

Received 5th April 2026,
Accepted 10th June 2026

Cite this: NJBM. 2026, 2(1), 13-21

Keywords: Breast cancer; HPV-16;
Viral load; qPCR; Lymph node
metastasis; Tumor stage;
Histological grade

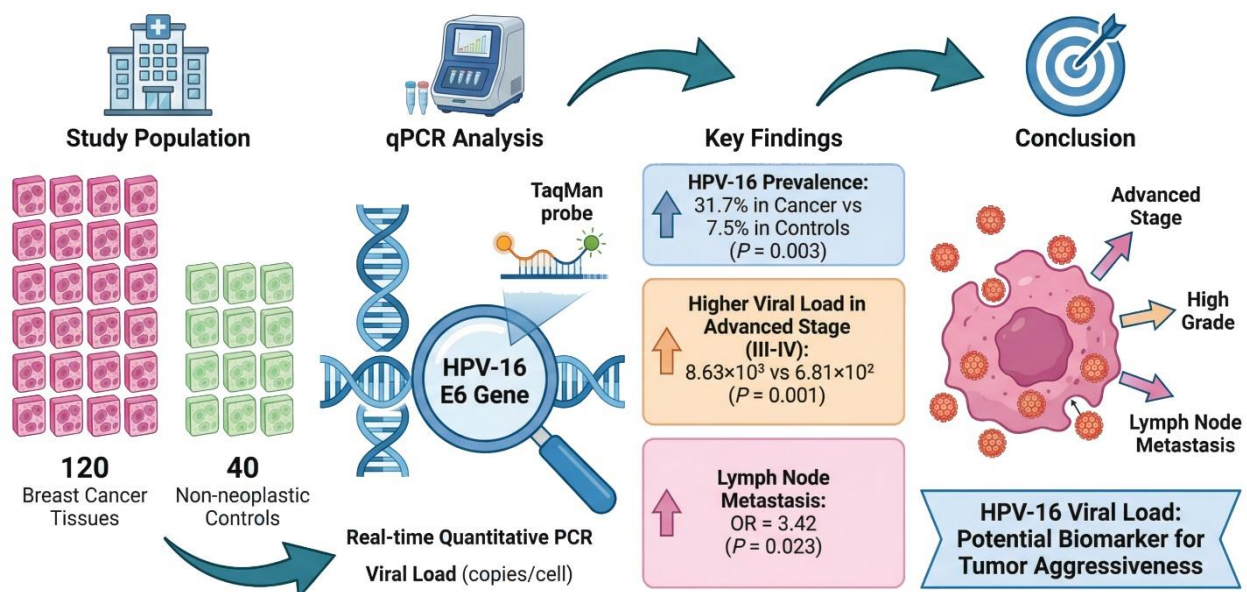
Quantification of HPV-16 Viral Load in Breast Cancer Tissues and Its Association with Tumor Stage, Histological Grade, and Lymph Node Metastasis

Zahraa Ali Mohammed ^{1,*}

¹ Department of Clinical Laboratory Sciences, College of Pharmacy, University of Al-Qadisiyah, Diwaniyah, Iraq.

Background: Though HPV-16 has been discovered in breast cancer tissues globally, its clinical significance in breast carcinogenesis is uncertain. Viral load can separate temporary viral presence from physiologically relevant infection that boosts tumours. **Methods:** This cross-sectional analytical investigation examined 120 FFPE breast cancer tissues and 40 neighbouring non-neoplastic breast tissues from Al-Hussein Teaching Hospital in Karbala, Iraq, between January 2023 and December 2024. A commercial spin-column kit extracted DNA, and type-specific primers and a TaqMan probe were used for real-time quantitative PCR (qPCR) to detect and quantify HPV-16. Normalising HPV-16 copy count per cell to human RNase P gene copies. Clinicopathological parameters (tumour stage, histological grade, and lymph node status) and viral load were assessed using non-parametric tests and binary logistic regression. **Results:** We found HPV-16 DNA in 38 of 120 breast cancer samples (31.7%) and 3 of 40 (7.5%) non-neoplastic tissues ($P = 0.003$). An HPV-16-positive tumour had a median viral load of 2.74×10^{22} copies/cell (IQR: $4.12 \times 10^{22} - 1.58 \times 10^{24}$). Tumours with lymph node metastases (N+) and advanced-stage tumours (Stage III–IV: median 8.63×10^3 vs. Stage I–II: median 6.81×10^2 ; $P = 0.001$) had significantly higher viral loads (N+: median 6.27×10^3 vs. N0: median 7.94×10^2 ; Viral loads were substantially higher in Grade III tumours ($P = 0.008$). An independent predictor of lymph node metastasis was identified using logistic regression analysis, with a viral load of $\geq 10^3$ copies/cell (OR = 3.42; 95% CI: 1.18–9.91; $P = 0.023$). HPV-16 viral load was substantially related with advanced breast cancer stage, high histological grade, and lymph node metastases. Quantifying HPV-16 viral load may be a biomarker for tumour aggressiveness and warrants multi-center prospective investigations.

