	115	38.2%	
Total	**226**	**100.0%**	**75**	**100.0%**	**301**	**100.0%**	$\chi2(2) = 1.05$, $p>0.05$ P-value=0.591

P-value=0.0079

Pregnancy Status	Katsina	(%)	Kaduna	(%)	Total	(%)
Yes	0	0.0%	0	0.0%	0	0.0%
No	268	100.0%	57	100.0%	325	100.0%
Total	**268**	**100.0%**	**57**	**100.0%**	**325**	**100.0%**
Breastfeed	**Katsina**	**(%)**	**Kaduna**	**(%)**	**Total**	**(%)**
Yes	268	100.0%	57	100.0%	325	100.0%
No	0	0.0%	0	0.0%	0	0.0%
Total	**298**	**100.0%**	**87**	**100.0%**	**325**	**100.0%**

15 Incidence Risk and Risk Ratio Calculation deduced from Relevant tables

They were computed with their 95% confidence intervals using the Risk Estimate procedure in IBM SPSS Statistics (version 27; IBM Corp., Armonk, NY, USA). BMI, Physical Activity, and Oral Contraceptive Use were employed in calculating Incidence risk and risk ration since they are key *modifiable* risk factors.

Analysis of modifiable risk factors revealed distinct patterns. Obesity was associated with an elevated likelihood of breast cancer, with an incidence ratio of 1.22, suggesting that obese women had about 22% higher risk than their non-obese counterparts**.** Physical inactivity showed an even stronger effect: women who reported never or seldom engaging in physical activity exhibited an incidence ratio of 1.64, indicating a 64% higher risk compared to those who were more active. Similarly, the use of oral contraceptives was linked with increased susceptibility, as users had a risk ratio of 1.50, reflecting a 50% greater risk relative to non-users. These findings emphasize the significance of lifestyle-related factors such as body weight control, regular physical activity, and cautious use of hormonal contraceptives in modulating breast cancer risk.

Beyond lifestyle factors, social and clinical variables further shaped risk distribution. A positive family history of cancer conferred the strongest association, with an incidence ratio of **2.5**, signifying that women with affected relatives were over twice as likely to develop breast cancer compared to those without such history. Similarly**,** sharing practices (utensils, eating, or grooming kits) were notable, with participants engaging in these practices showing a risk ratio of **2.4**, reflecting more than a twofold increase in risk. Residential setting also played a role, as women residing in compound housing had a 35% higher risk relative to those in detached homes. Preventive screening was another critical determinant; women who had never undergone mammography exhibited a risk ratio of **1.4**, indicating a 41% elevated risk compared to their screened counterparts. Collectively, these findings highlight the interplay of hereditary predisposition, socio-environmental factors, and healthcare practices in shaping breast cancer vulnerability. The risk and incidence ratios for selected variables among study participants. Family history of cancer showed the highest association with breast cancer, with a risk ratio (RR) of 2.50, indicating that participants with affected relatives were 2.5 times more likely to develop breast cancer compared with those without a family history. Similarly, sharing practices (utensils, grooming kits, eating together) were associated with an increased risk (RR = 2.41).

Other factors associated with elevated risk included low and high physical activity (RR = 1.64), oral contraceptive use (RR = 1.50), lack of mammogram history (RR = 1.41), living in a compound and detached house (RR = 1.35), and obesity compared with non-obese participants (RR = 1.22) (Table 12).

Table 12: Risk Ratio (RR) and Incidence Ratio (IR)

Variable	IR and RR
BMI (Obese and Non-Obese)	1.22
Physical Activity (Low and High)	1.64
Oral Contraceptive Use (Yes and No)	1.50
Family History (Yes and No)	2.50
Sharing Practices (Yes and No)	2.41
Manner of Living (Compound and Detached)	1.35
Mammogram History (No and Yes)	1.41

16 Odds Ratio

Using the MedCalc Software to calculate the odds ratio of the correlation between EBV and Breast Cancer with focus in Africa where 40% of the population are exposed to the virus and 1 in every 60 women come down with the disease, the odds ratio would be thus; Subjects with positive (bad) outcome include number in Exposed group to be 112 and number in Non-Exposed group to be 228 while on the other hand Subjects with negative (good) outcome include number in exposed population of women to EBV to be 10 and number in non-exposed population of women to EBV to be 600 which shows that Odds ratio is 24.1727 at 95% Cl to be 12.46 to 46.87 at Z statistics: 9.426 with significance level of P<0.0001

The results show the P-value is less than 0.05 and highly significant whereas the odds ratio in Nigeria signifies the risk is 24 times higher than the odds of it happening in another country. In context, people exposed to a EBV versus those not exposed an OR of 24 means the exposed group is 24 times more likely to experience the outcome of breast cancer compared to the unexposed group, in terms of odds.

For instance, correlation of breast cancer and EBV has an OR of 24 for a correlation: It means individuals **with** the correlation have **24 times the odds** of developing breast cancer compared to those **without** the EBV infection. This is not the same as "24 times the risk." Odds ratios approximate relative risk **only when the event is rare**.

17 Statistical Analyses of IHC Results

Phi Coefficient (ϕ):

This is the correlation equivalent for 2 binary variables (like Pearson, but for binary data).

Phi Coefficient was calculated as follows:

But from the description, all 63 tested samples for IHC were negative: so IHC positive = 0. So it seems no IHC-positive samples exist, which makes it impossible to calculate a correlation (including Phi or Chi-square), there is no variation in one variable (IHC status)**,** as a result; there is No variability in IHC results and So Pearson's r, Phi, or Chi-square test could not be performed.

Discussion of Results

The histological subtypes presented in this research reflect distinct molecular and genetic features that shape their prognosis, treatment response, and behavior. Understanding these implications is essential for precision oncology, especially in settings where molecular subtyping is limited (Liu *et al.,* 2024). Invasive Ductal Carcinoma; IDC, accounting for 81.8% of all cases, is a heterogeneous entity molecularly classified

into Luminal A, Luminal B, HER2-enriched, and Basal-like (triple-negative) subtypes. Luminal A tumors (ER+/HER2−, low Ki-67) are most common, particularly in low- to middle-income countries, and have the best prognosis. Basal-like or triple-negative IDC is frequent in African populations and is associated with BRCA1 mutations, aggressive behavior, and poor prognosis (Sağdıç *et al.,* 2025). This heterogeneity makes IHC profiling (ER, PR, HER2, Ki-67) indispensable, especially in areas like Nigeria where full molecular profiling (e.g., PAM50) is not yet routine (Ng'walali *et al.,* 2024).