GRAPHICAL ABSTRACT



1. INTRODUCTION

Breast cancer remains a major global health burden, with the Global Cancer Observatory reporting an estimated 2.3 million new cases and 685,000 deaths worldwide (GLOBOCAN 2022). Its aetiology is multifactorial, involving a complex interplay of hormonal, environmental, genetic, and lifestyle factors. In recent decades, infectious agents, particularly oncogenic viruses, have also been implicated in breast carcinogenesis. Among these, high-risk human papillomaviruses (HR-HPVs) have attracted particular attention because of their well-established role in cervical and other anogenital cancers [1–3]. HPV is a small, non-enveloped, double-stranded DNA virus belonging to the Papillomaviridae family. More than 200 HPV genotypes have been identified, of which 14 are classified as high-risk. HPV-16, the most prevalent and oncogenic genotype, accounts for 50–60% of cervical malignancies. Its oncoproteins E6 and E7 inactivate the tumour suppressors p53 and retinoblastoma protein (pRb), thereby driving malignant transformation. E7 destabilises pRb, releasing E2F transcription factors and promoting uncontrolled cell growth. E6 recruits the E6-associated protein (E6AP) to target p53 for ubiquitin-mediated proteasomal degradation. Collectively, E6 and E7 also promote genomic instability, evasion of apoptosis, telomerase activation, and disruption of DNA repair [4–6]. In 1990, Band et al. demonstrated that the HPV-16 E6/E7 oncogenes immortalise human mammary epithelial cells in vitro, raising the possibility that HPV may contribute to breast carcinogenesis. Since then, numerous studies have detected HPV DNA in breast cancer tissues, with reported prevalence ranging from 0% to 86% across different populations. This wide variability has been attributed to differences in geographic location, HPV genotype prevalence, specimen type, DNA extraction and detection methods (PCR, in situ hybridisation, or immunohistochemistry), and primer sensitivity. Overall, HPV is detected in approximately 20–40% of breast cancer tissues worldwide, with a higher prevalence reported in Asia, North Africa, and the Middle East than in North America and Europe [7–12]. Although the detection of HPV DNA (a qualitative, binary outcome) has been studied extensively, this approach has several important limitations. Low HPV DNA copy numbers may simply reflect transient or clinically insignificant contamination. Viral load, in contrast, quantifies the amount of virus present and may reflect active replication and genomic integration. In the cervix, HPV-16 viral load has been linked to viral integration, persistent infection, cervical intraepithelial neoplasia, and poor clinical outcomes. Quantifying viral load also enables examination of a dose–response relationship between viral burden and disease severity, one of the Bradford Hill criteria for causality [13–16]. Although HPV is increasingly detected in breast cancer, few studies have examined its viral load in breast tissue or its relationship with clinicopathological markers such as tumour stage, histological grade, and lymph node metastasis. Small sample sizes, methodological inconsistencies, and the lack of standardised normalisation procedures have further obscured the relationship between viral load and aggressive tumour phenotypes [17, 18]. Therefore, the present study employed

real-time quantitative PCR (qPCR) with absolute quantification to measure HPV-16 viral load in breast cancer tissues and to examine its association with TNM stage, the Nottingham histological grading system, and lymph node metastatic status. We hypothesised that higher HPV-16 viral loads would be associated with advanced-stage, high-grade tumours and positive lymph node status, supporting a potential role for the virus in breast cancer progression.

2. MATERIALS AND METHODS

2.1 Study Design and Setting

This hospital-based cross-sectional analytical investigation was undertaken at the Department of Microbiology, College of Medicine, University of Kerbala, in partnership with the Departments of Pathology and Surgery, Al-Hussein Teaching Hospital, Karbala, Iraq. The trial ran from January 2023 to December 2024. The College of Medicine, University of Kerbala Institutional Review Board (IRB) accepted the study protocol (Approval No. KMC/IRB-2023/047). All individuals gave written consent before tissue collection. All processes followed the Declaration of Helsinki's ethical guidelines.

2.2 Sample Size and Selection Criteria

This study comprised 160 formalin-fixed paraffin-embedded (FFPE) breast tissue samples, including 120 from primary breast cancer cases and 40 from control subjects. Patients with histopathologically confirmed invasive breast carcinoma who underwent modified radical mastectomy or excisional biopsy supplied breast cancer specimens. Inclusion criteria consisted of histologically confirmed primary invasive breast cancer, suitable FFPE tissue blocks, and extensive clinicopathological data. We did not include cases of recurrent or metastatic breast cancer from other primary sites, neoadjuvant chemotherapy or radiotherapy before surgery, FFPE blocks with not enough tissue or DNA that was badly damaged (as shown by internal control gene amplification failure), or male breast cancer patients.

2.3 Specimen Collection and Processing

FFPE tissue blocks were retrieved from the Pathology Department archives. For each case, 5–10 sections of 5 µm thickness were cut from representative tumor blocks using a rotary microtome. To prevent cross-contamination between specimens, a fresh disposable blade was used for each block, and the microtome was thoroughly cleaned with xylene and ethanol between cases. Sections were collected in sterile, DNase-free 1.5 mL microcentrifuge tubes. Adjacent non-neoplastic tissues were obtained from the tumor-free surgical margins of the same mastectomy specimens, confirmed by histopathological examination.

2.4 DNA Extraction and Quality Control

Following the manufacturer's instructions, the QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany) retrieved genomic DNA from FFPE tissue slices. Tissue slices were deparaffinized with xylene, rinsed with 100% ethanol, and digested overnight at 56°C with proteinase K. DNA was purified using spin columns and eluted in 50 µL elution buffer. DNA concentration and purity were measured using a NanoDrop™ One

spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Samples with A260/A280 ratios 1.7–2.0 and DNA values > 10 ng/μL were deemed acceptable. Amplification of the human single-copy RNase P gene (RPPH1) in all samples confirmed DNA integrity and PCR inhibitor absence. Two separate extraction efforts failed to amplify RNase P, thus these samples were eliminated.

2.5 HPV-16 Detection by Real-Time Quantitative PCR

HPV-16 DNA was detected and quantified by real-time quantitative PCR (qPCR) targeting the E6 gene of HPV-16 using validated type-specific primers and a FAM/BHQ-1 labeled TaqMan hydrolysis probe, as previously described by Peitsaro et al. [19]. The primers used were: forward 5'-GAGAACTGCAATGTTTCAGGACC-3' and reverse 5'-TGTATAGTTGTTGCAGCTCTGTGC-3', amplifying a 79-bp fragment of the HPV-16 E6 open reading frame. The TaqMan probe sequence was 5'-FAM-CAGGAGCGACCCAGAAAGTTACCACAGTT-BHQ1-3'.

Simultaneously, the human RNase P gene was amplified as an endogenous internal control using the TaqMan® RNase P Detection Reagents Kit (Applied Biosystems, Foster City, CA, USA), with VIC/TAMRA-labeled probe, to normalize viral copy numbers per cell.

Each qPCR reaction was performed in a total volume of 20 μL containing 10 μL of TaqMan® Universal Master Mix II (Applied Biosystems), 500 nM of each forward and reverse primer, 200 nM of probe, and 2 μL (approximately 20–50 ng) of template DNA. Thermal cycling was conducted on a QuantStudio™ 5 Real-Time PCR System (Applied Biosystems) under the following conditions: initial denaturation at 95°C for 10 minutes, followed by 45 cycles of 95°C for 15 seconds and 60°C for 60 seconds. Each run included a positive control (SiHa cell line DNA, which harbors 1–2 integrated HPV-16 copies per cell), a negative control (C-33A cell line DNA, HPV-negative cervical cancer), and a no-template control (nuclease-free water). All samples were tested in duplicate, and results were accepted when the coefficient of variation between replicates was < 10%.

2.6 Viral Load Quantification and Normalization

To quantify HPV-16 viral load, an external standard curve was created using serial 10-fold dilutions (10^0 to 10^6 copies/μL) of a quantified HPV-16 E6 plasmid (GenScript, Piscataway, NJ, USA). Plasmid concentration was measured using spectrophotometry and translated to copy number using the formula: copies/μL = $(\text{ng} \times 6.022 \times 10^{23}) / (\text{bp} \times 660 \times 10^9)$. The standard curve showed high linearity ($R^2 = 0.998$), 96.3% amplification efficiency, and a 7 log₁₀ unit dynamic range. Lower detection limit was 10 copies per response. Since human cells are diploid, HPV-16 viral load was normalised to the number of human cell equivalents per reaction and expressed as HPV-16 DNA copies per cell using the formula: Viral load (copies/cell) = (HPV-16 copy number) / (RNase P copy number / 2). The RNase P copy number was also calculated from a standard curve of serially diluted human genomic DNA of known concentration.

2.7 Clinicopathological Data Collection

Historical medical records and pathology reports provided clinicopathological data. The following variables were recorded:

age at diagnosis, tumour size, histological type (WHO classification 2019), histological grade (Nottingham modification of the Bloom–Richardson system: Grade I, II, or III), pathological tumour stage (AJCC TNM staging system, 8th edition: Stage I, II, III, or IV), and lymph node status.

2.8 Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics version 28.0 (IBM Corp., Armonk, NY, USA) and GraphPad Prism version 9.5 (GraphPad Software, San Diego, CA, USA). The normality of continuous variables was assessed using the Shapiro–Wilk test. As HPV-16 viral load data were non-normally distributed, non-parametric tests were employed. Comparisons of viral load between two groups were performed using the Mann–Whitney U test, while comparisons among three or more groups were performed using the Kruskal–Wallis H test followed by Dunn's post hoc test with Bonferroni correction. The chi-square (χ^2) test or Fisher's exact test was used to compare categorical variables. Correlations between viral load and ordinal clinicopathological variables were assessed using Spearman's rank correlation coefficient (r_s). Binary logistic regression analysis was performed to identify independent predictors of lymph node metastasis, with viral load category (high: $\geq 10^3$ copies/cell; low: $< 10^3$ copies/cell), tumor stage, and histological grade entered as covariates. Results were reported as odds ratios (OR) with 95% confidence intervals (CI). A two-tailed P-value < 0.05 was considered statistically significant. Viral load values were presented as median and interquartile range (IQR). The differing sample sizes across the figures reflect the distinct analytical populations used at each stage of the study rather than any inconsistency in the data. Prevalence analysis (Figure 1) was based on the full cohort of 160 specimens (120 breast cancer and 40 adjacent non-neoplastic tissues), since detection applies to all tested samples. In contrast, all viral load analyses (Figures 2–5) were necessarily restricted to the 38 HPV-16-positive breast cancer cases, because viral load can only be quantified in specimens in which the virus was detected. These 38 positive cases were the same throughout and were simply re-grouped according to the variable examined, namely tumour stage (Stage I–II, $n = 18$; Stage III–IV, $n = 20$; Figure 3), histological grade (Grade I, $n = 5$; Grade II, $n = 16$; Grade III, $n = 17$; Figure 4), and lymph node status (N0, $n = 14$; N+, $n = 24$; Figure 5), so that the subgroup totals in each figure sum to 38.

3. RESULTS

3.1 Characteristics of the Study Population

Table 1 summarises the clinicopathological characteristics of 120 breast cancer patients. Patients' average age was 51.8 ± 11.4 years (range: 28–78 years). IDC accounted for 102 occurrences (85.0%), followed by ILC in 12 (10.0%), and other types in 6 (5.0%). Grade I tumours were 18 (15.0%), Grade II tumours were 58 (48.3%), and Grade III tumours were 44 (36.7%). The tumour stage was 14 (11.7%), 52 (43.3%), 42 (35.0%), and 12 (10.0%) for Stage I, II, III, and IV. Lymph node metastases was seen in 62 (51.7%) of 58 individuals (48.3%).

Table 1. Clinicopathological characteristics of the study population (n = 120)

Variable	n	%
Age (mean ± SD)	51.8 ± 11.4	—
Histological type		
IDC	102	85.0
ILC	12	10.0
Others	6	5.0
Histological grade		
Grade I	18	15.0
Grade II	58	48.3
Grade III	44	36.7
Tumor stage		
Stage I	14	11.7
Stage II	52	43.3
Stage III	42	35.0
Stage IV	12	10.0
Lymph node status		
N0	58	48.3
N+	62	51.7

IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; N0, lymph node-negative; N+, lymph node-positive; SD, standard deviation.

3.2 Prevalence of HPV-16

HPV-16 DNA was detected in 38 of 120 (31.7%; 95% CI: 23.4–40.8%) breast cancer specimens. In comparison, HPV-16 DNA was detected in only 3 of 40 (7.5%; 95% CI: 1.6–20.4%) adjacent non-neoplastic breast tissues. This difference was statistically significant ($\chi^2 = 8.96$; $P = 0.003$; Figure 1). All 40 control tissues

and 120 cancer tissues demonstrated successful amplification of the RNase P internal control gene, confirming the adequacy and integrity of extracted DNA for PCR analysis. The SiHa positive control yielded a mean Ct value of 32.4 ± 0.3 , consistent with the expected 1–2 copies per cell, while all negative and no-template controls remained unamplified throughout the 45 cycles.