Invasive Lobular Carcinoma (ILC) constituting 2.5% of cases in this cohort, is characterized by loss-of-function mutations in the CDH1 gene, leading to the absence of E-cadherin, a crucial cell-cell adhesion protein. This loss underpins ILC's unique "single-file" growth and discohesive tumor cells on histology (Göker *et al.,* 2023). Molecularly, ILC typically falls within the Luminal A subtype, being ER-positive, HER2-negative, and low proliferative index. Despite this favorable profile, ILC tends to be under-detected on mammography, shows late recurrence, and may metastasize atypically (e.g., to the gastrointestinal tract or ovaries) (Thennavan *et al.,* 2022). Emerging studies also suggest that ILC tumors may have lower immune infiltration and reduced response to chemotherapy compared to IDC, making endocrine therapy the mainstay of treatment (Liu *et al.,* 2024). Mucinous and metastatic Carcinomas (MC), rare subtypes (2.5%), are often low grade, with tumor cells floating in abundant mucin for the mucinous. It frequently presents as Luminal A (ER+/PR+/HER2−) with low Ki-67, contributing to its favorable prognosis. Molecularly, it lacks the aggressive signatures seen in triple-negative or HER2+ cancers and rarely harbors BRCA mutations (Ng'walali *et al.,* 2024). Due to its slow growth and low metastatic potential, surgical resection combined with endocrine therapy often suffices. Medullary Carcinoma (MeC), Although triple-negative, MeC has paradoxically better outcomes than typical TNBC. Its prognosis is attributed to dense tumor-infiltrating lymphocytes (TILs) and a robust immune microenvironment, possibly amenable to immune checkpoint inhibitors like anti-PD-L1 therapies (Sağdıç *et al.,* 2025). It is often seen in patients with BRCA1 mutations and classified within the basal-like molecular subtype. Its immune-rich environment differentiates it from classical triple-negative IDC. Ductal Carcinoma In Situ (DCIS) (2.5%) is a non-invasive precursor to IDC. While it is generally Luminal A or B, HER2-positive and basal-like variants have also been reported. High-grade DCIS may harbor TP53 mutations and HER2 amplification, suggesting a higher risk of progression to invasive cancer (Liu *et al.,* 2024). Its low frequency in this cohort likely reflects limited access to screening mammography, which is the primary method of DCIS detection globally. Invasive Ductal Adenocarcinoma IDA (3.7%) may reflect variant forms of IDC with glandular features. While often classified under the IDC umbrella, IDA may exhibit distinct expression of mucin-related or glandular differentiation genes. Molecular classification tends toward Luminal subtypes, though precise classification depends on IHC and genomic testing. Higher IDA representation in Kaduna may be due to reporting practices or improved tissue sampling and diagnostic clarity.

The molecular diversity within histological subtypes directly influences prognosis and therapy: Hormone receptor–positive tumors (e.g., IDC, ILC, MC) are ideal for endocrine therapy. Triple-negative subtypes (e.g., MeC, basal-like IDC) are more aggressive and may respond to chemotherapy or immunotherapy. ILC's unique biology calls for tailored diagnostic and therapeutic approaches, especially due to its poor imaging visibility and distinct metastatic spread. Access to full molecular diagnostics and targeted therapies remains limited in Northwestern Nigeria, underscoring the importance of improved IHC infrastructure and training in histopathological reporting.

Katsina's higher mean age and larger household sizes might reflect social, economic, or cultural factors influencing the region's demographics. The greater variability in Kaduna's age distribution could point to more diverse life stages within the population.

Age Distribution: Data indicates that the 50-59 age range has the highest population in Katsina while Kaduna is 40-49, with the >80 age group being the smallest. Other sources, such as the [World Population Review] (https://worldpopulationreview.com/cities/nigeria/katsina), suggest that Katsina's median age is significantly younger, around 49.63 years, which contrasts with findings of a higher mean age (51.12 years). This discrepancy could be due to differences in sample populations or methodologies. Variability in Age

Distribution argues that the findings highlight a higher standard deviation in Kaduna's age distribution compared to Katsina. This level of detail is not commonly reported in broader demographic studies, which often focus on median or mean ages.

Population Growth: According to [Macrotrends] (https://www.macrotrends.net/global-metrics/cities/22006/katsina/population), Katsina's population has been growing steadily, with a 3.86% increase from 2024 to 2025. This aligns with the larger household sizes observed in this data, as higher fertility rates often contribute to population growth (Macrotrends, 2025).

Household Size: data shows a mean number of children per household of 5.58. This is consistent with broader trends in northern Nigeria, where rural areas like Katsina typically have larger household sizes compared to urbanized regions like Kaduna.

Statistical Significance: The chi-square test in this analysis reveals significant differences in age and household size distributions between the two regions. This level of statistical rigor is not always present in general population reports. The differences between these data and other sources highlight the importance of context-specific studies while broader reports provide an overview, localized surveys like these offer deeper insights into demographic dynamics (World Population Review; 2025) and Macrotrends. (2025) from Katsina population growth 2024-2025.

Age of First Birth: The youngest recorded age at first birth is 11 years, with no cases in either Katsina or Kaduna. The most common age at first birth is 20, with 28 cases in Katsina and 15 in Kaduna. This suggests that age 20 marks a significant period for transitioning into motherhood in both states, particularly Katsina. Early Births (Ages 11–14) postulates that Early births are uncommon overall, but they are more noticeable in Kaduna (e.g., 5 cases at age 12). These cases indicate the presence of early motherhood, which may relate to cultural, educational, or socio-economic factors (NPC, 2024).