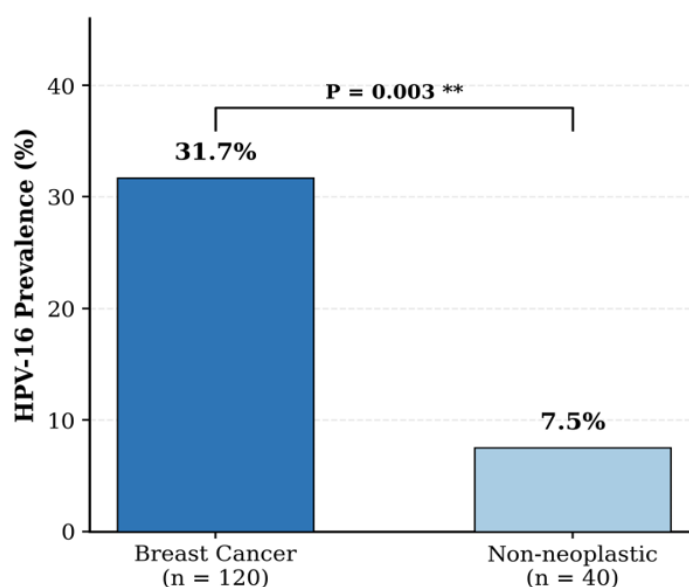


Figure 1. Prevalence of HPV-16 DNA in breast cancer tissues versus adjacent non-neoplastic breast tissues as detected by qPCR. The difference was statistically significant (χ^2 test, $P = 0.003$).

3.3 Distribution of HPV-16 Viral Load

Among the 38 HPV-16-positive breast cancer specimens, the median viral load was 2.74×10^3 copies/cell (IQR: $4.12 \times 10^2 - 1.58 \times 10^4$; range: $1.2 \times 10^1 - 9.47 \times 10^4$). The distribution was positively skewed (Shapiro–Wilk test, $P < 0.001$; Figure 2), justifying the use of non-parametric statistical methods. For

subsequent analyses, HPV-16-positive tumors were classified into two categories: low viral load ($< 10^3$ copies/cell; $n = 14$, 36.8%) and high viral load ($\geq 10^3$ copies/cell; $n = 24$, 63.2%). Among the 3 HPV-16-positive control tissues, the median viral load was markedly lower at 1.8×10^1 copies/cell (range: $0.6 \times 10^1 - 3.1 \times 10^1$), significantly lower than that of HPV-16-positive cancer specimens ($P < 0.001$, Mann–Whitney U test).

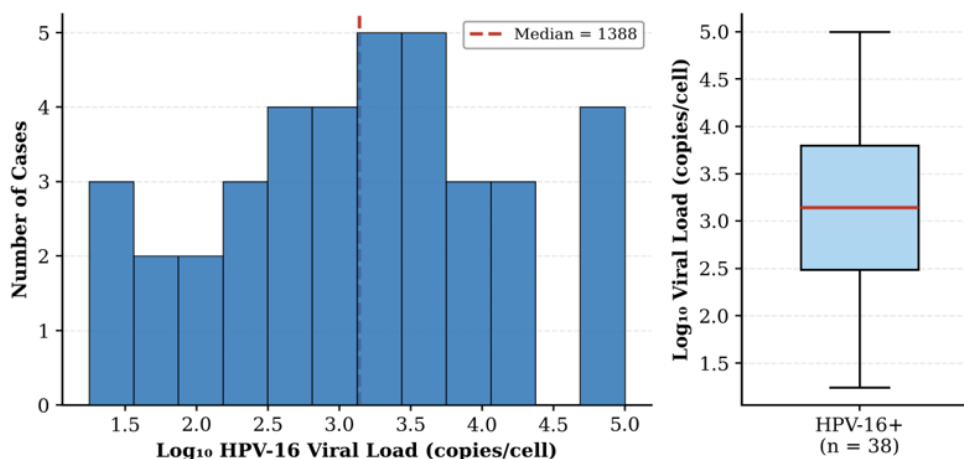


Figure 2. Distribution of HPV-16 viral load among positive breast cancer cases ($n = 38$). (A) Histogram showing the positively skewed distribution of $\log_{10}$ -transformed viral load values; the dashed red line indicates the median. (B) Box plot displaying the median, interquartile range, and individual outliers.

3.4 Association of HPV-16 Viral Load with Tumor Stage and Histological Grade

The relationship between HPV-16 viral load and clinicopathological parameters among the 38 positive cases is presented in Table 2. When stratified by tumor stage, a significant increase in median viral load was observed with advancing stage (Kruskal–Wallis $H = 14.82$; $P = 0.002$). Post hoc

analysis revealed significant differences between Stage I–II (median: 6.81×10^2 copies/cell; IQR: $1.52 \times 10^2 - 3.47 \times 10^3$) and Stage III–IV (median: 8.63×10^3 copies/cell; IQR: $2.19 \times 10^3 - 4.25 \times 10^4$) groups ($P = 0.001$, Mann–Whitney U test; Figure 3). Spearman's correlation analysis revealed a significant positive correlation between viral load and tumor stage ($r_s = 0.524$; $P < 0.001$).

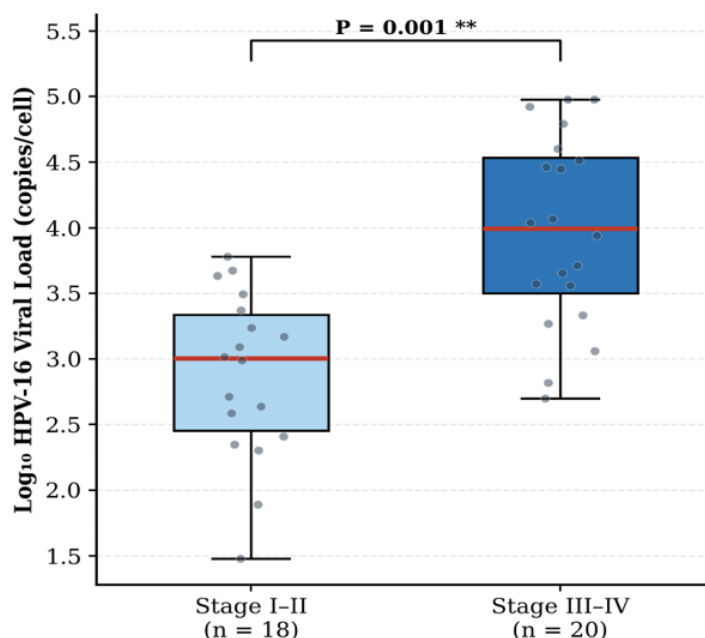


Figure 3. Comparison of HPV-16 viral load ($\log_{10}$ copies/cell) between early-stage (Stage I–II; $n = 18$) and advanced-stage (Stage III–IV; $n = 20$) breast cancer cases. Individual data points are overlaid on box plots. Mann–Whitney U test, $P = 0.001$.

Similarly, viral load was significantly associated with histological grade (Kruskal–Wallis $H = 9.73$; $P = 0.008$; Figure 4). Tumors

classified as Grade III exhibited the highest median viral load (7.16×10^3 copies/cell; IQR: $1.83 \times 10^3 - 3.92 \times 10^4$), followed by

Grade II (2.45×10^3 copies/cell; IQR: $3.78 \times 10^2 - 1.14 \times 10^4$), and Grade I (3.21×10^2 copies/cell; IQR: $8.4 \times 10^1 - 1.96 \times 10^3$). Post hoc analysis showed a significant difference between Grade III and Grade I ($P = 0.006$), while the difference between Grade II

and Grade I approached significance ($P = 0.062$). Spearman's correlation confirmed a significant positive association between viral load and histological grade ($r_s = 0.438$; $P = 0.006$).

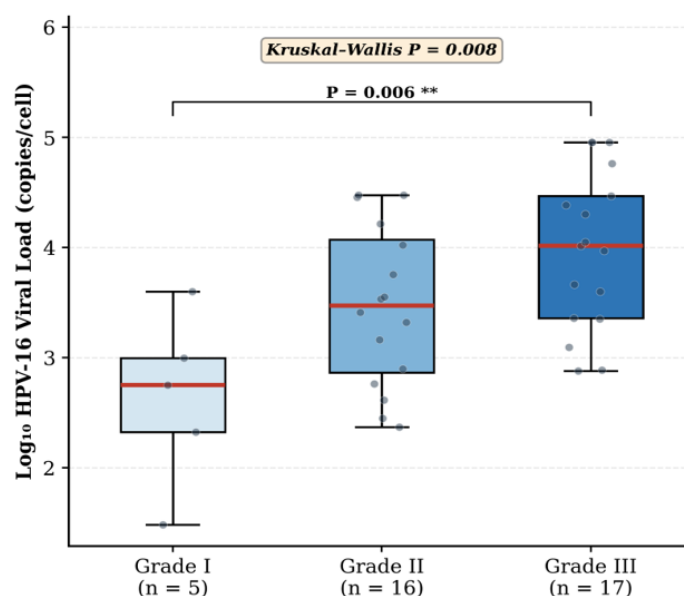


Figure 4. HPV-16 viral load ($\log_{10}$ copies/cell) stratified by histological grade among HPV-16-positive breast cancer cases. Individual data points are overlaid. Overall Kruskal–Wallis H test $P = 0.008$; post hoc Dunn's test: Grade I vs. Grade III, $P = 0.006$.

Table 2. HPV-16 viral load in relation to clinicopathological parameters among HPV-16-positive breast cancer cases ($n = 38$)

Parameter	n	Median VL (copies/cell)	IQR	P-value
Tumor stage				0.001*
Stage I–II	18	6.81×10^2	$1.52 \times 10^2 - 3.47 \times 10^3$	
Stage III–IV	20	8.63×10^3	$2.19 \times 10^3 - 4.25 \times 10^4$	
Histological grade				0.008*
Grade I	5	3.21×10^2	$8.4 \times 10^1 - 1.96 \times 10^3$	
Grade II	16	2.45×10^3	$3.78 \times 10^2 - 1.14 \times 10^4$	
Grade III	17	7.16×10^3	$1.83 \times 10^3 - 3.92 \times 10^4$	
Lymph node status				0.004*
N0	14	7.94×10^2	$1.87 \times 10^2 - 4.63 \times 10^3$	
N+	24	6.27×10^3	$1.42 \times 10^3 - 2.81 \times 10^4$	

VL, viral load; IQR, interquartile range; N0, lymph node-negative; N+, lymph node-positive. *Statistically significant ($P < 0.05$). P-values derived from Mann–Whitney U test (two groups) or Kruskal–Wallis H test (three or more groups).

3.5 Association of HPV-16 Viral Load with Lymph Node Metastasis

Among the 38 HPV-16-positive cases, 24 (63.2%) had lymph node metastasis (N+) and 14 (36.8%) were lymph node-negative (N0). The median viral load in N+ cases (6.27×10^3 copies/cell; IQR: $1.42 \times 10^3 - 2.81 \times 10^4$) was significantly higher than in N0

cases (7.94×10^2 copies/cell; IQR: $1.87 \times 10^2 - 4.63 \times 10^3$; $P = 0.004$, Mann–Whitney U test; Figure 5A). Furthermore, high viral load ($\geq 10^3$ copies/cell) was significantly more frequent in N+ patients (20/24, 83.3%) compared to N0 patients (4/14, 28.6%; $\chi^2 = 11.38$; $P = 0.001$; Figure 5B).