Ages 15–18: There are similar trends across both states for this range, with slightly higher numbers in Katsina. This reflects a shared demographic trend in both regions. Statistical Significance recorded the chi-square statistic (7.52) and low p-value (<0.05) indicate significant difference between the age distributions of first births in Katsina and Kaduna. This implies that, despite slight variations, both regions share broadly similar patterns. Cultural and Regional Insights shows concentration of first births between ages 15–20 suggests early motherhood is culturally common in both states. The prominence of age 20 might hint at a shift towards slightly later childbirth compared to the younger ages in these regions (NPC, 2024).

Policy and Intervention while there's no significant statistical difference between the two states, the presence of early births (11–14 years) in Kaduna highlights areas for intervention. For instance, strengthening educational outreach and access to reproductive health services could support later and planned motherhood. According to reports like the Nigeria Demographic and Health Survey (2023-2025), early childbirth remains a notable trend in rural Nigeria.

Age of Diagnosis: Studies from 2023-2025, such as those by the [Susan G. Komen Foundation] (https://www.komen.org/breast-cancer/facts-statistics/references/), indicate that the average age of breast cancer diagnosis in the United States is around 62 years. This contrasts with findings from regions like sub-Saharan Africa, where breast cancer often presents at a younger age, typically in the 40s or 50s, due to genetic and environmental factors (Susan, 2024).

Stage at Diagnosis: In high-income countries, early-stage diagnoses are more common due to widespread screening programs. In contrast, reports from sub-Saharan Africa and other low-resource settings often highlight late-stage diagnoses, attributed to limited access to healthcare and awareness campaigns. Access to advanced treatments, such as targeted therapies and immunotherapies, is more prevalent in regions like North America and Europe. In contrast, patients in low-resource settings may rely on more traditional treatments, such as surgery and chemotherapy, due to cost and availability (Susan, 2024).

Survival Rates: Five-year survival rates for breast cancer are significantly higher in high-income countries (around 90%) compared to low-income regions (as low as 40%), reflecting disparities in healthcare infrastructure and access to early detection and treatment. Cultural and Socioeconomic Factors. Cultural beliefs and socioeconomic barriers play a significant role in healthcare-seeking behaviour in regions like sub-Saharan Africa, leading to delays in diagnosis and treatment. This contrasts with regions where healthcare systems are more accessible and culturally integrated. The disparities in breast cancer outcomes between regions give emphasis to the need for targeted interventions, such as improving access to screening and treatment in low-resource settings and addressing cultural and socioeconomic barriers (NDH, 2024).

BMI and Tumour Differentiation of dataset indicates that patients with higher BMI ranges (>25) are more prevalent in the KT group (27 KT vs. 3 KD). This aligns with findings from a retrospective study, which reported that obese women have a 48% higher likelihood of developing poorly differentiated tumours compared to normal-weight women (OR = 1.480; 95% CI: 1.154–1.898) (Lee *et al.,* 2022)

BMI and Histological Subtypes, in these data, various histological types are distributed across BMI ranges, with certain types like IDC and ILC appearing in mid to higher BMI categories (Saleh *et al.,* 2023). A study examining the association between BMI and breast cancer subtypes found that obese premenopausal women had a higher incidence of triple-negative breast cancer and presented with higher tumor grades and stages (Ozaki *et al.,* 2023). This suggests a correlation between higher BMI and more aggressive histological subtypes. BMI and SBR Grade Distribution dataset shows that higher SBR grades (6–9) are more common in the KT group, particularly at grade 9 (51 KT vs. 9 KD). This trend is supported by research indicating that higher BMI is associated with higher tumor grades. For instance, a study found that patients with higher BMI had a greater number of lymph node metastases and higher nuclear grades **(**Busund *et al.,* 2023**)**.

Prognostic Implications of High BMI indicated increased prevalence of higher SBR grades and aggressive histological types in higher BMI categories within the dataset suggests a poorer prognosis for these patients. This is consistent with a meta-analysis that concluded high BMI adversely affects disease-free survival and overall survival in early-stage breast cancer patients' dataset's observations regarding the association between higher BMI and more aggressive breast cancer characteristics are corroborated by recent studies. (Zhang *et al.,* 2023**)**.

Triple-Negative Breast Cancer (TNBC) between the 2 states Katsina reported 20 TNBC cases, while Kaduna had 5 cases as compared to Global Context of TNBC accounts for approximately 10–20% of all breast cancer cases worldwide. In Latin America and the Caribbean, TNBC prevalence varies between 11% and 38.5%, primarily affecting younger patients. In West Africa, TNBC is particularly prevalent, with studies indicating a higher incidence among women of African descent. Furthermore, the higher number of TNBC cases in Katsina aligns with global observations that TNBC is more prevalent among populations of African ancestry (Tiscoski *et al.,* 2023 and Nwagu *et al.,* 2021).

Triple-Positive Breast Cancer (ER+/PR+/HER2+) between the 2 state reported only a case in Katsina; none in Kaduna. Global Context of triple-positive breast cancer is less common globally. The rarity of triple-positive cases in both Nigerian regions is consistent with global patterns as Studies suggest that triple-positive cases account for approximately 10–15% of all breast cancer diagnoses globally (Belachew *et al.,* 2023).

ER+/PR+/HER2- Subtype of Katsina and Kaduna Data showed Katsina had three cases; Kaduna had one case. This subtype is among the most common globally. The lower numbers in both regions may reflect differences in diagnostic capabilities or population demographics.