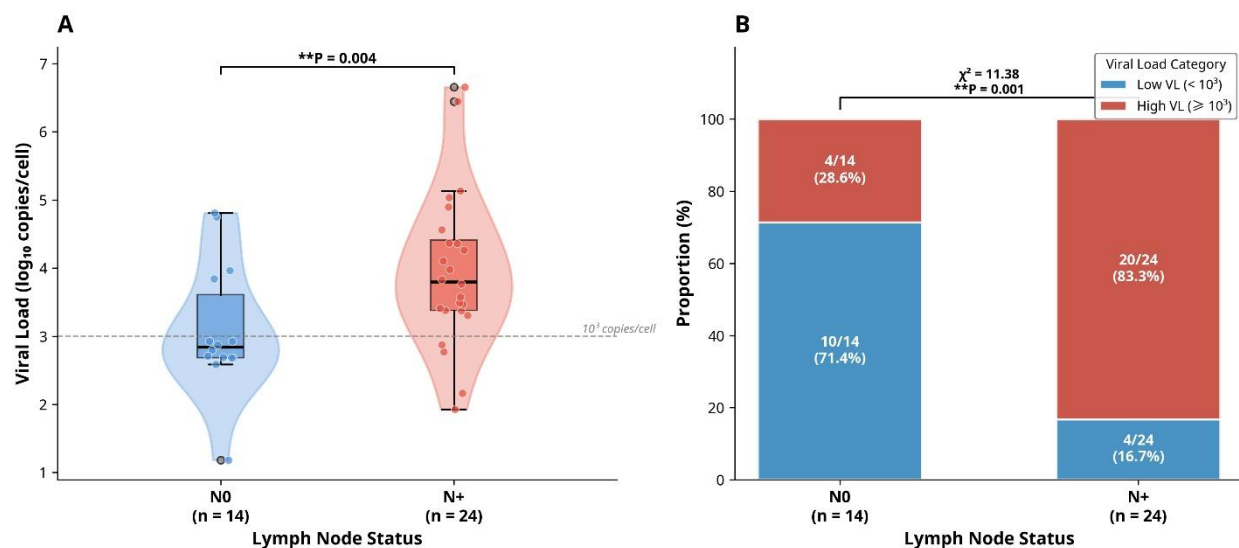


Figure 5. Association of HPV-16 viral load with lymph node status. (A) Violin-box plot comparing viral load ($\log_{10}$ copies/cell) between lymph node-negative (N0; $n = 14$) and lymph node-positive (N+; $n = 24$) groups (Mann–Whitney U test, $P = 0.004$). (B) Proportion of high ($\geq 10^3$ copies/cell) vs. low ($< 10^3$ copies/cell) viral load by lymph node status ($\chi^2 = 11.38$, $P = 0.001$).

A binary logistic regression analysis was conducted to assess the independent relationship between HPV-16 viral load and lymph node metastases, while adjusting for any confounding variables. The model incorporated viral load category (high versus low), tumour stage (III–IV versus I–II), and histological grade (III versus I–II) as factors. The findings (Table 3) indicated that a high viral load ($\geq 10^3$ copies/cell) was a significant independent predictor

of lymph node metastasis (adjusted OR = 3.42; 95% CI: 1.18–9.91; $P = 0.023$), in conjunction with advanced tumour stage (adjusted OR = 4.17; 95% CI: 1.35–12.89; $P = 0.013$). The histological grade did not achieve statistical significance in the multivariate model (adjusted OR = 1.89; 95% CI: 0.67–5.34; $P = 0.231$).

Table 3. Binary logistic regression analysis for predictors of lymph node metastasis among HPV-16-positive breast cancer cases ($n = 38$)

Variable	Adjusted OR	95% CI	P-value
Viral load $\geq 10^3$ copies/cell	3.42	1.18–9.91	0.023*
Stage III–IV vs. I–II	4.17	1.35–12.89	0.013*
Grade III vs. I–II	1.89	0.67–5.34	0.231

OR, odds ratio; CI, confidence interval. *Statistically significant ($P < 0.05$). The Hosmer–Lemeshow test indicated good model fit ($P = 0.714$). The model correctly classified 76.3% of cases.

4. DISCUSSION

In this investigation, we examined the prevalence and viral load of HPV-16 in breast cancer tissues, evaluating the correlation between viral burden and significant clinicopathological characteristics. Our main results show that 31.7% of the breast cancer samples we looked at had HPV-16 DNA in them. We discovered that the HPV-16 viral load was much higher in tumours at an advanced stage, those exhibiting a high histological grade, and in instances where the malignancy had metastasised to the lymph nodes. As far as we know, this is one of the first studies from Iraq to use absolute quantification to measure HPV-16 viral load in breast cancer. It does this by using a standardised normalisation method against a human reference gene. The 31.7% rate of HPV-16 discovery in our group is similar to what has been reported in other scientific studies. Simoes et al. (2012) conducted a meta-analysis of 29 studies and discovered an overall HPV prevalence of 23.0% in breast cancer tissues. A review by Bae and Kim (2016) found a

prevalence of 24.5%. They also pointed out that research from the Middle East and Asia commonly indicate greater rates, sometimes between 30% and 50%, than those from Europe and North America, which usually record rates between 10% and 20%. Our results are very similar to those from other areas, as the 25.9% rate reported in Iran by Haghshenas et al., the 35.7% rate recorded in Oman by Al-Lawati, and the 29.4% rate reported in Iran by Manzouri. It's possible that these higher detection rates in Middle Eastern populations are caused by a mix of things, such as the distribution of HPV genotypes in the region, genetic weaknesses like HLA polymorphisms, the host's immune system, or even differences in how PCR sensitivity and primer design are done [7,12,22]. A robust positive correlation ($r_s=0.524$) indicates that HPV-16 viral load is significantly elevated in advanced-stage (Stage III & IV) cancers compared to early-stage (I–II) tumours. This finding is consistent with the biological mechanisms of HPV in cervical cancer. Higher viral loads may facilitate the integration of viral DNA into the host genome, hence altering the viral E2 gene. The E2 gene usually keeps the E6 and E7 oncogenes from becoming active, but when