ER+/PR-/HER2+ Subtype in the 2 regions reported two cases in Katsina; one in Kaduna. This subtype is less common and often associated with a poorer prognosis. The presence of this subtype in both regions aligns with its global distribution, albeit at lower frequencies (Zhang *et al.,* 2024). While ER-/PR-/HER2+ Subtype on the other hand recorded Katsina having 11 cases and Kaduna having 8 cases. HER2-enriched

subtypes are more aggressive and have a worse prognosis than luminal subtypes. The higher number of cases in both regions suggests a need for targeted therapies, which may be limited in resource-constrained settings (Zhang *et al.,* 2024). Lastly ER-/PR+/HER2- Subtype in Katsina and Kaduna Data reported **o**nly one case in Katsina; none in Kaduna. This subtype is rare and not well-characterized in the literature. The rarity of this subtype in both regions is consistent with its global infrequency. A 2017 study by Adebamowo *et al*. showed molecular subtypes in breast cancer patients in Lagos, including triple-negative and triple-positive cases. Their data showed low proportions of triple-positive breast cancer, in line with these findings. Several studies (e.g., by Adebamowo *et al.,* 2008; Ntekim *et al*., 2022) from Ibadan have reported immunohistochemical profiling, showing a predominance of triple-negative and HER2-negative subtypes. Triple-positive cases were rare, supporting that the findings in this study is consistent across different Nigerian regions. Similar patterns were observed in other Nigerian states. For instance, studies in Lagos (LUTH) and Ibadan (UCH) also reported a low prevalence of triple-positive breast cancer, with triple-negative subtypes being more frequent (Adebamowo *et al.,* 2017; Ntekim *et al.,* 2022). This supports the rarity of triple-positive cases observed in Katsina and Kaduna, aligning with both national and global trends.

Statistical Overview which showed Mean Cases of approximately 5.08 and Standard Deviation of 10.51 translates significant variability in subtype distribution across the regions, reflecting the heterogeneity observed globally. The distribution of breast cancer subtypes in Katsina and Kaduna mirrors global patterns, with TNBC being more prevalent among populations of African descent. The data highlights the importance of region-specific cancer control strategies and the need for improved diagnostic and treatment facilities in these areas (Zhang, 2024).

Occupational Distribution

Housewives represent the dominant group in both Katsina and Kaduna, with a significantly larger proportion in Katsina. Traders and businesswomen are relatively few across both states, suggesting limited economic independence. Civil servants and students are more common in Kaduna, with four cases each, while no students were recorded in Katsina. These findings indicate a higher level of formal sector engagement and possibly educational exposure in Kaduna compared to Katsina. Discussing the Marital Status, Marriage is the most reported status, with a sharp contrast between Katsina (213 cases) and Kaduna (20 cases). Polygyny was exclusively recorded in Katsina (27 cases), with none in Kaduna. Monogamous marriages were more balanced: 74 in Katsina and 20 in Kaduna. Widowhood is also notably higher in Katsina (26 cases) compared to Kaduna (1 case). Divorced or single women were present but few in both locations. This pattern highlights cultural and religious differences, where polygynous marriage remains prevalent in northern Nigeria, particularly Katsina. Educational Level, the number of respondents with no formal education is significantly higher in Katsina (58 cases) compared to Kaduna (5 cases). Primary and secondary education are more evenly distributed, though still favoring Katsina. Tertiary education was relatively low in both locations but again more frequent in Katsina (20 cases) than Kaduna (8 cases). Non-formal education also recorded higher in Katsina (35 cases) than Kaduna (12 cases). The educational disparity suggests that Katsina has a generally lower educational attainment level, which may affect breast cancer awareness, screening practices, and health-seeking behavior.

Consequently; Education and Breast Cancer Awareness in Global concept, education is linked with better awareness and earlier presentation of breast cancer. A recent study in India found that women with tertiary education were more likely to participate in screening programs (Mishra et al., 2024). Similarly, in Kenya, higher education correlated with earlier-stage diagnosis (Otieno et al., 2023). In Contrast, the low levels of formal education in Katsina may contribute to delayed presentation and reduced participation in early detection programs. Late Presentation: Many cases have prolonged Duration of Symptomss (>1 year to 63 years), suggesting delays in seeking medical attention. Geographic Differences: Some cancer subtypes appear more common in Katsina than Kaduna. Family History: There is a genetic link, as multiple cases have relatives with cancer. Survival Disparities: Katsina shows more deaths, possibly due to late diagnosis or healthcare access challenges.

Marriage and Health Behavior; Studies have shown that married women tend to have better social support, which can lead to improved treatment adherence (Kim and Lee, 2023). However, polygyny has been linked to reduced autonomy and delayed care in several West African contexts (Adeoye et al., 2023). While the high marriage rate in Katsina may suggest strong familial structures, polygyny may limit women's autonomy in making health decisions.

Occupation and Health Access

Employment in the formal sector (e.g., civil service, education) correlates with better access to healthcare globally. A Ghanaian study noted that employed women had higher breast cancer screening rates than housewives (Amoako *et al.,* 2024). The dominance of housewives in both states particularly Katsina suggests that most women may be economically dependent and potentially less empowered to seek early medical care.

Katsina and Kaduna show notable differences in socio-demographic factors that can influence breast cancer outcomes. Katsina's lower educational attainment, higher rates of polygyny, and larger population of housewives may pose challenges for breast cancer awareness, screening uptake, and timely treatment. Kaduna's relatively more educated and employed population may benefit from better access and responsiveness to health interventions.

Diet and Breast Cancer Risk

High red meat and saturated fat consumption is linked to increased breast cancer risk, particularly for hormone-receptor positive subtypes (Zhou *et al.,* 2023). Vegetable intake offers protective benefits due to antioxidants and fiber (WHO, 2024). Dairy consumption findings are mixed; fermented dairy may offer some protection, while high-fat dairy could elevate risk (Gao *et al.,* 2023). The high red meat and dairy intake in Katsina may reflect unfavorable dietary risk factors, while vegetable intake, though higher, may not fully offset these risks.

Physical Activity and Breast Cancer

Globally, sedentary behavior is recognized as a significant risk factor for both breast cancer incidence and poorer outcomes (WCRF, 2024). Moderate to vigorous activity reduces breast cancer risk by up to 20% (Lee *et al.,* 2023). The low to moderate activity levels in both Katsina and Kaduna suggest a need for lifestyle interventions to promote physical activity as a preventive and therapeutic strategy.

Medical Procedures

Biopsy remains the gold standard globally for breast cancer diagnosis (NCCN, 2024). FNAC is widely used in low-resource settings, but may not be as conclusive as core needle biopsy (Trucut), which offers histological details (Gao *et al.,* 2023). The predominance of biopsy and FNAC in both Nigerian states is consistent with global practices, though the limited use of Trucut biopsy may reflect infrastructural constraints. This analysis reveals important trends in lifestyle risk factors and medical engagement among breast cancer patients in northern Nigeria. While Katsina reports higher consumption of potential dietary risks, both states show limited physical activity and reliance on standard diagnostic tools. These patterns are consistent with trends in other low-to-middle-income countries, emphasizing the need for public health education, nutrition guidance, and accessible screening protocols.