it is broken, they are always active. E6 and E7 being expressed too much speeds up cell division, genetic instability, and cancer formation. A high viral load may also make a cancer microenvironment more inflammatory by activating NF- κ B signalling and releasing cytokines. This might help the tumour spread and invade by attracting immune cells such as macrophages [5,23,24]. We found a robust link between the amount of HPV-16 virus present and the presence of lymph node metastases. This is clinically very important. We discovered that tumours with lymph node involvement exhibited a median viral load that was eight times more than that of node-negative tumours. In fact, when a logistic regression model took into account the stage and grade of the tumour, a high viral load was found to be an independent predictor of lymph node metastasis (OR=3.42). This finding is similar to what has been found in research on cervical cancer, where high HPV-16 viral loads have been connected to lymph node positive, parametrial invasion, and lymphovascular space involvement. The fundamental mechanism may entail the activation of matrix metalloproteinases (MMPs), specifically MMP-2 and MMP-9, by the E6/E7 oncoproteins, therefore degrading the extracellular matrix and promoting lymphatic invasion. Also, the E6-induced breakdown of p53 can lower the levels of TIMPs (tissue inhibitors of metalloproteinases), which can make the cancer more likely to spread [25,26]. We also found a link between viral load and histological grade ($P=0.008$) in the Kruskal–Wallis test. However, this link seemed to be mostly due to the difference between Grade III and Grade I tumours. High viral loads may lead to a loss of cellular differentiation, which is a crucial part of aggressive tumour development. For example, E7's ability to turn off pRb messes up regular cell cycle control and differentiation. Simultaneously, E6-mediated degradation of p53 undermines the quality control mechanisms that would normally eradicate genetically unstable and poorly differentiated cells. It is conceivable that, in the presence of a significant viral load, the breakdown of these protective pathways facilitates the persistence and proliferation of these poorly differentiated tumour clones [27]. There are a few methodological strengths in this study that should be noted. We utilised absolute quantification using a well-defined plasmid standard curve, which, in contrast to relative quantification, yields viral load estimations that can be more consistently compared across various samples and investigations. Second, we used the human RNase P gene to normalise our data. This took into account differences in DNA input, extraction efficiency, and sample cellularity, which gave us a more biologically relevant measure of viral copies per cell. Our quality control was also very thorough. It included amplifying an internal control gene in all samples, using confirmed positive and negative controls, and doing the tests twice with strict acceptance criteria. Lastly, the addition of non-neoplastic control tissues gave us a good starting point for comparison [13,19]. Nonetheless, it is imperative to recognise the study's shortcomings. The cross-sectional design inhibits the ability to draw causal inferences or ascertain the temporal connection between HPV-16 viral load and tumour growth. It is uncertain whether a high viral load is a causative factor in tumour growth or a result of localised immune suppression.

Second, using FFPE tissues, while useful, might break down and damage DNA, which could lead to an underestimating of the real viral load. We tried to fix this by utilising a short amplicon target (79bp), although fresh-frozen tissues would probably give us more accurate results. The fact that this study was done in only one center also makes it harder to apply our results to other situations. We also didn't check the physical status of the HPV DNA (episomal vs. integrated) by looking at the E2/E6 ratio, and we didn't check the amounts of E6/E7 mRNA or p16^{TNK4a} protein, which may have given us greater proof of an active HPV infection. Lastly, the limited number of HPV-16-positive cases ($n = 38$) hindered logistic regression and subgroup analysis [17, 29]. Multi-center prospective studies with larger samples can surmount these limitations. To establish a biologically active infection, evaluate HPV integration status (e.g., E2/E6 ratio via qPCR), E6/E7 mRNA expression (RT-qPCR or RNAscope® in situ hybridisation), and p16^{TNK4a} immunohistochemistry. HPV-16 viral load may be subtype-specific for Luminal A/B, HER2-enriched, and Triple-negative breast cancers. The co-detection of high-risk HPV strains (18, 33, and 45) may elucidate HPV's role in breast cancer. Ultimately, functional investigations regarding HPV-16 copy number and the transformation and invasion of mammary epithelial cells in vitro may reveal critical pathways [30–33].

5. CONCLUSION

The present study shows that HPV-16 DNA is found in almost one-third of breast cancer tissues, exceeding the prevalence in non-neoplastic breast tissue samples. Higher HPV-16 viral load was associated with advanced tumour stage, higher histological grade, and metastatic disease involvement of regional lymph nodes. A high viral load was an independent predictor of lymph node positivity on multivariate analysis, supporting the idea that HPV-16 infection may drive breast cancer aggressiveness. In HPV-positive breast cancer patients, measuring HPV-16 viral load may help determine tumour behaviour. These findings need to be confirmed by larger, well-designed multi-center prospective studies that include viral integration analysis and functional validation experiments to better understand HPV-16's role in breast carcinogenesis.

DECLARATIONS

Ethics Approval and Consent to Participate. This study was approved by the Institutional Review Board (IRB) of the College of Medicine, University of Kerbala (Approval No. KMC/IRB-2023/047). Written informed consent was obtained from all participants. All procedures were performed in accordance with the Declaration of Helsinki.

Consent for Publication. Not applicable. No individually identifiable patient data are presented in this manuscript.

Availability of Data and Materials. The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Competing Interests. The authors declare that they have no competing interests.

Funding. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions. 1 conceived and designed the study, performed molecular experiments, analyzed the data, and drafted the manuscript. 2 contributed to histopathological evaluation, specimen selection, and critical revision of the manuscript. 3 contributed to clinical data collection, surgical specimen provision, and manuscript review. All authors read and approved the final manuscript.

Acknowledgements. The authors gratefully acknowledge the staff of the Department of Pathology and the Central Laboratory at Al-Hussein Teaching Hospital for their valuable technical assistance during specimen processing and archival retrieval.

REFERENCES

- Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–249.
- Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2024;74(3):229–263.
- De Villiers EM. Relationship between steroid hormone contraceptives and HPV, cervical intraepithelial neoplasia and cervical carcinoma. *Int J Cancer.* 2003;103(6):705–708.
- Zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer.* 2002;2(5):342–350.
- Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer.* 2010;10(8):550–560.
- Scheffner M, Werness BA, Huibregtse JM, et al. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell.* 1990;63(6):1129–1136.
- Band V, Zajchowski D, Kulesa V, Sager R. Human papilloma virus DNAs immortalize normal human mammary epithelial cells and reduce their growth factor requirements. *Proc Natl Acad Sci U S A.* 1990;87(1):463–467.
- Simoes PW, Medeiros LR, Simoes Pires PD, et al. Prevalence of human papillomavirus in breast cancer: a systematic review. *Int J Gynecol Cancer.* 2012;22(3):343–348.
- Bae JM, Kim EH. Human papillomavirus infection and risk of breast cancer: a meta-analysis of case-control studies. *Infect Agents Cancer.* 2016;11:14.
- Li N, Bi X, Zhang Y, Zhao P, Zheng T, Dai M. Human papillomavirus infection and sporadic breast carcinoma risk: a meta-analysis. *Breast Cancer Res Treat.* 2011;126(2):515–520.
- Lawson JS, Glenn WK, Salyakina D, et al. Human papilloma viruses and breast cancer. *Front Oncol.* 2015;5:277.
- Glenn WK, Heng B, Delprado W, Iacopetta B, Whitaker NJ, Lawson JS. Epstein-Barr virus, human papillomavirus and mouse mammary tumour virus as multiple viruses in breast cancer. *PLoS One.* 2012;7(11):e48788.
- Josefsson AM, Magnusson PK, Ylitalo N, et al. Viral load of human papilloma virus 16 as a determinant for development of cervical carcinoma in situ: a nested case-control study. *Lancet.* 2000;355(9222):2189–2193.
- Flores R, Papenfuss M, Klimecki WT, Giuliano AR. Cross-sectional analysis of oncogenic HPV viral load and cervical intraepithelial neoplasia. *Int J Cancer.* 2006;118(5):1187–1193.
- Sun CA, Liu JF, Wu DM, Nieh S, Yu CP, Chu TY. Viral load of high-risk human papillomavirus in cervical squamous intraepithelial lesions. *Int J Gynaecol Obstet.* 2002;76(1):41–47.
- Moberg M, Gustavsson I, Gyllensten U. Real-time PCR-based system for simultaneous quantification of human papillomavirus types 6, 11, 16, and 18. *J Clin Microbiol.* 2003;41(7):3221–3228.
- Khan NA, Castillo A, Zafar N, et al. Human papillomavirus detected in female breast carcinomas in Pakistan. *Br J Cancer.* 2008;98(5):812–815.
- Heng B, Glenn WK, Ye Y, et al. Human papilloma virus is associated with breast cancer. *Br J Cancer.* 2009;101(8):1345–1350.
- Peitsaro P, Johansson B, Syrjänen S. Integrated human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative real-time PCR technique. *J Clin Microbiol.* 2002;40(3):886–891.
- Haghshenas MR, Mousavi T, Kheradmand M, Afshari M, Moosazadeh M. Human papillomavirus and breast cancer in Iran: a meta-analysis. *Iran J Basic Med Sci.* 2016;19(3):231–237.
- Al-Lawati Z, Khamis AH, Al-Hamdani A, et al. Prevalence of human papillomavirus in breast cancer in Oman. *Gulf J Oncol.* 2020;1(33):59–64.
- Manzouri L, Salehi R, Shariatpanahi S, Rezaian M. Prevalence of human papillomavirus among women with breast cancer since 2005–2009 in Isfahan. *Adv Biomed Res.* 2014;3:75.
- Williams VM, Filippova M, Soto U, Duerksen-Hughes PJ. HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress. *Future Virol.* 2011;6(1):45–57.
- Fernandes JV, De Araujo JM, Fernandes TA. Biology and natural history of human papillomavirus infection. *Open Access J Clin Trials.* 2013;5:1–12.
- Talora C, Sgroi DC, Crum CP, Dotto GP. Specific down-modulation of Notch1 signaling in cervical cancer cells is required for sustained HPV-E6/E7 expression and late steps of malignant transformation. *Genes Dev.* 2002;16(17):2252–2263.
- Cardeal LB, Boccardo E, Termini L, et al. HPV16 oncoproteins induce MMPs/RECK-TIMP-2 imbalance in primary keratinocytes: possible implications in cervical carcinogenesis. *PLoS One.* 2012;7(3):e33585.
- Munger K, Basile JR, Duensing S, et al. Biological activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene.* 2001;20(54):7888–7898.
- Akil N, Yasmeen A, Kassab A, Ghabreau L, Aboussekhra A, Al Moustafa AE. High-risk human papillomavirus infections in breast cancer in Syrian women and their association with Id-1 expression: a tissue microarray study. *Br J Cancer.* 2008;99(3):404–407.
- Yasmeen A, Bismar TA, Kandouz M, Foulkes WD, Heuer T, Al Moustafa AE. E6/E7 of HPV type 16 promotes cell invasion and metastasis of human breast cancer cells. *Cell Cycle.* 2007;6(16):2038–2042.
- De Cremoux P, Thioux M, Lebigot I, Sigal-Zafrani B, Salmon R, Sastre-Garau X. No evidence of human papillomavirus DNA sequences in invasive breast carcinoma. *Breast Cancer Res Treat.* 2008;109(1):55–58.
- Lindel K, Forster A, Altermatt HJ, Suter P, Rieben R. Breast cancer and human papillomavirus infection: no evidence of HPV etiology of breast cancer in a group of Swiss women. *Breast.* 2007;16(2):172–177.
- Salman NA, Davies G, Majidy F, et al. Association of high risk human papillomavirus and breast cancer: a UK based study. *Sci Rep.* 2017;7:43591.
- Khodabandehlou N, Mostafaei S, Etemadi A, et al. Human papillomavirus and breast cancer: the role of inflammation and viral expressed proteins. *BMC Cancer.* 2019;19(1):61.