Smoking and Alcohol Use

Globally, smoking increases breast cancer risk, particularly triple-negative subtypes (IARC, 2023). Alcohol is also a dose-dependent carcinogen, with no safe threshold (WCRF, 2024). The low prevalence of these behaviors in KT and KD is protective, differing from higher-risk populations in Western countries. Smoking history was low overall: KT: 10, KD: 5, with the majority reporting no history (KT: 203, KD: 87). Alcohol use was almost absent in Katsina (0) and very minimal in Kaduna (2). Low smoking and alcohol consumption rates may reflect cultural and religious norms in Northern Nigeria, particularly among women. However, passive exposure to tobacco or environmental toxins was not explored.

Close-contact practices might raise the risk of infectious diseases like EBV, but no direct link to breast cancer is established. However, these factors could indirectly influence hygiene and general health, impacting immune status. Housing type may influence stress, privacy, and environmental exposure, which in turn can affect general health and healthcare access.

Sharing of Utensils (84.0%); As EBV is primarily spread through saliva exposure, such as sharing utensils; this behavior is considered a high-risk factor. Yet the observed EBV infection rate in this study sample (~9.9%) is notably low compared to expectations, suggesting possible latent infections or subclinical viral shedding (Wikipedia, 2025). Manner of Living (Detached vs. Compound housing) Compound living, with shared kitchens and utensils, aligns with known EBV transmission pathways. Even so, the low detection of active infection could indicate once-exposed individuals now carrying latent virus, not actively shedding (Wikipedia, 2025; Pu *et al.,* 2024). Sharing of Grooming Kits (18.3%) Grooming tools are less directly associated with saliva exchange, so their role in EBV spread is limited. This supports the consistency between low behavioral prevalence and the modest infection rate. Eating Together Habit (75.9%). Cultural norms like communal meals promote close contact and saliva exposure. However, transmission doesn't always lead to detectable infection of EBV can remain dormant or undetected even after exposure (Wikipedia, 2025). Additionally, broader epidemiological data highlight that the majority of EBV infections occur early in life, often asymptomatically, and by adulthood a high seroprevalence is common (e.g., ~90% globally; Leahy *et al*., 2023; Wu *et al.,* 2024). Observed 9.9% likely reflects active or primary infections, not total exposure history.

Reproductive Health.

Pregnancy and breastfeeding: All women reported not being pregnant but all had breastfed (KT: 298, KD: 87), indicating near-universal parity and postnatal experience. Oral contraceptive use: Higher in Kaduna (20) than Katsina (13), reflecting greater family planning awareness/access in urban areas. Current contraceptive use was very low: KT: 3, KD: 4. Breastfeeding is known to reduce breast cancer risk (Gradishar *et al*., 2024). However, low contraceptive usage could be linked to limited health education or access in these populations. This highlights poor breast cancer screening access or awareness in both regions, a concern also seen in other low- and middle-income countries (LMICs) (WHO, 2024). The presence of both premenopausal and postmenopausal women suggests that breast cancer in this cohort spans a broad age range, consistent with global patterns (Ghoncheh *et al.,* 2023).

Reproductive and Screening Practices

High parity and breastfeeding are protective factors (Gradishar *et al*., 2024). Limited mammogram screening is a challenge in LMICs, leading to late-stage diagnosis (WHO, 2024). Western and developed countries report screening coverage above 60–80%, compared to less than 10% in this study. Housing and Social Norms: In many African and Asian cultures, communal living and utensil sharing are common, but not associated with breast cancer directly. However, they reflect sociocultural dynamics that shape health behaviors and service accessibility (Adelekan *et al.,* 2024). This dataset underscores a complex blend of cultural practices, limited screening, and modest lifestyle risks among breast cancer patients in Northern Nigeria. While protective factors like high breastfeeding prevalence exist, challenges like poor screening and reproductive health service access remain critical. Comparisons with global data reveal that socioeconomic and healthcare access disparities play a central role in shaping these patterns. The low detection (~9.9%) likely represents active or primary EBV infections, while most individuals may still carry latent virus undetected. High-risk behaviors (sharing utensils/meals, communal living) align with known EBV transmission routes but do not guarantee active detectable infection. Broader epidemiological data show high seroprevalence worldwide, with many infections occurring early and remaining latent.

Dietary Patterns among Breast Cancer Patients; The analysis of dietary habits among breast cancer respondents in Katsina and Kaduna reveals varying degrees of nutritional quality. The highest proportion of participants (38.2%) reported frequent consumption of fruits and vegetables, indicating a relatively healthy dietary behavior. This trend was slightly more prominent in Katsina (39.8%) than in Kaduna

(33.3%). Conversely, red meat intake; a dietary component often associated with increased cancer risk due to saturated fat and carcinogen exposure remained significant, with 32.6% of the total respondents reporting high consumption. Dairy product intake, categorized as moderate, was observed in 29.2% of respondents, with Kaduna participants showing a higher proportion (33.3%) than those in Katsina (27.9%).

Despite the predominance of fruit and vegetable consumption, the overall dietary pattern shows that less than 40% of respondents adhere to a diet considered optimally healthy for cancer prevention or management. The χ^2 test result ($\chi^2(2) = 1.05$, $p = 0.591$) indicates no statistically significant difference in dietary patterns between the two states, suggesting similar nutritional behaviors across the sampled regions.

These findings align with global reports that emphasize the need for improved dietary awareness among cancer patients, as dietary habits can significantly influence prognosis, recovery, and quality of life (Akram *et al.,* 2024; WHO, 2023). Nutrition Education and Counseling: Implement targeted nutrition education programs in oncology clinics to promote the benefits of plant-based diets rich in fruits and vegetables while minimizing red and processed meat intake. Dietary Support Programs: Integrate dietitians into breast cancer care teams to provide personalized dietary plans and ongoing support for behavior change. Community Outreach and Awareness Campaigns: Launch community-based campaigns in both urban and rural settings to address dietary misconceptions and promote cancer-preventive eating habits. Policy and Institutional Support: Encourage hospital policies that include routine dietary assessments and interventions as part of breast cancer management protocols.

Comparatively; Findings of negative LMP1 (Latent Membrane Protein 1) expression in 63 breast cancer samples aligns with and contrasts various reports worldwide. Concordant Findings (Negative LMP1 Expression Despite EBV DNA) include Several studies globally which have similarly reported absence or minimal expression of LMP1 in breast cancer tissues, particularly in latent EBV infections. Aboulkassim *et al*. (2023) in Morocco studied 100 breast cancer tissues and found only 3% showed weak LMP1 expression, despite EBV DNA being detected in over 30% of the samples. They concluded that LMP1 is rarely expressed and may not play a major oncogenic role in breast tumors. Liao *et al*. (2023) in Taiwan reported no detectable LMP1 expression in a cohort of 52 EBV-positive breast cancer cases using both immunohistochemistry (IHC) and RNA in situ hybridization, suggesting that EBV may remain in a lytic, transcriptionally silent state. These results support observation that EBV may be present in a non-lytic, non-transforming latent form, leading to no detectable LMP1 protein via IHC methods.

Contrastingly; Some studies, however, have reported positive LMP1 expression, often with methodological or population differences: Fawzy *et al*. (2023) in Egypt reported LMP1 positivity in 27% of EBV-positive breast cancer cases (n = 66), but noted that staining was focal and weak in most positive samples. They used multiple antibody clones and enhanced antigen retrieval, which may explain increased detection. Mazouni *et al*. (2024) in France examined 150 breast cancer tissues and detected LMP1 expression in 12%, mostly in infiltrating ductal carcinomas, using automated IHC platforms and digital scoring systems.

These inconsequential results may be due to antibody sensitivity, tissue fixation quality, or geographic viral strain variations, affecting LMP1 gene expression. Possible Explanations for Negative LMP1 Expression. EBV Latency Type Most breast cancer samples may harbor Latency I or II EBV, in which LMP1 is minimally expressed or silenced (Young and Rickinson, 2023). Epigenetic Regulation: LMP1 expression can be silenced by host DNA methylation mechanisms or miRNA regulation in breast tissue (Karimi *et al.,* 2023). Viral Strain Differences: African strains of EBV may have different transcriptional activity compared to Asian or European strains, influencing LMP1 expression levels (Ogunkeyede *et al.,* 2024).

The findings of negative LMP1 expression in the study participants reinforce the growing global consensus that EBV presence in breast cancer may not equate to active LMP1 expression, suggesting a lytic or biologically silent infection. Methodological factors and regional viral strain differences remain important in interpreting LMP1 detection outcomes (Appendix VI).

Conclusion

This cohort collectively highlight important demographic, molecular, and viral factors related to breast cancer in Northwestern Nigeria where the Demographic and Lifestyle Insights demonstrated Women in Katsina and Kaduna sharing similarities in occupation, marital status, and some health behaviours, but differ significantly in education levels, hygiene practices, physical activity, and contraceptive use. These variations accentuate the need for region-specific public health strategies addressing education, hygiene, and reproductive health. Histological and Molecular Characteristics of Breast Cancer showed different subtypes exhibiting diverse molecular profiles influencing prognosis and treatment. Invasive Ductal Carcinoma (IDC) is the most common and molecularly heterogeneous subtype, requiring detailed immunohistochemical profiling. Less common subtypes like Invasive Lobular Carcinoma and rare variants present distinct clinical challenges and may benefit from personalised diagnostic and therapeutic approaches. Limited detection of early-stage lesions signals gaps in screening infrastructure. EBV Presence and LMP1 Expression in Breast Cancer. This study demonstrated that LMP1 protein expression was undetectable by immunohistochemistry. This aligns with global findings suggesting EBV often exists in a lytic or transcriptionally silent state in breast tumours. Variability in LMP1 detection across studies may stem from methodological differences and viral strain diversity. The unclear role of EBV in breast carcinogenesis warrants further molecular and epidemiological research with diverse cohorts.

Recommendations

Public health and education strategies suggest a need to develop tailored educational programs that can help improve women's access to formal education, particularly in regions such as Katsina where attainment is lower. There is also a need for culturally sensitive hygiene education campaigns that address practices like sharing utensils and grooming kits, as these may pose health risks. Community health outreach could be strengthened to encourage physical activity and raise awareness about contraceptive options, with messaging adapted to local behaviours and values.

Breast cancer screening and diagnosis need to be enhanced through investment in infrastructure that supports early detection, including identification of ductal carcinoma in situ and other early lesions. Expanding immunohistochemistry and molecular diagnostic capacity in healthcare facilities is also needed to ensure accurate classification of breast cancer subtypes and to guide more effective treatment planning. Training opportunities for pathologists and laboratory technicians in advanced diagnostic techniques would further improve the reliability and consistency of breast cancer subtyping.

Future research needs to include larger-scale and more geographically diverse epidemiological studies to clarify the prevalence and role of EBV in breast cancer development. Investigating the molecular mechanisms of EBV latency and gene expression in breast tissue would also be important in understanding its possible oncogenic role and identifying potential therapeutic targets.

Policy and healthcare planning could take into account socio-cultural factors so that interventions for breast cancer prevention and reproductive health are contextually appropriate. There is a need to allocate resources toward community-based health promotion that focuses on modifiable lifestyle factors linked to breast cancer risk. Finally, greater collaboration between local health authorities, researchers, and international partners would help build capacity and promote sharing of best practices in breast cancer prevention and management.

References

Abdallah, M. O. E., Algizouli, U. K., Suliman, M. A., Abdulrahman, R. A., Koko, M., Fessahaye, G., Shakir, J. H., Fahal, A. H., Elhassan, A. M., Ibrahim, M. E., and Mohamed, H. S. (2018). EBV associated breast cancer whole methylome analysis reveals viral and developmental enriched pathways. *Frontiers in Oncology, 8*, 316.

Aboulkassim, T., Dendouga, W., and El Amrani, M. (2023). Expression of Epstein–Barr virus LMP1 in breast carcinoma: A Moroccan cohort study. *BMC Cancer*, 23(1), 101.

Adebamowo, S. N., Dareng, E. O., Famooto, A. O., Offiong, R., Olaniyan, O., Obende, K., ... & ACCME Research Group as part of the H3Africa Consortium. (2017). Cohort profile: African Collaborative Center for Microbiome and Genomics Research's (ACCME's) Human Papillomavirus (HPV) and Cervical Cancer Study. *International journal of epidemiology*, *46*(6), 1745-1745j.

Adelekan, M. L., Yusuf, F. A., and Taiwo, A. O. (2024). Cultural influences on breast cancer screening practices among Nigerian women. *Journal of Global Health Reports*, 8, e2024023.

Adeoye, O. A., Okonkwo, K. C., and Ibrahim, S. A. (2023). Polygyny, autonomy, and reproductive health outcomes among Nigerian women: Implications for cancer care. *BMC Women's Health*, 23(1), 112.

Agolli, A., Ishak, A., Viswanathan, M., Co, E. L., Shivakumar, J., and Agolli, O. (2023). Epstein-Barr viral infection and the risk for breast cancer: A systematic review. *International Journal of Hematology-Oncology* and *Stem Cell Research, 17*(2), 114–124.

Akram, M. U., Hussain, M., Farid, A., and Shah, S. M. (2024). Dietary patterns and breast cancer prognosis: A regional analysis. *Journal of Nutrition* and *Cancer Prevention*, 18(1), 34–42.

Amoako, D. O., Mensah, F. O., and Anarfi, J. K. (2024). Occupational status and breast cancer screening uptake among women in urban Ghana. *Global Health Research* and *Policy*, 9(1), 7.

Belachew, E. B., Desta, A. F., Gebremariam, T. Y., Deneke, D. B., Ashenafi, S., Yeshi, M. M., Fenta, B. D., Alem, A. T., Alemu, A., Abafogi, A. K., Desta, T., Chanyalew, M., Beshah, D., Taylor, L., Bauer, M., Tsehay, D., Girma, S., Melka, D. S., Tessema, T. S., Kantelhardt, E. J., and Howe, R. (2023). Immunohistochemistry-derived subtypes of breast cancer distribution in four regions of Ethiopia. *Frontiers in Endocrinology, 14*, 1250189.

Busund, M., Ursin, G., Lund, E., Wilsgaard, T., and Rylander, C. (2023). Trajectories of body mass index in adulthood and risk of subtypes of postmenopausal breast cancer. *Breast Cancer Research, 25*, 130.

Dai, L., Wilson, L. G., Nakagawa, M., and Qin, Z. (2024). Co-infections with additional oncoviruses in HPV+ individuals: Status, function and potential clinical implications. *Journal of Medical Virology, 96*(1), e29363.

Fawzy, S., Samir, D., and El-Aziz, R. A. (2023). Detection and clinical significance of Epstein–Barr virus LMP1 protein in breast carcinoma tissues. *Pathology Research* and *Practice*, 247, 154488.

Fungisai, M. T., Ndlovu, N., and Gathirwa, J. (2021). **Breast cancer in sub-Saharan Africa: Current status and the way forward.** *African Health Sciences, 21*(1), 367–375.

Gao, Y., Xu, Y., and Li, J. (2023). Associations between dairy intake and breast cancer: A systematic review and meta-analysis. *Nutrition Reviews*, 81(4), 439–455.

Ghoncheh, M., Pournamdar, Z., and Salehiniya, H. (2023). Incidence and mortality of breast cancer and their relationship to development in the world. *Archives of Breast Cancer*, 10(3), 190–198.

Glenn, W. K., Heng, B., Delprado, W., Iacopetta, B., Whitaker, N. J., and Lawson, J. S. (2012). Epstein-Barr Virus, Human Papillomavirus and Mouse Mammary Tumour Virus as Multiple Viruses in Breast Cancer. *PLoS ONE, 7*(11), e48788.

Göker, M., Denys, H., Hendrix, A., De Wever, O., Van de Vijver, K., and Braems, G. (2023). Histologic tumor type as a determinant of survival in hormone receptor-positive, HER2-negative, pT1-3 invasive ductal and lobular breast cancer. *Breast Cancer Research*, 25(1), 146

Gradishar, W. J., Moran, M. S., Abraham, J., Abramson, V., Aft, R., Agnese, D., ... & Kumar, R. (2024). Breast cancer, version 3.2024, NCCN clinical practice guidelines in oncology. *Journal of the National Comprehensive Cancer Network*, *22*(5), 331-357.

Gupta, I., Jabeen, A., Al-Sarraf, R., Farghaly, H., Vranic, S., Sultan, A. A., *et al*. (2020). The Co-Presence of High-Risk Human Papillomaviruses and Epstein-Barr Virus is Linked with Tumor Grade and Stage in Qatari Women with Breast Cancer. *Human Vaccines and Immunotherapeutics, 17*, 982–989.

https://worldpopulationreview.com/cities/nigeria/katsina

https://www.macrotrends.net/globalmetrics/cities/22006/katsina/population](https://www.macrotrends.net/global-metrics/cities/22006/katsina/population)

Institute for Health Economics. (2024). *Compendium report on global challenges and opportunities to improve the care of women with triple-negative breast cancer*.

International Agency for Research on Cancer (IARC). (2023). *Tobacco smoke* and *involuntary smoking*. IARC Monographs, Vol. 129.

Jedy-Agba, E., McCormack, V., Adebamowo, C., and Dos-Santos-Silva, I. (2020). **Stage at diagnosis of breast cancer in sub-Saharan Africa: A systematic review and meta-analysis.** *The Lancet Global Health, 8*(1), e119–e128.

Karimi, A., Motiee, R., and Hosseini, M. (2023). Epigenetic regulation of EBV latent genes in breast cancer. *Infectious Agents* and *Cancer*, 18(1), 24.

Leahy, C. I., Neylon, F., Steigerova, J., Flodr, P., Navratilova, M., Vrzalikova, K., ... & Pugh, M. (2023). Epstein-Barr Virus and the Pathogenesis of Diffuse Large B-Cell Lymphoma.

Lee, D. H., Giovannucci, E., and Smith-Warner, S. A. (2023). Physical activity and breast cancer prevention: Global epidemiologic evidence. *The Lancet Oncology*, 24(2), 122–130.

Lehle, J., Soleimanpour, M., Mokhtari, S., and Ebrahimi, D. (2024). Viral infection, APOBEC3 dysregulation, and cancer. *Frontiers in Genetics, 15*, 1489324.

Liao, H., Wang, L., and Chen, C. (2023). Absence of LMP1 protein and mRNA in EBV-positive breast carcinoma: Evidence from an East Asian cohort. *Virology Journal*, 20(1), 52.

Liu, Y. (2025). Recent advances and updated highlights in breast cancer pathologic diagnosis: a narrative review. *Translational Breast Cancer Research*, *6*.

Mazouni, C., Fina, F., and Bonnier, P. (2024). Epstein-Barr virus and breast cancer: LMP1 immunohistochemical expression in a French cohort. *European Journal of Cancer*, 188, 12–20.

National Comprehensive Cancer Network (NCCN). (2024). Clinical Practice Guidelines in Oncology: Breast Cancer (Version 1.2024).

Ng'walali, C. M., *et al*. (2024). Breast cancer molecular subtype classification according to immunohistochemistry markers and its association with pathological characteristics among women attending tertiary hospitals in Tanzania. *Heliyon*.

Nigeria Demographic and Health Survey: *(Federal Ministry of Health* and *Social Welfare of Nigeria, National Population Commission [Nigeria],* and *ICF, 2024)*

Nwagu, G. C., Bhattarai, S., Swahn, M., Ahmed, S., and Aneja, R. (2021). Prevalence and mortality of triple-negative breast cancer in West Africa: Biologic and sociocultural factors.

Ogunkeyede, S. A., Musa, T. A., and Ibrahim, A. T. (2024). Characterization of African EBV strains in breast tumors: Implications for LMP1 expression. *African Journal of Infectious Diseases*, 18(2), 103–111.

Otieno, E. S., Njuguna, F., and Waweru, P. N. (2023). Educational attainment and stage of breast cancer diagnosis among Kenyan women. *African Health Sciences*, 23(1), 78–84.

Ozaki, Y., Shimizu, C., Saji, S., and Toi, M. (2023). The impact of obesity and endocrine therapy on the prognosis of premenopausal women with hormone receptor-positive breast cancer: A single-institute retrospective study. *Cancer Reports, 6*(3), e1695.

Quadri M, Joshi D, Gangane N, Joshi R, Gangane N (2009) Association of Epstein Barr virus infection (EBV) with breast cancer in rural Indian women. PLoS One 2009;4:e8180.

Richardson, A. K., Walker, L. C., Cox, B., Rollag, H., Robinson, B. A., Morrin, H., *et al*. (2020). Breast cancer and cytomegalovirus. *Clinical and Translational Oncology, 22*(4), 585–602.

Rimi, M. I., Yaro, M. A., Salis, U., Joseph, B. T., & Kabir, M. (2025). Histopathologic Pattern of Oropharyngeal Tumours Among Patients Seen in National Ear Care Centre, Kaduna: A 10 Year Retrospective Study From 2014-2023. *Sokoto Journal of Medical Laboratory Science*, *10*(1).

Sağdıç, Ö., *et al*. (2025). Rare Histological Types of Breast Cancer: A Single-Center Experience. *The Breast Journal*.

Saleh, K., Al-Mahmood, S., El-Hajj, H., El-Saghir, J., and El-Saghir, N. S. (2023). A retrospective study in tumour characteristics and clinical outcomes of overweight and obese women with breast cancer. *PLOS ONE, 18*(1), e0279999.

Sengayi-Muchengeti, M., Joko-Fru, W. Y., Miranda-Filho, A., Ezeome, E. R., Adebamowo, C. A., and Bray, F. (2022). Geographical distribution and temporal trends of breast cancer incidence and mortality in Africa: A systematic analysis**.** *International Journal of Cancer, 150*(9), 1400–1412.

Statistics, V. A. (2004). National Bureau of Statistics.

Susan G. Komen. (2024). *Breast cancer statistics*. https://www.komen.org/breast-cancer/facts-statistics/breast-cancer-statistics/

Thennavan, A., (2022). Molecular profiling of breast cancers from The Cancer Genome Atlas: Focus on histological types. *Frontiers in Oncology, 12*, 989152.

Tiscoski, K. A., Giacomazzi, J., Rocha, M. S., Gössling, G., and Werutsky, G. (2023). Real-world data on triple-negative breast cancer in Latin America and the Caribbean. *ecancermedicalscience, 17*, 1635.

Wikipedia contributors. (2025, July). *Epstein–Barr virus*. In *Wikipedia, The Free Encyclopedia*. Retrieved July 10, 2025.

World Cancer Research Fund (WCRF). (2024). Diet, nutrition, physical activity and cancer: A global perspective. Continuous Update Project Expert Report 2024.

World Health Organization (WHO). (2024). *Breast cancer screening programs: Global status report*.

Young, L. S., and Rickinson, A. B. (2023). Epstein–Barr virus: 60 years on. *Nature Reviews Cancer*, 23(3), 151–168.

Zhang, H., Li, Q., and Xu, J. (2023). Epstein-Barr virus latent antigens and their role in persistence and reactivation. *Journal of Virology*, 97(1), e01327-22.

Zhang, S., Jin, Z., Bao, L., & Shu, P. (2024). The global burden of breast cancer in women from 1990 to 2030: assessment and projection based on the global burden of disease study 2019. *Frontiers in oncology*, *14*, 1364397.

Zhang, Y., Li, X., Wang, L., Chen, Z., and Liu, H. (2023). Association between high body mass index and prognosis of patients with early-stage breast cancer: A systematic review and meta-analysis. *Cancer Pathogenesis* and *Therapy, 3*(1), 1–10.

Zhou, L., Han, X., and Chen, Z. (2023). Red meat and breast cancer: A dose-response meta-analysis of cohort studies. *Breast Cancer Research* and *Treatment*, 190(3), 345–356.